A mathematical kinetic model of memory in *Bacillus subtilis* spore germination

Chris G. de Koster[a,†], Xiuping Lin[c], Yong-qing Li[c], Stanley Brul[b*], Leo J. de Koning[a] and Peter Setlow[d]

a) Mass Spectrometry of Biomacromolecules and (b) Molecular Biology & Microbial Food Safety, Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, 1098 XH, Amsterdam, The Netherlands.

c) Department of Physics, East Carolina University, Greenville, North Carolina, USA[c].

d) Department of Molecular Biology and Biophysics, UConn Health, Farmington, Connecticut, USA[d].

† In memory of prof. Chris de Koster who passed away 5th August 2019
* Author to whom correspondence should be addressed s.brul@uva.nl

Keywords: Bacillus subtilis, germination, memory, germinosome, model

Abstract

Dormant *Bacillus subtilis* spores germinate through interaction of germinants with germination receptors (GRs). Subsequently, the activated GR signal is transduced to SpoVA proteins that constitute transport channels. SpoVA channel opening then leads to calcium dipicolinic acid (CaDPA) release and completion of spore germination. Spores are known to exhibit memory in germination, as spores given an initial short germinant pulse respond more readily to a second pulse. We have developed a mathematical model to identify a minimal reconstructed molecular network that is crucial for germination kinetics leading to memory of germinant exposure; the model reproduces experimental double germinant pulse germination curves. Analysis of the reconstructed network indicates that a minimal set of inactive and active GRs and a SpoVA channel in three states - closed inactive, closed active and open - is needed to reproduce memory. Spore germination memory is introduced in the

network by the activation and deactivation rates of GRs, and by the interplay between activation of closed SpoVA channels and their rate of opening and closing.

Author summary

Dormant bacterial spores return to life in germination when exposed to germinants which are recognized by germinant receptors (GRs). Spores have memory of germinant exposure, such that more spores germinate after a 2nd germinant pulse than after the 1st one, if intervals between pulses are short. In this work, a mathematical model has been developed that reproduces experimental memory behavior. Computer simulations of the model reveal the crucial factors in spore memory of germinant exposure. Active and inactive states GR states and a membrane channel that exists in closed inactive, closed active and open states are indispensable for spore memory. The model will be part of an iterative cycle of modelling of network dynamics and generating hypotheses for further experimentation.

Introduction

Spores of bacteria of *Bacillus subtilis* are dormant forms which can survive harsh environmental conditions for many years. However, when nutrients become available and conditions are favorable for growth, spores can germinate and outgrow into vegetative cells via a sequence of molecular processes and signaling through pathways and networks (1). In the 1st stage of germination the spore probes its environment for the availability of nutrient germinants. Upon interaction of germinants with germinant receptors (GRs) spores germinate. The binding of germinants, most likely to the GRs, leads to release of the spore's huge depot (~ 10% of spore dry wt) of pyridine-2,6-dicarboxylic acid (DPA) from its core; spore DPA is chelated with divalent metal ions, predominantly Ca^{2+} (CaDPA). CaDPA release then triggers the activation of two cortex-lytic enzymes (CLEs), CwlJ and SleB. In the 2nd stage of germination, these CLEs hydrolyze the large spore cortex peptidoglycan (PG) layer. Cortex PG degradation then allows core expansion with full hydration of the spore core which allows resumption of normal metabolism, macromolecular synthesis, and finally outgrowth into a vegetative growing cell.

In *B. subtilis* the GR protein assemblies GerA, GerB and GerK are each encoded by homologous tricistronic operons *gerA*, *gerB* and *gerK* (2-4). GR subunits A and B are integral membrane proteins that are embedded in the spores' inner membrane (IM) through multiple predicted α-helices. Subunit C is a peripheral membrane protein, attached to the IM by a diacylglycerol anchor (4-6). The GerA GR responds to L-alanine or L-valine, the GerB and GerK GRs cooperatively respond to a mixture of L-asparagine, D-glucose, D-fructose and K^+. A GerB variant, GerB*, germinates in response to L-asparagine alone. Average wild-type GR spore copy numbers have been determined by Western blotting and are ~2,500 total GRs/spore, with GerA at ~1,100 molecules/spore; and GerB and GerK each at ~700 molecules/spore (7). As expected, GR levels in spores affect the germination rate, and GR overexpression increases rates of spore germination by these GRs (8). All GRs co-localize in the spores' IM in a cluster termed the germinosome (9). The GerD protein co-localizes with GRs in the germinosome and is essential for germinosome formation. GerD is also essential for normal GR-dependent spore germination, as *gerD* spores germinate extremely poor with L-valine via GerA and AFGK via GerB plus GerK. Like GR C subunits, GerD is also likely attached to the IM through an N-

terminal diacylglycerol anchor attached to an N-terminal cysteine residue. GerD presumably acts as a scaffold protein for germinosome assembly (10).

The signal generated upon GR activation is transduced to SpoVA proteins that make up a channel for CaDPA in the spore IM. However, the signaling mechanism is not fully understood, although no proteins that act in between the GRs and the SpoVA channel protein complex have been identified. The *Bacillus subtilis spoVA* operon encodes SpoVAA, SpoVAB, SpoVAC, SpoVAD, SpoVAEb, SpoVAEa and SpoVAF proteins. SpoVAA, SpoVAB, SpoVAC, SpoVAEb and SpoVAF are predicted to be integral membrane proteins based on their primary sequences, and SpoVAD and SpoVAEa are hydrophilic, soluble proteins that are on the outer IM surface (11-13). The copy numbers of SpoVAD and SpoVAEa have been determined by western blotting and are ~6,000 - 6,500 and 750 molecules/spore, respectively (7,11) SpoVAC alone functions as a mechanosensitive channel for small molecules such as CaDPA, with several treatments including membrane tension opening this channel in vivo and in vitro (13). While the arrangement of SpoVA proteins in a supramolecular structure is not known, it seems most likely that the SpoVA proteins are subunits of a protein assembly in the spore's IM that constitutes a channel for CaDPA movement into the developing spore during sporulation and out of the germinating spore (11). In contrast to GRs which are in the multiprotein germinosome, SpoVA proteins are more evenly distributed over the spore IM (9). Opening of the CaDPA channel is triggered not only by GR activation, but also in a GR-independent fashion by dodecylamine or hydrostatic pressures of ≥ 4000 atmospheres (14-17). Notably, CaDPA release from the spore core in germination is paralleled by water uptake into the core.

In the 2nd germination stage, CaDPA activates the CLEs which hydrolyze the cortex PG. CwlJ is activated first by CaDPA released from the spore core and SleB is activated only after substantial CaDPA release, perhaps by changes in cortex PG structure accompanying CaDPA release and water uptake. CwlJ can also be activated by exogenous CaDPA. The structure of the dormant spore cortex PG and physical effects due to elastic and contractile properties of the cortical PG sacculus presumably affect the susceptibility of opening of SpoVA channels (18). Cortex hydrolysis degrades the cortex PG macromolecular network restricting expansion of the spore core, thus enabling further uptake of water.

As core water content increases, eventually full core hydration allows onset of metabolic reactions including biosynthesis of molecules needed for cell growth.

Spore germination is a heterogeneous process, as in populations some spores germinate rapidly and some extremely slowly (19). This heterogeneity is due primarily to a variable time period between the application of a germinant stimulus and the onset of rapid CaDPA release from the spore core. This time interval is denoted as the lag time T_{lag}; all CaDPA is released by $T_{release}$, and cortex PG hydrolysis and core swelling are complete by T_{lysis}. Spores that are exposed to germinant molecules for a short time interval can be committed to germination at time T_c even when the germinants are removed and further binding to GRs is blocked (17,20,21); slow leakage of 15-25% of CaDPA begins at or close to the T_c and this slow release ends at T_{lag}, when rapid CaDPA release begins (22). T_{lag} values are decreased by spore heat activation, higher GR levels and higher levels of GR-dependent germinants up to saturation (19). Spores with ~4 fold higher SpoVA copy numbers have lower T_{lag} values in GR germination than those for wild type spores (15). The times needed for rapid release of all CaDPA remaining at T_{lag} ($\Delta T_{release} = T_{release}-T_{lag}$) (0.5-3 min), and for cortex hydrolysis and core swelling ($\Delta T_{lysis} = T_{lysis} - T_{release}$) (10-20 min) are generally shorter than T_{lag}. GR levels and nutrient germinant concentrations do not appreciably affect $\Delta T_{release}$ or ΔT_{lysis} values (19,23). However, loss of the CwlJ CLE, but not SleB, markedly decreases rates of CaDPA release during spore germination (24,25).

Recently, spores were found to exhibit memory of prior germinant exposures. Thus spores given a brief exposure to a germinant that activates the GerA or GerB* GRs responded more readily to a 2nd short germinant pulse (21,26). However, the response to the 1st germinant pulse diminished when the time interval in between pulses was increased, indicating that there is memory in the germination system. It was concluded that GRs and the SpoVA channel proteins are the most likely candidates for germination memory generation and storage (26), but mechanisms underlying spore germination memory are not clear.

To begin to address possible mechanisms of germination memory, in the current work we have attempted to build a mathematical mass action kinetics germination model to identify a minimal molecular network and reactions thereof that are crucial for spore germination memory. Obviously, this model must be able to predict both the heterogeneity of T_{lag} and the percentages of spore germination in 1st and 2nd germinant pulses. In addition, with computer simulations we can identify network topologies that exhibit minimal deviation between predicted and experimental data, and reveal reaction dynamics of molecular processes between GR activation and rapid CaDPA release. Finally, analysis of the optimized model could shed light on the question of how the network's molecules and their interactions determine its biological function.

Results

Spore germination model.

In our modelling strategy, we follow a top-down approach. First we develop an overall mathematical model and then we dissect the mathematics into a set of mass action kinetics based differential equations based on the components of the germinosome and the SpoVA channel which must somehow interact. As a boundary condition we opted for the most parsimonious number of building blocks, and that the model must be able to recreate published experimental spore memory data (26). One common premise for commitment in molecular networks is a bistable switch (27). In such bistable systems two stable states coexist, and thus spores are dormant or are committed to germination (28). Upon activation, the system is irreversibly committed to one state. A condition for bistability is non-linear dynamics involving positive feedback in which a molecular species upregulates its production or autoactivates its function, *e.g.* channel opening, a mechanism to filter out noise and a mechanism to prevent numeric explosion of the upregulated or activated species (28,29). A simple deterministic molecular kinetic model for the autoactivation and deactivation rates of the state variable X commonly used to model bistability is given in Eq (1) (30,31).

$$\frac{dX}{dt} = P + k_s \frac{X^n}{X^n + \theta^n} - k_d X \qquad (1)$$

In this equation, X denotes the concentration [molecules/spore] of a crucial state of a germination protein, k_s is its maximum production rate, n represents cooperativity of variable X, θ^n sets the value at which the rate of production of X is half its maximum value, k_d is the decay rate of protein X and P is a basal production rate (30-32). Positive feedback is introduced by cooperative action of the system variable X for $n > 1$. The multistabilty of Eq (1) depends on the settings of the various parameters. There are some settings in which X gradually increases upon an inductive stimulus and there is a parameter subspace where X exhibits a bistable response. In the latter case, there are two stable states, one at low X and one at high X, and an unstable state in between. Fluctuations of X mediated by P around the lower stable state (state 1) but not progressing to the metastable state fall back to the lower stable state. Fluctuations of X through P surpassing the unstable state are auto-activated and increase to the stable state 2 at high X. Fluctuations in X above the high stable state fall back to stable state 2, thereby preventing a runaway of variable X to infinity. Eq (1) is thus an autocatalytic off-on switch with parameter subspace for bistability and is the core of our model.

Starting from this equation we designed an input section, signal transduction and finally we describe SpoVA channel opening kinetics to estimate T_{lag} times in germination. The molecular wiring of the reconstructed molecular network of the model is given in Fig 1. In our model, we do not differentiate between GerA, GerB and GerK GRs, and thus do not differentiate between germinants. GRs can be in an inactive R_i state or an active R_a state, and germinant S binding converts an R_i GR to R_a with a rate constant k_1. Low copy protein numbers in individual cells in bacterial populations usually vary according to a gamma distribution (33). Measurements of levels of GFP-labelled GRs in spores have shown that GR copy numbers also vary according to a gamma distribution function (33,34). Thus, the distribution of total levels of R_i in spores in the absence of germinant is given by the probability density (gamma) function

$$f(R_i; a, b) = R_i^{a-1} \frac{e^{-R_i/b}}{b^a \Gamma(a)} \tag{2}$$

in which $\Gamma(a) = (a-1)!$, a is the shape parameter and b is an inverse scale parameter. The distribution mean $E(R_i) = ab$ and the variance $Var(R_i)$ of the distributions is ab^2. The skewness of the distribution is equal to $2/a^{½}$. The skewness depends only on the shape parameter a and approaches a normal distribution when approximately $a > 10$.

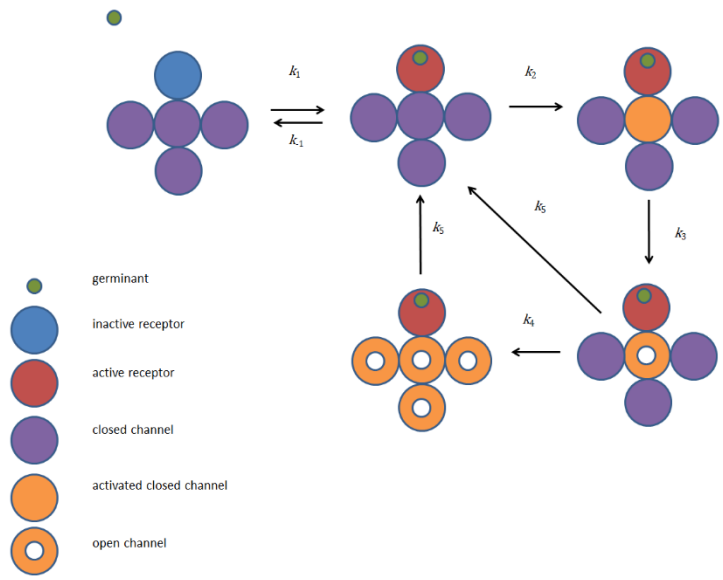

Fig 1 Architecture of a *Bacillus subtilis* germinosome that exhibits spore germination memory.

The molecular network is composed of germinant receptor proteins that are activated by a germinant, and channel proteins in a closed, activated closed and an open state and their mutual interactions.

Germinant binding in our model is reversible and dissociation takes place with rate constant k_{-1}. Eq (3) and Eq (4) describe rates of formation of active GR molecules R_a and rates of depletion of the inactive GR population R_i. In contrast to GRs which we propose exist in only two states, R_i and R_a, we need a SpoVA channel that can adopt three states, although there is no experimental proof for this assumption. Our model with two SpoVA states does not reproduce experimental germination curves. The SpoVA channel C_c is closed and inactive, the channel C_{ca} is closed and in an activated state, or the channel C_o is open and in an activated state. Interaction of activated GRs R_a with closed SpoVA channels C_c activates the latter (Eq (5)). We assume a direct interaction, although there is no conclusive experimental data about signaling from activated GRs to the SpoVA channel. However, yeast two hybrid experiments suggest that there is interaction between *B. subtilis* GerA proteins and several SpoVA proteins (35). The function of GerD is ignored in the model. A premise for commitment in molecular networks is a bistable switch with delay. Bistability is introduced here through intersection of a production function comprising a Hill function with $n > 1$ and a linear decay function in Eq (7). The production function describes the rate of opening of SpoVA channels. Cooperativity of open SpoVA channels also contributes to the opening rate. This means that open SpoVA channels (C_o) have multiple mutual interactions which enhance the opening rate of closed channels. Cooperativity of SpoVA channel gating is introduced through a Hill function with apparent dissociation constant θ and coefficient n. The decay function describes closure of channels with reaction rate $-k_5C_o$. Channel closure is first order in the concentration of open channels. Upon closure the closed channel falls back to the inactive ground state C_c.

$$\frac{dR_i}{dt} = -k_1SR_i + k_{-1}R_a \tag{3}$$

$$\frac{dR_a}{dt} = k_1SR_i - k_{-1}R_a \tag{4}$$

$$\frac{dC_c}{dt} = -k_2 R_a C_c - \frac{k_4 C_o^n}{C_o^n + \theta^n} + k_5 C_o \tag{5}$$

$$\frac{dC_{ca}}{dt} = k_2 R_a C_c - k_3 C_{ca} \tag{6}$$

$$\frac{dC_o}{dt} = k_3 C_{ca} + \frac{k_4 C_o^n}{C_o^n + \theta^n} - k_5 C_o \tag{7}$$

In these equations, S is germinant concentration, R_i inactive GR, R_a active GR, C_c closed SpoVA channel, C_{ca} activated closed SpoVA channel, and C_o open SpoVA channel. The units of the GR and SpoVA variables are numbers of the various complexes/spore. Constants k_1, k_{-1}, k_2, k_3, k_4 and k_5 are rate constants, n is a Hill coefficient, and θ^n sets the value at which the production rate in the second term is half its maximum value. The final stage of the model is SpoVA channel opening. Whether core CaDPA is exported by concentration differences and/or an electrochemical gradient of ions inside and outside the inner membrane, CaDPA efflux will be proportional to the number of open SpoVA channels and we assume that T_{lag} is determined by a minimal number of open SpoVA channels needed for rapid CaDPA release. It is known that CaDPA activates the CLE CwlJ and this enzyme's action stimulates CaDPA efflux. However, CwlJ activation and its effect on the SpoVA channel opening rate is not yet a part of the model, and we expect that this effect will become significant only after T_{lag}. If there is a contribution of CwlJ activation to SpoVA opening than this will lead to a minor shift of T_{lag}. In our computer model, we use a cycle of 500 numeric simulations to solve the set of coupled ordinary differential equations. In each simulation, we predict the number of active open SpoVA channels as a function of time and we determine for each simulation the time of opening, T_{open}, of SpoVA channels. This time point corresponds with the time point of the beginning of fast CaDPA release (19), and is related to T_{lag} through $T_{\text{lag}} = T_{open} - T_{\text{block}}$, where T_{block} is the time of the end of a germinant pulse. We

define T_{open} as the time point in a simulation where C_o exceeds 50 copies/spore. This number is arbitrarily selected as experimental data are lacking.

Modeling of spore memory of germinant stimuli

First, we reproduce with our mathematical model an earlier published percentage germination curve showing spore memory for stimulation of wild-type *B. subtilis* spores with L-valine (26). Next, we study the relationship between the mean $E(R_i)$ of the GR distribution and the percentage germination upon the 1st and the 2nd germinant pulses. We then compare the predictions with experimental results of the percentage germination of wild-type spores and mutant spores with elevated GR levels. Wild-type spores can acquire memory upon short pulses of L-valine to induce germination as follows. After the 1st pulse a fraction of the spores is committed to germinate and does so. From the subpopulation that does not germinate after the 1st pulse, a fraction germinates upon a 2nd pulse and finally a fraction does not germinate at all. Interestingly, spore populations can exhibit 2- to 3-fold more germination in the 2nd germinant pulse following the 1st one. Apparently, there is memory in the molecular system for external perturbation such as a germinant pulse that occurred before the 2nd pulse (26). Consequently, we optimize model parameters to reproduce the spore germination curve for an experiment with two short germinant pulses (26). In our simulation, we select two pulses of germinant exposure from $0 < t < 5$ min and $30 < t < 35$ min. We set the germinant concentration to 3.5 mM. A percentage germination curve in our definition is a cumulative distribution function of T_{open} to give rapid CaDPA release. Spores that do not germinate have an infinite T_{open} and do not contribute to the distribution in the time frame of the simulations. Heterogeneity of T_{open} and as result T_{lag} determines the shape of the curve. Here, we introduce heterogeneity through a gamma distribution of GR copy numbers (Eq (2)). For each simulation of a cycle of 500 we randomly select a GR copy number from a gamma distribution, but keep SpoVA copy numbers constant. This means that $C_c + C_{ca} + C_o$ is constant in each cycle of a set of simulations. This approach is justified by the experimental observation that overexpression or deletion of GR's has no effect on levels of SpoVAD and by inference levels of all proteins encoded by the

heptacistronic *spoVA* operon (8). From Western blotting, it is known that on average ~700 to ~1,100 of each GR are present in a spores' germinosome. The starting value of at least SpoVAD is 6,500 as determined earlier by Western Blotting (12). By tuning of k_4, the SpoVA copy number of activated and open channels in our simulations asymptotically approaches 6,500 copies/spore (*vide infra*). This copy number is also an arbitrary selection as we do not know by experimentation the number of open SpoVA channels at the end of rapid CaDPA release. At the end of the simulations we construct a predicted percentage germination curve. The experimental germination curve was determined by monitoring germinating spores with DIC microscopy. However, most of the intensity change seen by DIC microscopy are caused by rapid CaDPA release, and the kinetic parameters for which averages were determined are T_{lag}, T_{release} and $\Delta T_{\mathrm{release}} = T_{\mathrm{release}}-T_{\mathrm{lag}}$ (19,26,36).

We start with varying the mean $E(R_{\mathrm{i}})$ of the GR distribution and keep the standard deviation $Var^{½}(\mathrm{R})$ fixed. The parameters of the randomly accessed gamma distribution are estimated at the end of a set of 500 cycles from the fitted gamma distribution. A typical GR protein distribution is depicted in Fig 2A. The average of this GR copy number distribution is set to 1100 copies/spore and we selected combinations of distribution parameters a and b with the condition that $Var^{½}(\mathrm{x}) = a^{½}b = 220$. With this setting the GR distribution tails to higher copy numbers, and a fraction of the spores germinate after the 1st pulse and a fraction germinates after the 2nd germinant pulse (Fig 2A, B). Notably, with the 1st pulse 15% of the spore population germinates, while after the 2nd pulse 68% of the remaining spores germinate. Although the germinant concentration and pulse lengths are equal for both pulses, germination after the 2nd pulse is approximately four-fold higher than after the 1st pulse. Thus, the model reproduces published experimental data (26). Simulations for other $E(R_{\mathrm{i}})$ are compiled in S1 Figures.

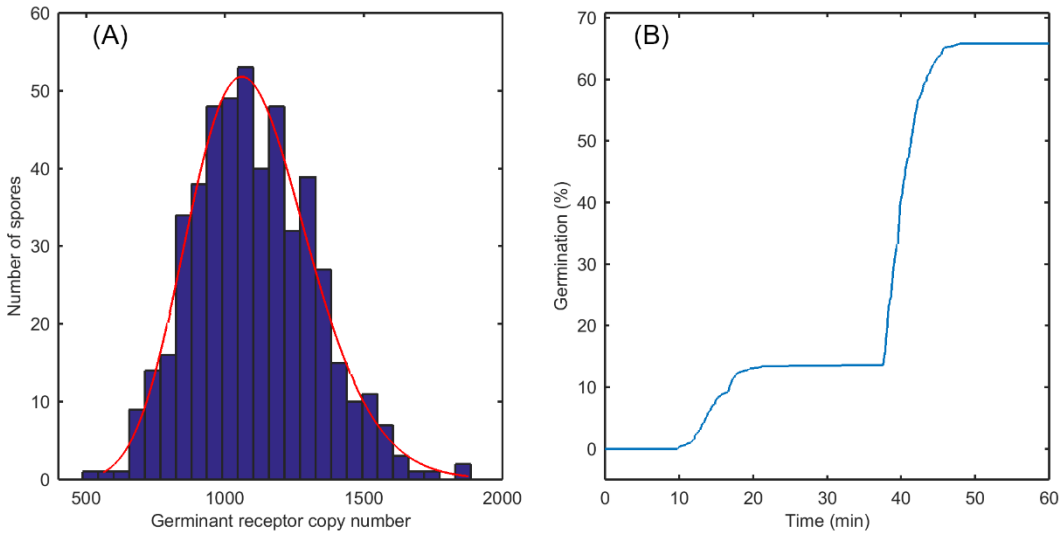

Fig 2 Simulation of spore germination memory

(A) Gamma distribution of GR copy numbers used for simulation of the percentage germination efficiency curve (Fig.2b) with $E(R_i)$ = 1100 at t is 0 min. The red line is the fitted distribution. The GR estimated mean is 1105 molecules/spore, and the estimated standard deviation is 219 molecules/spore; the parameters of the fitted gamma distribution are: a is 25.4491 [22.5002, 28.7844] and b is 43.4004 [38.3247, 49.1484]; Right panel: Percentage germination of spores in a two pulse germination experiment.

(B) Percentage germination in a two pulse germination experiment. After the 1st pulse the curve levels of to an asymptote at 15%. After the 2nd pulse the curve approaches 68%. Two pulses of 3.5 mM germinant for receptor R_i activation are administered from 0 < t < 5 and 30 < t < 35 min. At zero time R_a = 0 copies/spore; C_c = 6500 copies/spore. $C_a = C_o$ = 0 copies/spore; k_1 = 14e-3; k_{-1} = 100e-3; k_2 = 25e-7; k_3 = 60e-3; k_4 = 195e2; k_5 = 305e-2; n = 3; θ = 20.

We repeated a double pulse germination experiment with *B. subtilis* wild-type spores and spores of an isogenic strain that overexpresses the GerA GR ~8-fold as determined by Western Blotting, with an experimental error of 25% (7). The experimental percentage germination curves are depicted in Fig 3A, B. We simulated the curves with our model (Fig 3C, D). Compared to germination of wild-type spores we had to adjust reaction rate constants for a 2 min pulse for wild-type spores that overexpresses GerA with 10 mM L-valine as germinant. A broad GR copy number distribution, with $E(R_i)$ is 1100, a is 1 and b is 1100, reproduces approximately 4% germination upon the first pulse and approximately 6% germination at the second pulse (Fig 3C). Next, we overexpress the GerA GR by increasing $E(R_i)$ and keep the reaction rate constants similar. In this approach, we put forward the constraint that overexpression of GR GerA from a more active promotor does not affect the rate constants for binding and dissociation of the germinant to the receptor. A second assumption is that expression of GerA from another promotor affects the parameters of the gamma distribution. It is known from modelling of burst like stochastic gene expression that the rate at which the gene transitions from the inactive state to the active state and vice versa determines the momenta of the mRNA copy number distribution and consequently the protein copy number distribution (37). Our modelling results are in accord with an earlier prediction using a stochastic gene expression model that a high burst frequency produces a narrow mRNA copy number distribution and a low burst frequency a broad mRNA distribution (37). However, experimental data for the distribution of copy numbers of an overexpressed GFP-labelled GerA in individual spores are lacking. With $E(R_i)$ at 3350 a is 67 and b is 50 we reproduce 3% germination at the first pulse and 64% germination at the second pulse (Fig 3D). In this case, we had to adjust the variance of the GR distribution to reproduce the percentage germination curve. For $E(R_i)$ is 3350 the overexpression is approximately 3.3-fold. For further increase of $E(R_i)$ we have to diminish k_1 to keep the second pulse at 63% without surpassing the 3% germination at the first pulse. These results validate simulations of the relation between GR protein distributions and percentage germination curves in a double pulsed memory experiment. Interestingly, the model predicts constant $\Delta T_{release}$ times and that $\Delta T_{release}$ is independent of T_{lag}, corroborating observations from earlier experimental work (19). Next, we probed for the effect of varying the concentration of germinants used during the pulses, from

2 mM to 5mM (S2 Figures). At 2 mM germinant no germination is predicted with the 1st pulse and approximately 1% after the 2nd pulse. This latter percentage increases to 12% at 2.5 mM germinant, but still with no germination after the 1st pulse. This implies that the 1st pulse primes 12% of the spore population for germination after the 2nd pulse. For higher germinant concentrations, significant germination with the 1st pulse as well as substantial memory effects are predicted, *e.g.* S2 Fig 2C. The influence of the standard deviation of the distribution is further highlighted in S3 Figures. There is an obvious relationship between the variance of the distribution and the fraction of spores germinating after the 1st and 2nd pulses (S3 Figures). As the variance of the distribution increases and tails to higher GR copy numbers, more spores germinate with the 1st germinant pulse.

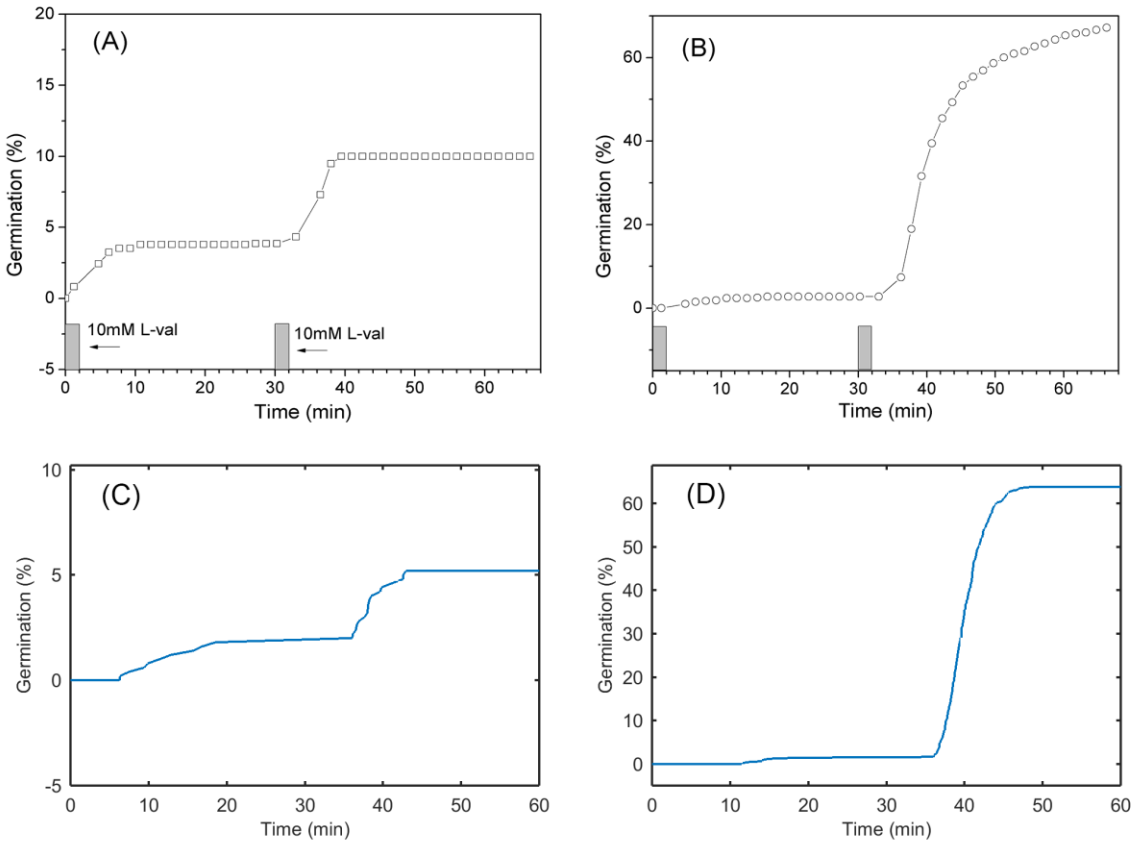

Fig 3 A double pulse germination experiment with *B. subtilis* wild-type spores and spores of an isogenic strain that overexpresses the GerA GR.

(A) Wild-type *B. subtilis* spores (PS533) were given two 2-min 10 mM L-valine (L-val) germinant pulses separated by 30 min.

(B) Spores of *B. subtilis* PS3476 (GerA GR over-expressed ~ 8-fold) were given two 2-min 10 mM L-valine germinant pulses separated by 30 min. Gray bars above the horizontal axis indicate pulse durations.

(C) Simulated percentage germination of WT spores in a two pulse germination experiment. The estimated distribution mean is 1118 and the standard deviation is 1064 with a is 1.10476 [0.989585, 1.23334] and b is 1012.19 [881.615, 1162.1]. At zero time R_a = 0 copies/spore; C_c = 6500 copies/spore. $C_a = C_o$ = 0 copies/spore; k_1 = 350e-5; k_{-1} = 100e-3; k_2 = 25e-7; k_3 = 60e-3; k_4 = 195e2; k_5 = 305e-2; n = 3; θ = 20.

(D) Simulated percentage germination of overexpressed GR spores in a two pulse germination experiment. The estimated distribution mean is 3367 and the standard deviation is 399 with a is 71.2149 [62.9306, 80.5898] and b is 47.2855 [41.7667, 53.5336]. At zero time R_a = 0 copies/spore; C_c = 6500 copies/spore. $C_a = C_o$ = 0 copies/spore; k_1 = 350e-5; k_{-1} = 100e-3; k_2 = 25e-7; k_3 = 60e-3; k_4 = 195e2; k_5 = 305e-2; n = 3; θ = 20.

Dynamics of SpoVA channel opening

The kinetics of SpoVA channel opening are studied in more detail for GR copy numbers ranging from 600 to 1,800 molecules/spore. The increment of GRs is 50 copies per cycle of numeric evaluation of the set of ordinary differential equations (ODE's) in a simulation. The number of GRs is defined by R_i at time point t is 0 seconds. The results C_o and the copy number time profiles of the state variables R_i, R_a, C_c, and C_{ca} are given in Fig 4. There are three subpopulations in the uniform GR distribution. With 600 to 950 GRs/spore, there is no germination, with 1000 to 1350 GRs/spore there is germination after the 2nd pulse and with 1400 to 1800 GRs/spore there is germination after the 1st and 2nd pulses. Fig 4 shows that only in a fraction of the spore population are GRs sufficiently activated by germinant binding

to lead to commitment to germinate. We can tune this fraction by variation of the rate constants k_1 and k_{-1} for binding of S to GR and dissociation thereof. This tuning also affects the number of spores opening SpoVA channels during the 1st and 2nd pulses. With the parameter settings of our simulation only a few SpoVA channels are activated by interaction with GRs. Because of different copy numbers between GRs and channels there cannot be a one to one stoichiometry between GR activation and channel opening. Approximately 1,100 GerA GRs are in the germinosome, while ~6,500 SpoVA channels are dispersed over the IM. The surface area of the spore IM can be approximated by an ellipsoid. The surface area of an ellipsoid (A_{im}) is calculated as $A_{im} = 2\pi(s^2+ls)$, where l is the length of the long axis and s of the short axis (38). The long axis l and short axis s of the IM are estimated by electron microscopy and l is 837 nm and s is 554 nm for a typical spore (39). These are obtained by subtraction of coat and cortex layers' thickness from the spore long axis and short axis (39). The surface area of the IM is A_{im} = ~ 4,835,800 nm^2, and on average there will be one SpoVA molecule per 744 nm^2 of IM. Next, we calculate the area of the germinosome. As input we use the normalized average distribution of the fluorescence signal along the long axis of a spores. The full width at half max intensity for GerKB-GFP signal is 300 nm (9). Assuming a planar circle on the surface of the IM for the contour of the germinosome its surface $A_{ger} = \pi r^2$ = 502,655 nm^2. Assuming a homogeneous distribution of 744 nm^2 per SpoVAE then the germinosome contains at most 676 SpoVA molecules. In a stoichiometry where one GR activates one channel maximally, 1,100 channels can be activated. Thus, there are ~ 5,800 SpoVA channels that cannot directly interact with a GR. Thus, SpoVA channel cooperativity is introduced in the model to relay the signal from GRs over the entire IM surface to open more than approximately 676 channels. In our model, closed SpoVA channels are directly opened by activated and open SpoVA channels. Immediately after the 1st germinant pulse the number of open SpoVA channels gradually increases. This slow increase is the gross difference between opening by activated SpoVA channels through k_4C_a and closing of SpoVA channels through decay $-k_5C_o$ (Eq (7)). Beyond 10 min sufficient channel opening takes place to force cooperative opening of the dark blue subpopulation through the Hill function term in Eq (5). Channel opening is autocatalytically driven, all SpoVA channels open, and hence the C_o curve approaches 6,500 molecules. The SpoVA channels from

spores that germinate after the 2nd pulse are partially activated by the 1st pulse but remain below a critical copy number needed to force an autocatalytic opening mediated by the Hill function in Equation 6. However, the extent of channel opening does not fully return to the baseline. A 2nd pulse at $30 < t < 35$ min contributes additionally to channel opening and drives through cooperative action channel opening to completion at 6,500 copies/spore. As a result, channels open that are not opened by the 1st germinant pulse. Thus, the system exhibits memory for an earlier germination stimulus. Finally, the curve levels off to a plateau at 6500 open SpoVA channels. The level of the plateau of open SpoVA channels is mainly determined by k_4 and k_5. A third fraction does not open at all, as the 2nd germinant pulse is insufficient to trigger autocatalytic SpoVA opening and the number of open channels decays to the baseline.

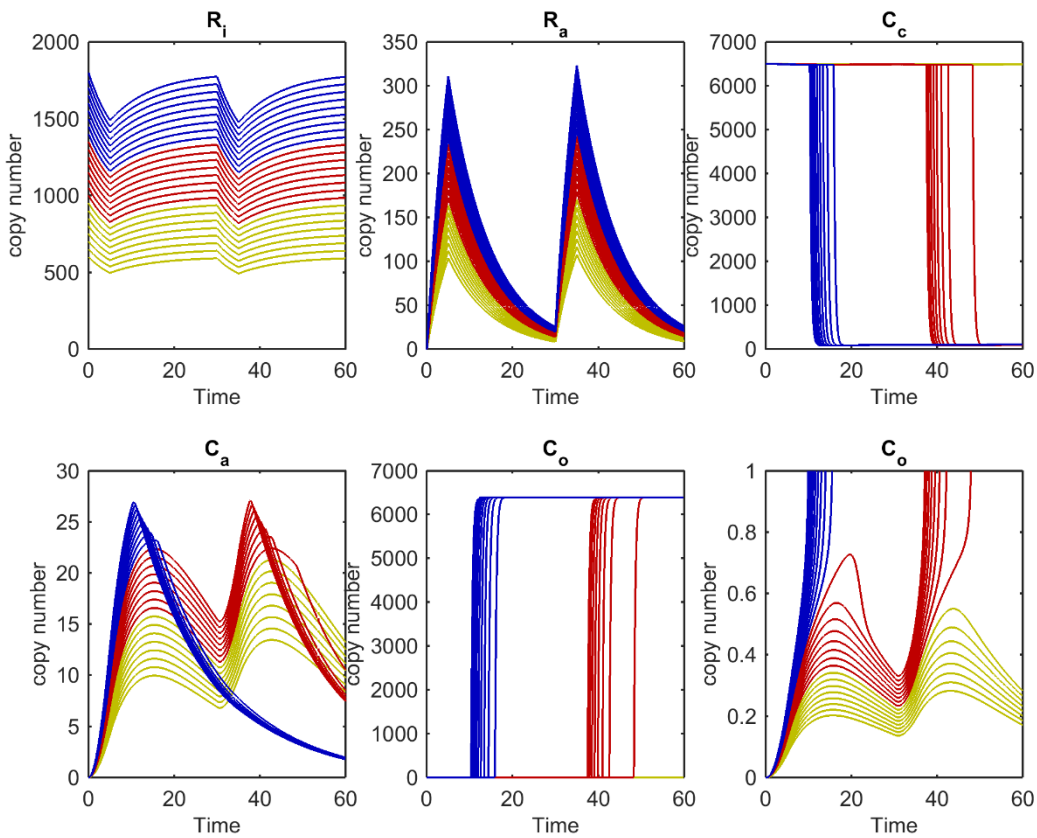

Fig 4 The copy numbers of R_i, R_a, C_c, C_a, C_o and an extension of C_o as a function of time.

Each trace represents a state variable profile for a variable GR R_i copy number. The starting GR values are $600 \leq R_i \leq 1800$ copies/spore and the increment is $\Delta R_i = 50$ copies/spore. The yellow population ($600 \leq R_i \leq 950$) do not germinate, the red one ($1000 \leq R_i \leq 1350$) germinates at the 2nd pulse and the blue ones ($1400 \leq R_i \leq 1600$) germinate at the first pulse. Two pulses of 3.5 mM germinant for R_i activation are administered from $0 < t < 5$ and $30 < t < 35$ min. At time zero, $R_a = 0$ copies/spore; $C_c = 6500$ copies/spore. $C_a = C_o = 0$ copies/spore; k_1 = 14e-3; k_{-1} = 100e-3; k_2 = 25e-7; k_3 = 60e-3; k_4 = 195e2; k_5 = 305e-2; $n = 3$; $\theta = 20$.

Discussion

Spore germination experiments with sequential germinant stimuli separated in time are very suitable to construct and optimize a kinetic germination model. The shape of the percentage germination curves with multiple inflection points and asymptotes enables a more robust fitting of model parameters than a single stimulus sigma-shaped germination response experiment. The germination model based on the minimal molecular network as presented in Fig 1 and described by differential Eq (3-7) reproduces percentage germination curves in double pulse memory experiments (26) and newly acquired data (Fig 3A, B). Upon germinant exposure in two 5 min pulses separated by thirty min, two subpopulations of the whole population germinate. The second one exhibits a memory effect, in that the amounts of spores that germinate after the 2nd pulse are larger than those germinating after the first pulse. Memory has been attributed to a conformational change in some germination protein, that activates this protein for germination, although the activated protein can also decay to the inactivate ground state. When the activated protein is not fully decayed, then a 2nd germinant pulse leads to a larger population of activated proteins and a higher percentage of germination after the 2nd pulse (26). The reconstructed network shows that a minimal set of an inactive and active GR and a transport channel in three states - closed inactive, closed active and open is needed to reproduce memory. Thus, memory can reside with GRs and/or be stored in the CaDPA channel complex. This is consistent with generation of memory in spores

that lack all GRs when germination is triggered by direct or indirect stimulation of CaDPA release by exogenous CaDPA itself that activates CwlJ, or by dodecylamine that may cause the SpoVA channels to open (26).

Memory is stored in an activated GR R_a, although this can decay to the inactive state R_i. As is seen in Fig 4, GR inactivation does not fully return to the baseline before the 2nd germinant pulse and thus GR inactivation characteristics contribute to memory. The fraction of activated GRs and its relaxation to an inactive state is determined in our model by the ratio k_1/k_{-1}, although we do not know this ratio from experimental data. However, memory in our model is not uniquely confined to GR activation and inactivation kinetics, as there is a model parameter setting with $k_{-1} = 225.10^{-3}$ and $k_2 = 45.10^{-7}$ where R_a returns to baseline before the 2nd germinant pulse and the network still shows memory for germination without adjusting the other rate constants. In this case memory is introduced in the network by the interplay between activation of closed channels and their rate of opening to C_o. The spore channel activation rate is proportional to k_2R_a and the number of closed inactive SpoVA channels C_c (Equation 6), and the rate of opening is proportional to k_3C_a. We assume that the number of activated GRs at 3.5 mM is approximately 17% and we have optimized k_2 accordingly. This needs further experimentation to dissect k_2R_a into k_2 and R_a. Rate constant k_3 has an effect on T_{lag} because it feeds the rate of open channel formation (Eq (7)). The Hill coefficient n, θ together with k_4 and k_5 determine the bistability of the model needed to irreversibly open CaDPA channels. In this simulation n is 3 pointing to cooperativity effects of open activated SpoVA channels (Eq (7)). Furthermore, rate constants k_4 and k_5 in combination with θ determine the number of open SpoVA channels at the end-point of simulations. This number is not known by experimentation and is arbitrarily set.

The time to the onset of rapid CaDPA release T_{lag} is heterogeneous and shows a probability distribution. We introduce germination time heterogeneity through the GR distribution in spore population. In a cycle of computer simulations, we keep the rate constants fixed. Rate constants k_1 and k_{-1} are on-rate and off-rate constants and determined by the 3-D structure of the GR proteins. As the primary sequence of the GR proteins is not variable within an isogenic strain, we expect that a k_1 and k_{-1} probability distribution does not contribute to heterogeneity of T_{lag} at the population level. At present, our model is

a minimal one, and we do not differentiate between actions of GerA, GerB and GerK GRs. Consequently, we do not integrate the signal from different GRs and germinants. We also do not know the fraction of SpoVA channels in an inactive state that is activated by a population of activated GRs. The network in our model is capable of signal amplification, as with current parameter settings approximately 1-27 closed SpoVA channels are activated (Fig 4). This is sufficient to irreversibly open the entire channel population through cooperative opening which is needed to open channels outside the perimeter of the germinosome. Another benefit from the network wherein only fractions of molecular populations are active is that it dampens fluctuations in germinant exposure. This ensures that there is a threshold for GR activation that will prevent unwanted early germination.

The model generated in this work has been optimized to reproduce memory of germinant exposures. Other combinations of reaction rate constants will likely also give kinetics that show memory, but the molecular wiring and the kinetic differential equations will likely remain identical. However, further, experimentation is needed to narrow down the parameter space. In particular, the contributions of GR activation and deactivation, and SpoVA channel activation, deactivation and opening should be studied. The endpoint of germination in our model is the rapid opening of SpoVA channels. In a future extension of the model, the effects of the CLEs SleB and CwlJ should also be included. In addition, as noted above GR-independent germination by CaDPA or dodecylamine also show memory, presumably through effects on the SpoVA channel alone, and this can also be introduced into the model. The model in its current state will be used in future research to reduce the number of experimental variables in an iterative cycle of prediction of the dynamics of a hypothetical molecular network, the design of an experiment, comparison with the outcome, and further tuning of the model until agreement between theoretical predictions and experimental data is reached.

Materials and methods

Simulations were carried out using the MATLAB package of software (the Mathworks, Natick, USA). M-files with source code for simulation of percentage germination curves and copy number time profile plots are given in S Matlab m-file one and S Matlab m-file two.

Phase contrast microscopy.

Spores (1 μl; ~10^8 spores/ml in water) were spread on the surface of a microscope coverslip that was then dried in a vacuum desiccator for 5 min. Coverslips were suspended in water and then mounted on and sealed to a microscope sample holder kept at a constant temperature of 37°C. A phase contrast microscope was used to record the images of ~500 individual spores at a rate of 15 s per frame. Spores that contain DPA appear as bright and spores that do not contain DPA appear as dark. Percentage of spores containing DPA was defined as the number of bright spores divided by the total number of spores.

Spore germination with germinant pulses.

B. subtilis spores of strains PS533 (wild-type) and PS3476 (overexpresses the GerA GR ~ 8-fold) were prepared, purified and stored as described previously (7). The purified spores were free (>98%) of sporulating or growing cells or germinated spores as observed by phase contrast microscopy. *B. subtilis* spores were heat activated prior to nutrient germination by incubation in water at 70°C for 30 min and then by cooling on ice for at least 15 min. *B. subtilis* spores attached on the surface of a glass coverslip were germinated in 25 mM K-HEPES buffer (pH 7.4) with a pulse of a germinant (10 mM L-valine) for 2 or 4 min before the germinant was removed and spores were rinsed 5 times with 25 mM HEPES buffer (pH 7.4) using vacuum pump suction ((26)). The rinsed spores were then incubated at 37°C in 25 mM K-HEPES buffer with a pH of 7.4 unless noted otherwise. After various incubation times, the spores were given a 2nd exposure to germinants for various times at 37°C in 25 mM K-HEPES buffer

(pH 7.4) followed by germinant removal, rinsing by vacuum pump suction, and further incubation at 37°C in 25 mM K-HEPES buffer (pH 7.4).

Acknowledgements

CdK, LdK and SB acknowledge the University of Amsterdam for funding of their research.

Author contributions

CdK developed the model and performed computational analyses with input from LdK and SB. XL and YQL conducted experiments. PS provided spores of *Bacillus subtilis* strains. CdK wrote the manuscript with input from LdK, SB, YQL and PS. YQL, JL and PS designed the germination experiments.

Conflict of interest

The authors declare that they have no conflict of interest.

References

(1) Setlow P. Germination of spores of *Bacillus species*: what we know and do not know. J Bacteriol 2014 Apr;196(7):1297-1305.

(2) Paredes-Sabja D, Setlow P, Sarker MR. Germination of spores of *Bacillales* and *Clostridiales* species: mechanisms and proteins involved. Trends Microbiol 2011 Feb;19(2):85-94.

(3) Ross C, Abel-Santos E. The Ger receptor family from sporulating bacteria. Curr Issues Mol Biol 2010;12(3):147-158.

(4) Setlow P. Spore germination. Curr Opin Microbiol 2003 Dec;6(6):550-556.

(5) Igarashi T, Setlow B, Paidhungat M, Setlow P. Effects of a *gerF* (*lgt*) mutation on the germination of spores of *Bacillus subtilis*. J Bacteriol 2004 May;186(10):2984-2991.

(6) Wilson MJ, Carlson PE, Janes BK, Hanna PC. Membrane topology of the Bacillus anthracis GerH germinant receptor proteins. J Bacteriol 2012 Mar;194(6):1369-1377.

(7) Stewart KA, Setlow P. Numbers of individual nutrient germinant receptors and other germination proteins in spores of *Bacillus subtilis*. J Bacteriol 2013 Aug;195(16):3575-3582.

(8) Stewart KA, Yi X, Ghosh S, Setlow P. Germination protein levels and rates of germination of spores of *Bacillus subtilis* with overexpressed or deleted genes encoding germination proteins. J Bacteriol 2012 Jun;194(12):3156-3164.

(9) Griffiths KK, Zhang J, Cowan AE, Yu J, Setlow P. Germination proteins in the inner membrane of dormant *Bacillus subtilis* spores colocalize in a discrete cluster. Mol Microbiol 2011 Aug;81(4):1061-1077.

(10) Li Y, Jin K, Ghosh S, Devarakonda P, Carlson K, Davis A, et al. Structural and functional analysis of the GerD spore germination protein of *Bacillus* species. J Mol Biol 2014 May 1;426(9):1995-2008.

(11) Perez-Valdespino A, Li Y, Setlow B, Ghosh S, Pan D, Korza G, et al. Function of the SpoVAEa and SpoVAF proteins of *Bacillus subtilis* spores. J Bacteriol 2014 Jun;196(11):2077-2088.

(12) Vepachedu VR, Setlow P. Localization of SpoVAD to the inner membrane of spores of *Bacillus subtilis*. J Bacteriol 2005 Aug;187(16):5677-5682.

(13) Velásquez J, Schuurman-Wolters G, Birkner JP, Abee T, Poolman B. *Bacillus subtilis* spore protein SpoVAC functions as a mechanosensitive channel. Mol Microbiol 2014 May;92(4):813-823.

(14) Setlow B, Cowan A, Setlow P. Germination of spores of *Bacillus subtilis* with dodecylamine. J Appl Microbiol 2003;95(3):637-648.

(15) Wang G, Yi X, Li YQ, Setlow P. Germination of individual *Bacillus subtilis* spores with alterations in the GerD and SpoVA proteins, which are important in spore germination. J Bacteriol 2011 May;193(9):2301-2311.

(16) Doona CJ, Ghosh S, Feeherry FF, Ramirez-Peralta A, Huang Y, Chen H, et al. High pressure germination of *Bacillus subtilis* spores with alterations in levels and types of germination proteins. J Appl Microbiol 2014 SEP;117(3):711-720.

(17) Setlow P. Summer meeting 2013-when the sleepers wake: the germination of spores of *Bacillus* species. J Appl Microbiol 2013 Dec;115(6):1251-1268.

(18) Zhang P, Thomas S, Li YQ, Setlow P. Effects of cortex peptidoglycan structure and cortex hydrolysis on the kinetics of Ca(2+)-dipicolinic acid release during *Bacillus subtilis* spore germination. J Bacteriol 2012 Feb;194(3):646-652.

(19) Zhang P, Garner W, Yi X, Yu J, Li YQ, Setlow P. Factors affecting variability in time between addition of nutrient germinants and rapid dipicolinic acid release during germination of spores of *Bacillus* species. J Bacteriol 2010 Jul;192(14):3608-3619.

(20) Yi X, Setlow P. Studies of the commitment step in the germination of spores of *Bacillus* species. J Bacteriol 2010 Jul;192(13):3424-3433.

(21) Zhang P, Liang J, Yi X, Setlow P, Li YQ. Monitoring of commitment, blocking, and continuation of nutrient germination of individual *Bacillus subtilis* spores. J Bacteriol 2014 Jul;196(13):2443-2454.

(22) Wang S, Setlow P, Li YQ. Slow leakage of Ca-dipicolinic acid from individual *Bacillus* spores during initiation of spore germination. J Bacteriol 2015 Mar;197(6):1095-1103.

(23) Chen D, Huang SS, Li YQ. Real-time detection of kinetic germination and heterogeneity of single *Bacillus* spores by laser tweezers Raman spectroscopy. Anal Chem 2006 Oct 1;78(19):6936-6941.

(24) Peng L, Chen D, Setlow P, Li YQ. Elastic and inelastic light scattering from single bacterial spores in an optical trap allows the monitoring of spore germination dynamics. Anal Chem 2009 May 15;81(10):4035-4042.

(25) Setlow B, Peng L, Loshon CA, Li YQ, Christie G, Setlow P. Characterization of the germination of *Bacillus megaterium* spores lacking enzymes that degrade the spore cortex. J Appl Microbiol 2009 Jul;107(1):318-328.

(26) Wang S, Faeder JR, Setlow P, Li YQ. Memory of germinant stimuli in bacterial spores. MBio 2015 Nov 24;6(6):e01859-15.

(27) Xiong W, Ferrell JE,Jr. A positive-feedback-based bistable 'memory module' that governs a cell fate decision. Nature 2003 Nov 27;426(6965):460-465.

(28) Smits WK, Kuipers OP, Veening JW. Phenotypic variation in bacteria: the role of feedback regulation. Nat Rev Microbiol 2006 Apr;4(4):259-271.

(29) Tiwari A, Ray JC, Narula J, Igoshin OA. Bistable responses in bacterial genetic networks: designs and dynamical consequences. Math Biosci 2011 May;231(1):76-89.

(30) Frigola D, Casanellas L, Sancho JM, Ibanes M. Asymmetric stochastic switching driven by intrinsic molecular noise. PLoS One 2012;7(2):e31407.

(31) Tyson JJ, Chen KC, Novak B. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr Opin Cell Biol 2003 Apr;15(2):221-231.

(32) Becskei A, Seraphin B, Serrano L. Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. EMBO J 2001 May 15;20(10):2528-2535.

(33) Friedman N, Cai L, Xie XS. Linking stochastic dynamics to population distribution: an analytical framework of gene expression. Phys Rev Lett 2006 Oct 20;97(16):168302.

(34) Zhang JQ, Griffiths KK, Cowan A, Setlow P, Yu J. Expression level of *Bacillus subtilis* germinant receptors determines the average rate but not the heterogeneity of spore germination. J Bacteriol 2013 Apr;195(8):1735-1740.

(35) Vepachedu VR, Setlow P. Analysis of interactions between nutrient germinant receptors and SpoVA proteins of Bacillus subtilis spores. FEMS Microbiol Lett 2007 SEP;274(1):42-47.

(36) Zhang P, Kong L, Wang G, Setlow P, Li YQ. Combination of Raman tweezers and quantitative differential interference contrast microscopy for measurement of dynamics and heterogeneity during the germination of individual bacterial spores. J Biomed Opt 2010 Sep-Oct;15(5):056010.

(37) Raj A, van Oudenaarden A. Single-molecule approaches to stochastic gene expression. Ann Rev Biophys 2009;38:255-270.

(38) Chada VG, Sanstad EA, Wang R, Driks A. Morphogenesis of *Bacillus* spore surfaces. J Bacteriol 2003 Nov;185(21):6255-6261.

(39) Driks A. *Bacillus subtilis* spore coat. Microbiol Mol Biol Rev 1999 Mar;63(1):1-20.