

# Stochastic simulation of prokaryotic two-component signalling indicates stochasticity-induced active-state locking and growth-rate dependent bistability†

Cite this: *Mol. BioSyst.*, 2014, 10, 2338

Katy Wei,<sup>a</sup> Maxim Moinat,<sup>b</sup> Timo R. Maarleveld<sup>bcd</sup> and Frank J. Bruggeman<sup>\*be</sup>

Signal transduction by prokaryotes almost exclusively relies on two-component systems for sensing and responding to (extracellular) signals. Here, we use stochastic models of two-component systems to better understand the impact of stochasticity on the fidelity and robustness of signal transmission, the outcome of autoregulatory gene expression and the influence of cell growth and division. We report that two-component systems are remarkably robust against copy number fluctuations of the signalling proteins they are composed of, which enhances signal transmission fidelity. Furthermore, we find that due to stochasticity these systems can get locked in an active state for extended time periods when (initially high) signal levels drop to zero. This behaviour can contribute to a bet-hedging adaptation strategy, aiding survival in fluctuating environments. Additionally, autoregulatory gene expression can cause two-component systems to become bistable at realistic parameter values. As a result, two sub-populations of cells can co-exist—active and inactive cells, which contributes to fitness in unpredictable environments. Bistability proved robust with respect to cell growth and division, and is tunable by the growth rate. In conclusion, our results indicate how single cells can cope with the inevitable stochasticity occurring in the activity of their two-component systems. They are robust to disadvantageous fluctuations that scramble signal transduction and they exploit beneficial stochasticity that generates fitness-enhancing heterogeneity across an isogenic population of cells.

Received 29th April 2014,  
Accepted 13th June 2014

DOI: 10.1039/c4mb00264d

[www.rsc.org/molecularbiosystems](http://www.rsc.org/molecularbiosystems)

## 1 Introduction

The prevalence and diversity of two-component signalling systems across the prokaryotic domain is remarkable given their structural simplicity. They are composed of only two proteins: a histidine kinase ‘sensor’ protein (S), which is typically a membrane protein, and a cytosolic response regulator protein (R) that acts as a transcription factor. Protein-complex formation, (de)-phosphorylation, and phosphotransfer are the only types

of reactions that take place. Despite this, two-component systems are capable of sensing and responding to a wide range of environmental changes. They play a role in chemotaxis,<sup>1</sup> osmoregulation,<sup>2</sup> adaptation to Mg<sup>2+</sup>-limiting environments<sup>3</sup> and the induction of sporulation in *Bacillus subtilis*,<sup>4</sup> among others. This broad scope is partly due to the modularity of the system, which allows the basic mechanism to be extended with auxiliary components, but is likely also due to the intrinsic flexibility of the basic system. These aspects make two-component systems of major interest for synthetic biology. The simplicity of this molecular sensing circuit and the apparent lack of any internal regulatory interactions raise the question how this simple system copes with the inherent stochasticity of the reactions between its signalling components as well as the stochasticity introduced during cell growth and division. This is the question we address here, using stochastic simulations of two-component systems.

Two-component systems have been studied with mathematical models in previous studies. Ortega *et al.*<sup>5</sup> analysed the consequences of the bifunctionality of the histidine kinase—which acts as both an autokinase and a phosphatase—for signal sensitivity, and found that bifunctionality reduces sensitivity.

<sup>a</sup> Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands

<sup>b</sup> Systems Bioinformatics, Amsterdam Institute for Molecules, Medicines & Systems, VU University, De Boelelaan 1087, 1081 HV Amsterdam, The Netherlands.  
E-mail: f.j.bruggeman@vu.nl

<sup>c</sup> Life Sciences, Centrum voor Wiskunde en Informatica (CWI), Science Park 123, 1098 XG Amsterdam, The Netherlands

<sup>d</sup> BioSolar Cells, P.O. Box 98, 6700 AB Wageningen, The Netherlands

<sup>e</sup> Kluwyver Centre for Genomics of Industrial Fermentation/NCSB, P.O. Box 5057, 2600 GA Delft, The Netherlands

† Electronic supplementary information (ESI) available: Further modelling details, mathematical derivations and data have been provided together with modelling files and scripts used to produce the plots in the main article. See DOI: 10.1039/c4mb00264d

Shinar *et al.*<sup>6</sup> and Batchelor & Goulian<sup>7</sup> discovered that the output of two-component systems, the concentration of phosphorylated response regulator proteins, is robust with respect to the total concentration of their signalling components. Batchelor & Goulian<sup>7</sup> provided experimental evidence for this surprising behaviour. Hoyle *et al.*<sup>8</sup> and Igoshin *et al.*<sup>9</sup> analysed the bistability of two-component signalling transduction systems when coupled to autoregulatory gene expression. Hoyle *et al.*<sup>8</sup> further demonstrated this behaviour experimentally, using a synthetic two-component system.

The existing theoretical studies all use different models of two-component systems, regarding the precise reactions and kinetic parameter values that they consider. Here, we present a model of two-component signal transduction that is intermediate in size, considers bifunctionality and captures all the essential reactions: autophosphorylation, complex formation, phosphotransfer, and dephosphorylation. With this model, we revisit the phenomena of robustness and autoregulation-induced bistability, using stochastic simulations. We study both these system properties in more detail from the perspective of the impact of stochastic fluctuations on the beneficial properties of the signalling circuit. We also report a new behaviour of two-component systems that arises only in stochastic simulations and study the signalling system when embedded in a growing and dividing cell. This paper hopes to offer a model to the field that is of manageable complexity, yet sufficiently realistic and generic to be of practical use. By benchmarking these models against the established system behaviour of two-component systems, we hope to show that this model is flexible and realistic enough to be useful in other applications, such as in the design of new two-component systems in synthetic biology.

The model we propose captures the basic design of two-component systems and differs in several ways from previous models. Upon binding a signalling ligand (L), the sensor S changes conformation and autophosphorylates, yielding the phosphorylate species SP. Next, SP forms a complex with R, after which phosphotransfer can take place to yield RP and S. Typically, RP acts as a transcription factor inducing a gene expression response.<sup>4</sup> As aforementioned, in our model the sensor is bifunctional: when S is neither phosphorylated nor bound to L it can also dephosphorylate RP. Most sensor proteins have this property, including EnvZ from the EnvZ/OmpR model system responsible for osmoregulation in *Escherichia coli* (*E. coli*).<sup>2</sup> This captures the basic model without gene regulation (Fig. 1). Positive autoregulation, in which RP enhances transcription of the genes encoding R and S, is also not uncommon; the archetypal model system being the PhoP–PhoQ system first identified in *Salmonella enterica*.<sup>3</sup> We therefore also extend the basic model with autoregulatory gene expression and study the appearance of bistability which may result. Finally, we place the autoregulatory model in a cell that grows exponentially in a steady-state manner to investigate whether bistability is impacted by cell division and the required synthesis of signalling components accompanying cell volume growth.

We find that the basic, generic two-component model that we propose (1) protects itself spontaneously against hazardous

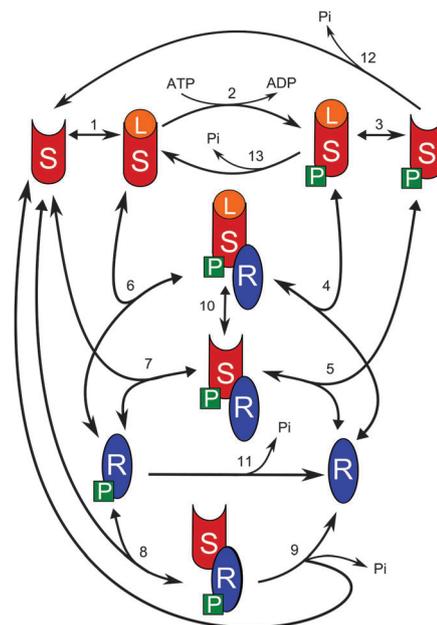


Fig. 1 Network reaction model. The quadrilateral arrowheads indicate the 'forward' directions of the (numbered) reactions, while the triangular arrowheads indicate the reverse directions. Pictorially represented network nodes are treated as individual species in the *in silico* model, other reactants are considered to have fixed concentrations in the cell. Rate constants corresponding to each reaction are provided in Table S1.†

fluctuations that scramble signal transduction and (2) manages to profit from fluctuations that generate beneficial heterogeneity in the cell population, which is advantageous for organisms living in unpredictable, dynamic environments.

## 2 Results

### 2.1 A model of the two-component signalling system

The topology of the signalling network we consider in this work is shown in Fig. 1. This model contains the core reactions known to occur in basic two-component systems.<sup>4</sup> When the L concentration is high, high levels of RP are expected to be produced and the system is in the 'active-state'. Conversely, when the signal is weak (or absent) the amount of RP is low; we will refer to this situation as the 'off-state'. The main simplification made was to consider the sensors and response regulators as monomers, though in reality they are often dimers.<sup>10</sup> Shinar *et al.*<sup>6</sup> have already shown that the robustness of two-component systems with respect to the total sensor and regulator concentration is preserved in monomer and dimer models.

In our simulations we describe all reactions in terms of mass-action kinetics. For stochastic simulations, we used the Gillespie direct method in StochPy,<sup>11</sup> a Python-based stochastic simulation platform. For the incorporation of cell division we use an extension of StochPy described in the ESI† (Fig. S2). We will later also consider 'total' R and S numbers,  $R_T$  and  $S_T$ , which refer to the total number of R and S molecules in the cell, irrespective of phosphorylation or complex formation. The ligand concentration was considered fixed—equivalent to assuming

a large extracellular reservoir of the signalling molecule. Also, the intracellular concentrations of ATP and inorganic phosphate are considered fixed and therefore not explicitly modelled. Our estimated rate constants are provided in Table S1 of the ESI.† In making the estimations, we ensured that the parameters were physically relevant by considering critical factors, such as the diffusion limit which restricts the rates of second-order reactions, as well as matching published experimental results with the model, as described in Section 2.2.

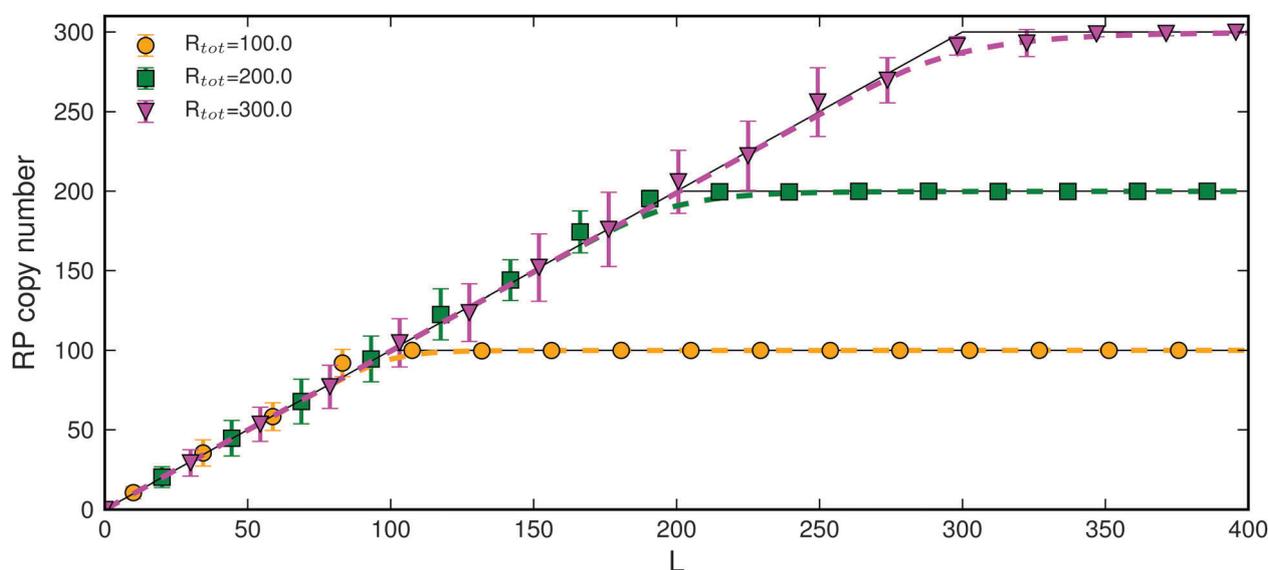
## 2.2 Robustness of signal transmission against fluctuations in signalling protein abundance

In order to validate our model, the first step we took was to compare its output with existing experimental data. The best validated experimental results pertain to steady-state activation levels of the system, typically measured indirectly by placing a reporter gene under transcriptional control of a RP-responsive promoter.<sup>7</sup> Such data indicate that steady-state concentrations of RP increase linearly in proportion to concentrations of the extracellular signalling molecule  $L$ .<sup>7</sup> Furthermore, there is also data available which suggest that the steady-state concentrations of RP are substantially robust to changes in the total number of  $R$  and  $S$  molecules in the cell.<sup>1,6</sup> This allows the cell to maintain a consistent response despite fluctuations in molecule numbers due to stochasticity in protein production, degradation and cell division events. Robustness is, however, limited because the number of RP molecules cannot exceed  $R_T$ , so RP concentrations always saturate under high stimulus.

To test whether our model displayed the same (qualitative) characteristics, we performed a mathematical analysis† based

on the approach reported by Shinar & Alon<sup>6</sup> as well as deterministic and stochastic simulations. The results of the mathematical analysis indicate that the steady-state RP concentration of the model is indeed expected to increase linearly with  $L$ , independent of  $R_T$  and  $S_T$ , until saturation.† The parameter requirements for this behaviour are:† (1) the binding of the sensor to the ligand should occur on a fast enough time scale such that this reaction always operates close to thermodynamic equilibrium and (2) the spontaneous dephosphorylation rates of RP and SP should be negligible when compared to the catalysed reactions in the model. For additional discussion we refer to Shinar & Alon.<sup>6</sup>

In the numerical simulations, the network shown in Fig. 1 was considered at fixed levels of  $R_T$  and  $S_T$ , defined by the initial conditions. The results of these simulations are shown in Fig. 2. The deterministic steady-state RP concentrations indeed increase linearly with  $L$  for relatively low levels of stimulation, independently of  $R_T$  (and  $S_T$ , though these data are not shown). As expected, the response also saturates ( $RP = R_T$ ) when the signal strength is sufficiently high. Furthermore, the mean steady-state concentration of RP in the stochastic simulations also follows the same trend. Hence, we conclude that fluctuations in  $R_T$  and  $S_T$ —which may arise from stochasticity during transcription, translation or cell division—do not scramble the output of the signalling network. This acts as a built-in mechanism to ensure high signal transmission fidelity in two-component systems despite inevitable fluctuations in their components. We emphasise that this behaviour is only dependent on the two parameter requirements (mentioned above) and that experimental evidence for his behaviour exists.<sup>7</sup>

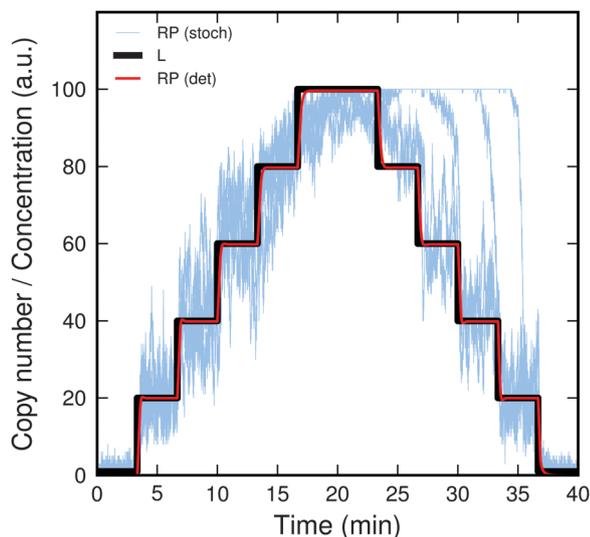


**Fig. 2** Steady-state RP dependence on the signal level,  $L$ . Deterministic and stochastic simulation results for three different  $R_T$  values (100, 200, 300) are coloured orange, green and purple respectively. Deterministic results are given by dashed lines, while time-averaged means of stochastic simulations (calculated over 1000 seconds after the steady-state levels were reached) are plotted with markers. Standard deviations of the stochastic simulations are also shown as error bars; where they are not visible, they are smaller than the marker size. The mathematical predictions are also shown as thin back lines. Note that since the concentration of  $L$  is considered fixed, the units are arbitrary and hence the gradient of the linear portion of the plot is also arbitrary. In this study, saturation is reached when  $L = R_T$ .

### 2.3 Stochasticity-induced active-state locking upon sudden removal of the signal

Next, we studied the response time of the system—the time that the system requires to attain a steady-state RP concentration. In order for our model to be physiologically relevant, we must consider that organisms have to respond within a timescale significantly shorter than their generation time.<sup>12</sup> We therefore expected this response time to be on the order of about one minute. This is also what one would expect if the reaction rates are considered in their diffusion limit (not shown). To address the response time, we performed stochastic simulations of the signalling network (Fig. 1) and subjected the system to a stepwise-varying signal strength, ranging from negligible to saturating L concentrations (Fig. 3). Indeed, the deterministic simulation shows that the response time ranges from approximately 20 to 60 seconds. The stochastic simulation agrees with this timescale when the signal is increasing. However, when the signal is reduced from a saturating level, some stochastic trajectories show surprising behaviour. Namely, the RP concentration has a tendency to stay ‘locked’ in the active-state for a prolonged time period—sometimes over ten minutes—after the signal has been reduced. We call this interesting phenomenon ‘stochasticity-induced active-state locking’.

The parameter conditions for stochasticity-induced active-state locking only concern a small subset of all the parameters, as was also the case for the robustness. The cause of stochasticity-induced active-state locking is that the number of (unbound, unphosphorylated) S molecules can drop to zero when the system is subjected to saturating signal strengths. As the S molecules are the dominant means of dephosphorylation of RP,



**Fig. 3** The response of the system to time-varying signal strengths, where copy numbers of R and S are fixed at 100. The concentration of L, shown in black, was increased stepwise at 200 second intervals from 1 to 100 a.u., then subsequently decreased. The deterministic prediction for the output RP concentration is shown in red, and 10 stochastic simulations are shown in blue. The deterministic response time ranges from 20–60 seconds, but for the stochastic simulations active-state locking is seen to occur after L reaches 100 a.u., when saturation occurs.

their absence prevents the number of RP molecules from declining. In our model this is achieved in the simplest way as only the free sensor state, S, can dephosphorylate RP. The autodephosphorylation rates are slow, and although the phosphorylation reactions are reversible, the forward rates are overwhelmingly favoured. Thus, the new lower equilibrium value of RP at the reduced level of L can only be reached once unbound, unphosphorylated S molecules are recovered by unbinding/dephosphorylation reactions. Since these reactions occur stochastically and at a low rate, the observed waiting time can be remarkably long; in our simulations, we have observed locking for up to ten minutes. As long as those reactions are reversible and occur at a low enough rate the time that it takes before S molecules reappear can be quite variable amongst individual cells.

We emphasize that this result is not reproducible in deterministic simulations. The reason is that, deterministically, the concentration of S is a continuous variable that would take a small but non-zero value under saturating conditions and reactions which increase S (and decrease RP) concentrations start to occur immediately (albeit at a low rate) upon decreasing the signal strength.

In order to further investigate this locking behaviour, we studied a single stepwise drop in signal strength, from a saturating level to a negligible level (Fig. 4A). As we now expected, the stochastic trajectories tended to lock in the active state for up to ten minutes, but over time an increasing number of trajectories managed to escape the locked state. The histograms taken at the different time points (Fig. 4B) show that trajectories that remain locked have completely saturated RP populations, and also illustrate the transiently bimodal distribution just after the signal drops. If each stochastic trajectory is interpreted as an individual cell in an isogenic population, this locking phenomenon should thus create a similarly transient bimodal distribution, which should be experimentally observable. A quantification of the variability of the response time can be obtained by numerically determining the distribution of the time that it takes for cells to reach a threshold % of unphosphorylated R molecules. This resembles a hitting time and is defined in the analysis of the stochastic time series. As long as cells do not communicate *via* cell-to-cell signalling, stochasticity-induced active state locking can transiently diversify populations of cells which can be advantageous in unpredictable environments.

### 2.4 Autoregulatory gene expression leads to bistability of signalling activity

At this point, we extended the stochastic model with autoregulatory gene expression. The two-component system then up-regulates the transcription of the genes encoding its own sensor and response regulator, as illustrated in Fig. 5. We based our transcription model on the structure of the PhoP–PhoQ system, best studied in *Salmonella enterica*.<sup>3</sup> The coding sequence for S follows that of R in the same operon, and they are under the transcriptional control of two promoters; one constitutive, and the other activated by a RP dimer.

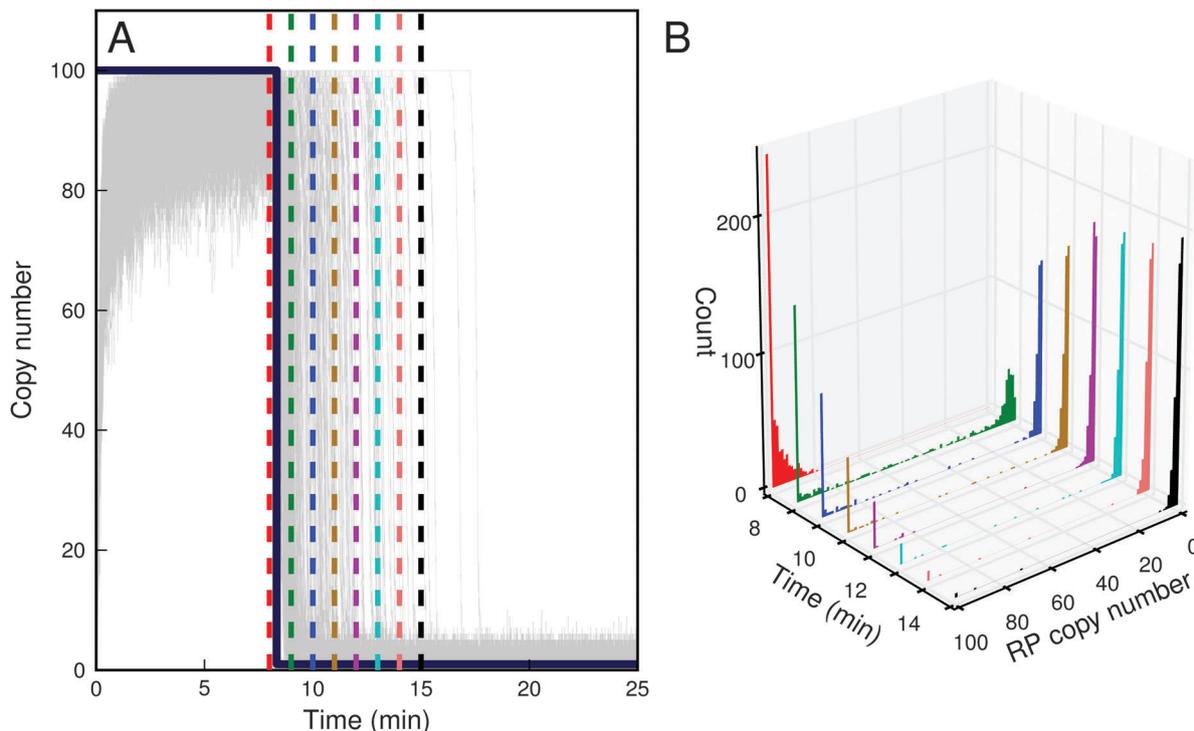


Fig. 4 Active-state locking for 500 stochastic simulations. (A) Time traces of the output RP numbers from 500 stochastic simulations. Histograms were prepared for time points indicated by colour-coded dashed lines. (B) Histograms for RP numbers at the different time-points, illustrating the evolution of the transient bimodal population.

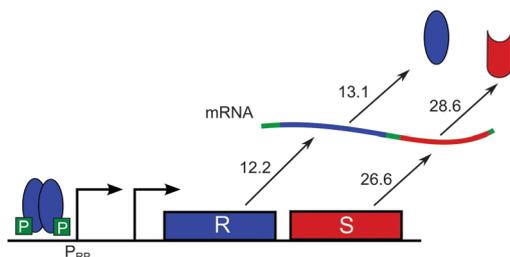


Fig. 5 Genetic diagram illustrating autoregulation of transcription as used in this model. R and S genes are coded on the same operon, and transcribed onto a single mRNA strand, which is subsequently translated to produce the protein. Times taken for transcription and translation is indicated in seconds, and modelled by gamma distributions in the stochastic simulations. Protein folding times are assumed to be negligible relative to the translation time.

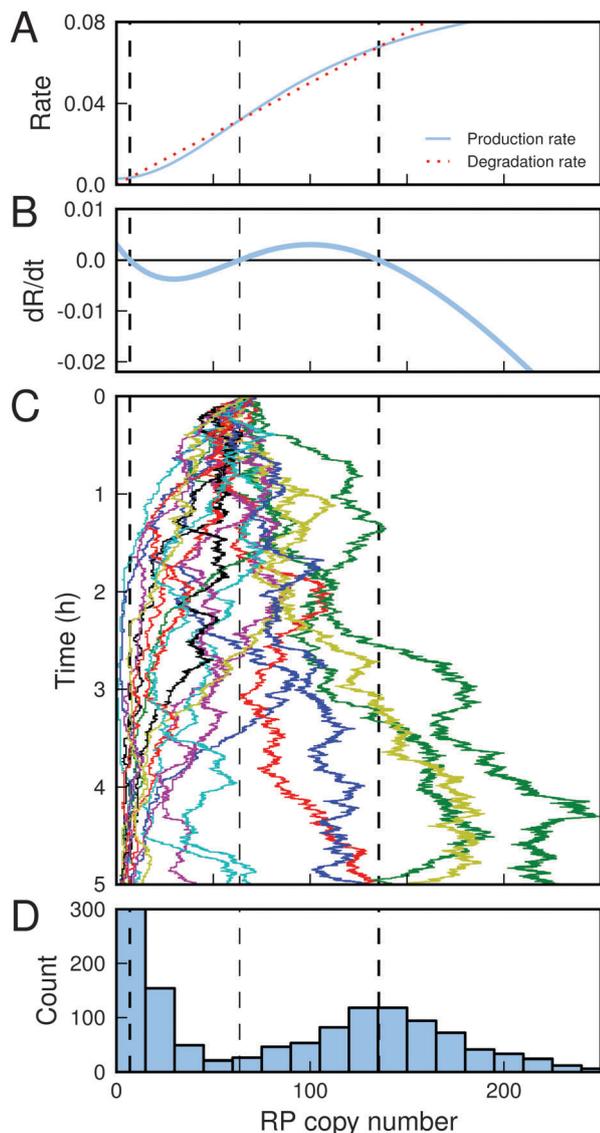
To investigate the steady-state behaviour of the autoregulatory system, we began with a simple (deterministic) mathematical model consisting of the single equation:

$$\frac{d[R_T]}{dt} = k_{p1} + k_{p2} \frac{[RP]^2}{K_d^2 + [RP]^2} - k_{deg}[R_T] \quad (1)$$

where  $k_{p1}$ ,  $k_{p2}$  and  $k_{deg}$  are the basal production rate constant, the maximal activated production rate constant and the degradation rate constant of R proteins, respectively. Furthermore,  $k_{p2}$  precedes a Hill function that describes the saturation of the promoter with RP. The  $K_d$  is the microscopic dissociation constant and the Hill coefficient has been taken equal to 2 to

account for RP dimerisation. We have therefore approximated the kinetics of RP dimers binding on a single operator site as the binding of RP monomers on a pair of operator sites with infinite cooperativity. To further simplify our calculations, we considered saturating signal strengths, such that  $[R_T] = [RP]$ . Setting the right-hand side of the equation to zero—*i.e.* investigating the steady-state concentrations—thus gives a cubic equation for  $[RP]$  (or, equivalently,  $[R_T]$ ) which can in principle have one, two or three solutions. When three solutions are exhibited, two are stable, while if only one or two solutions exist only one state is stable (Fig. 6A and B). Depending on the parameter values, therefore, the system may exhibit bistability. We found that, for physically reasonable parameter values ( $ESI^\dagger$ ), both bistable and non-bistable states were possible. When the system is bistable, the system may settle in either an active-state or an off-state under high stimulus, depending on the initial conditions. We note that the conditions for bistability derive from the parameters in eqn (1); in fact the parameter ratios  $k_{p2}/k_{p1}$  and  $k_{deg}/k_{p1}$  and the  $K_d$ . When the R molecules have a long life time, then  $k_{deg}$  is sensitive to the cellular growth rate otherwise not. So, bistability depends on a relation between gene expression, protein stability and growth rate parameters. Also the separating border between two attractive regions (grey dashed line in Fig. 6) is set by those parameters.

The largest simplification in the above equation is the presence of the Hill function, which estimates the average proportion of time that the system spends activated without considering the individual dimerisation and binding-unbinding



**Fig. 6** Bistability in the full autoregulatory model. (A) RP production and degradation rates, given by the positive and negative terms in eqn (1) respectively. Where production and degradation rates intersect, a steady-state RP level is reached. (B) The total rate of change in [RP] (eqn (1)), for saturating signal strengths and parameters permitting bistability. The first and last fixed points are stable, while the central one is unstable, separating the basins of attraction for the two stable states. (C) 20 stochastic simulations with an initial condition  $RP = 70$  (close to the unstable point) illustrate the bistability in the system, as trajectories may tend towards either the upper or the lower steady state when they are stochastically pushed into either basin of attraction. (D) Histogram of RP levels after 16.66 hours for 3000 stochastic simulations, revealing a bimodal population. The first bin had a count of 2046, but the total height was cut in the figure for clarity.

reaction rates. For our stochastic simulations, we considered all the individual reaction steps in order to create a more realistic model (see ESI<sup>†</sup>), by adding these reactions to the signalling network depicted in Fig. 1. Also, we accounted for (stochastic) delay times in transcription and translation using average transcription/translation rates from *E. coli* and gene lengths taken from the autoregulatory PhoP–PhoQ two-component

system in *E. coli* (see ESI<sup>†</sup>). Using parameters that we found to cause bistability in the reduced mathematical model given in eqn (1), we reproduced the phenomenon in the stochastic model (Fig. 7A and B). Occasionally, stochasticity caused bistable switching, in which a simulation initially settled in one stable activation state spontaneously switches to the other. The overall effect on a large number of such cells is a bimodal population, as shown in Fig. 6D, where a proportion of cells remain in an ‘off’ state despite high signal strengths. This could be advantageous as an evolutionary bet-hedging strategy for signalling systems responding to toxic stimuli, which we will consider further in the Discussion. The parameter requirement for bistable switching is that cells should be able to pass the border separating the two attractive regions within a single generation time (grey dashed line in Fig. 6).

### 2.5 Growth-rate dependent stochastic bistable switching between inactive and active signalling states

Two notable features of the bistable system prompted us to investigate the influence of cell growth and division on the stochastic behaviour of the autoregulatory circuit. Firstly, the timescale over which the system settles into a steady-state—on the order of tens of minutes as shown in Fig. 6C—can easily exceed the generation time of a typical prokaryotic organism. For example, *E. coli* can have a generation time of under 20 minutes when growing on glucose in rich media. During this time, the cell doubles in size and cell growth is thus often the dominant source of protein concentration ‘degradation’, particularly when the protein is not targeted for enzymatic degradation.<sup>13</sup> Cell division also has the dramatic effect of (approximately) halving the protein copy numbers nearly instantaneously. The division of molecules between daughter cells can be well-modelled by a volume-dependent binomial distribution.<sup>14</sup> With bistability-switching timescales exceeding generation times, it is therefore important to check how bistability is influenced by cell growth and division of the host cell. Secondly, *via* the mathematical analysis, we found that the presence of bistability was especially sensitive to changes in the protein degradation rate  $k_{\text{deg}}$  (data not shown). Since the protein degradation rate can equivalently be interpreted as a protein dilution rate, we hypothesised that if cell growth is taken as the only source of protein degradation we would find that the bistability of the system becomes growth-rate dependent. This would then constitute a parameter requirement for the occurrence of growth-rate induced bistability.

In implementing growth and division processes in StochPy, (exponential) growth rates were assumed to be fixed but the generation time, daughter cell volume and division of cellular species between daughter cells were modelled stochastically (ESI<sup>†</sup>). We found that a generation time of about 45 minutes—typical for bacteria such as *E. coli* and *Bacillus subtilis* under laboratory conditions—could indeed lead to bistable steady-state RP concentrations (Fig. 7C and D). The activation state is also occasionally seen to switch spontaneously due to stochasticity. This indicates that the phenomenon of bistability is also robust to growth and division processes, providing further

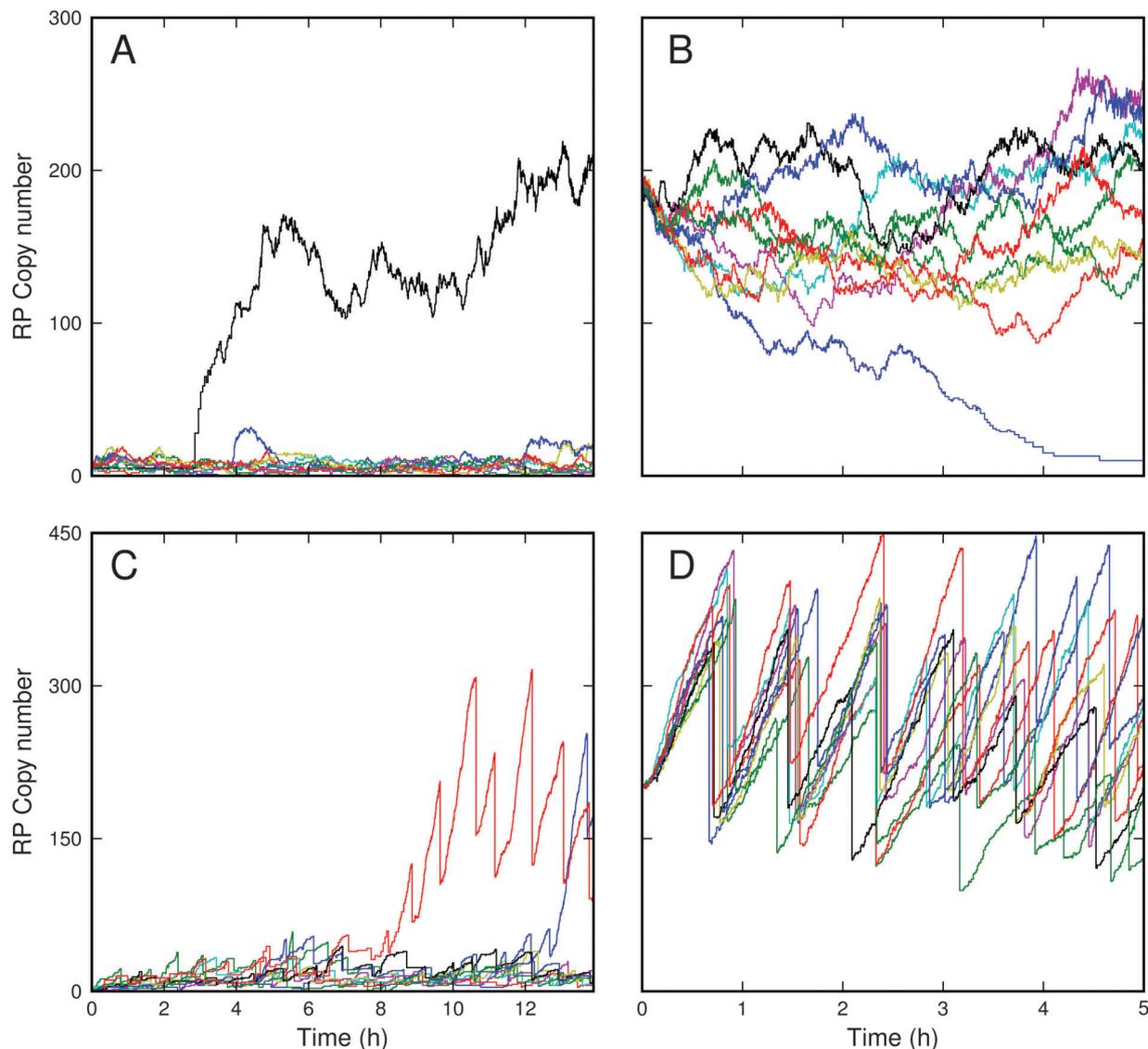


Fig. 7 10 stochastic simulations with two different initial conditions, demonstrating bistability in the system without (A and B) and with cell division (C and D). Switching between the two stable states is occasionally enabled by stochasticity. (A) Initial RP population is low, so the steady-state output is biased towards the 'off-state'. (B) Initial RP population is high, so the steady-state output is biased towards the active-state. (C) Initial RP population is low, so the steady state output is biased towards the 'off-state'. (D) Initial RP population is high, so the steady-state output is biased towards the active-state.

support for the possibility of either finding or creating such a system in nature.

We also found that this bistability is growth-rate dependent. An upper bound to bistability-permitting growth rates is determined by cell division events—the cell growth rate must be slow enough to allow the protein population to at least double during the generation time. Otherwise, the protein population in mother cells decreases with each generation that passes, so the active-state can never be maintained. This upper limit on the growth rate can be adjusted by varying the protein production rates,  $k_{p1}$  and  $k_{p2}$ . There was also a lower bound on growth-rate which arose from the original mathematical model, because the degradation rate must exceed a critical value in order for the 'off-state' to be maintained. For our parameters, cell generation times of between 45 and 110 minutes permitted bistability.

### 3 Discussion

In this work, we studied with stochastic simulations a model of two-component signalling that has a generic, realistic network structure and is of manageable complexity. We studied how this system copes with its inherent stochasticity, deriving from its reactions as well as the growth processes of the cell, and how induction of bistability by auto-regulatory gene expression is affected by cell growth. We found that two-component systems are robust with respect to disadvantageous stochasticity while they can also exploit stochasticity to cause diversification of the cell population.

We discovered a new behaviour of two-component systems that we termed 'active-state locking'. Active-state locking transiently creates a bimodal population after a sudden drop in

signal strength. This behaviour can contribute to bet-hedging strategies where a subpopulation of cells remains active and 'anticipates' a return of the signal. One can imagine that this is advantageous when the ligand signals toxicity. The locked fraction of the population remains in the 'safe-mode' after the signal decreases and is hence well-prepared if the toxicity level rises again. The remainder of the population loses the anti-toxic response rapidly after the signal is removed, enabling them to free more resources for other metabolic processes. In this way, the population as a whole is better prepared for survival in fluctuating and/or heterogeneous environments.

To our knowledge, the stochasticity-induced active-state locking phenomenon has not yet been reported by either theoretical or experimental research groups. The lack of experimental evidence may initially raise doubts about the biological applicability of our findings. However, there are numerous factors that should be considered. Firstly, the bimodal population is only transient—unless an active search is made for this phenomenon, the probability of observation is low. Secondly, the phenomenon also relies on a saturating signal strength that causes all R molecules to be phosphorylated and all S molecules to be either ligand-bound or phosphorylated. This requires a high signal strength, and is also facilitated if  $R_T$  and  $S_T$  are relatively low. This may not be feasible in all two-component systems, but suggests a direction for future research.

Several two-component systems are also known to be autoregulatory. Our results and those of Igoshin *et al.*<sup>9</sup> and of Kierzek *et al.*<sup>15</sup> indicate that this property can lead to bistability. Bistability persists even when cell growth and division effects are (stochastically) accounted for. We found that the bistability exhibits a growth-rate dependence and is sensitively dependent on protein production and degradation parameters. The growth-rate dependence has a particularly interesting interpretation for the case of toxicity sensing. In this case, the growth-rate dependency may be interpreted as a 'back-up mechanism' to judge whether or not the cell should respond to a detected high-toxicity environment. If the growth-rate is sufficiently fast, indicating that the cell's metabolism is still functioning soundly (and little toxin is likely present), then the system remains in the off-state and the high detected level of toxicity is treated as a false alarm. For example, a morphologically similar molecule could have been mistaken for the toxin. Alternatively, if the growth-rate is sufficiently slow, indicating that the organism has been badly affected, then the system switches to the active-state, where a large amount of RP molecules is produced and the cell can respond to the toxic threat. For intermediate growth-rates bistability exists, leading to a bimodal population in which a proportion of the cells are activated while the rest are off—another example of a bet-hedging evolutionary strategy that helps the species survive in fluctuating environments.

Experimental evidence for bistability in the autoregulatory PhoB-PhoR system was obtained by Zhou *et al.*,<sup>16</sup> where flow cytometry data showed bimodal distributions qualitatively similar to Fig. 6D. Moreover, Kierzek *et al.* reproduced these distributions by using a model accounting for stochasticity in transcription, translation and degradation in growing cells.<sup>15</sup>

We have incorporated several additional sources of stochasticity in our more detailed signalling network model, and showed that the bistability can still be maintained. The additional stochasticity arose not just from the inclusion of more stochastic reactions, but also from more detailed modelling of cell growth, division and gene expression.

We considered stochasticity of the cellular generation time, daughter cell volumes, volume-dependent binomial division of cell contents, and gamma-distributed transcription/translation delay times to account for the polymerisation of each nucleotide/amino acid at an exponentially-distributed rate. Growth dependence of bistability was also investigated in a simple synthetic system by Tan *et al.*,<sup>13</sup> but is yet to be found in a natural two-component system. The advantages of such a system could also lend itself very well to applications in synthetic biology, where persistence in a fluctuating or heterogeneous environment may often be required.

An important feature of our findings is that their qualitative features are quite robust with respect to the reaction rate parameters used in our model. Robustness of the steady-state signal transmission relies on just two assumptions: (1) S and L binding reactions are relatively fast and (2) autodephosphorylation reaction rates for RP and SP are relatively low. Active-state locking and bistability only require the signalling system to saturate at high signal strengths (*i.e.* all R molecules are phosphorylated, and all S molecules are ligand-bound and/or phosphorylated when L is high). Bistability further requires an autoregulatory gene network in combination with more stringent restrictions on the protein production and degradation rates. It is, however, well established that these rates vary widely in nature for different signalling systems, depend heavily on bacterial growth conditions and are easily variable in synthetic systems.<sup>17,18</sup> We expect that our results are robust to uncertainties in precise parameter values in the signalling network, which are often difficult to quantify and control. Furthermore, our findings are likely broadly applicable to two-component signal transduction systems in general, both natural and synthetic.

## 4 Conclusions

The rich and varied dynamics that we report in this study highlights the remarkable versatility of a biological sensing system that consists of only two components. This suggests a possible explanation for why two-component systems are so widespread among prokaryotes and why a single species may rely on tens of such signalling systems: these systems are reliable, versatile and small enough to be rapidly evolvable. Furthermore, it also bodes well for the future of synthetic biology and minimal cell research, which hope to achieve advanced functionality using minimally complex building blocks. Finally, our work demonstrates the importance of conducting stochastic simulations of molecular circuit models, as they can uncover startling new dynamics with biological significance—such as active-state locking—which may not be evident from deterministic simulations.

## Acknowledgements

The authors thank Anne Schwabe for very helpful discussions. TRM acknowledges funding from the project BioSolar Cells, co-financed by the Dutch Ministry of Economic Affairs. FJB acknowledges funding of NWO-VIDI Project 864.11.011.

## References

- 1 M. Goulian, *Curr. Opin. Microbiol.*, 2010, **13**, 184–189.
- 2 S. J. Cai and M. Inouye, *J. Biol. Chem.*, 2002, **277**, 24155–24161.
- 3 E. A. Groisman, *J. Bacteriol.*, 2001, **183**, 1835–1842.
- 4 A. M. Stock, V. L. Robinson and P. N. Goudreau, *Annu. Rev. Biochem.*, 2000, **69**, 183–215.
- 5 F. Ortega, L. Acerenza, H. V. Westerhoff, F. Mas and M. Cascante, *PNAS*, 2002, **99**, 1170–1175.
- 6 S. Shinar, R. Milo, M. Rodriguez Martinez and U. Alon, *PNAS*, 2007, **104**, 19931–19935.
- 7 E. Batchelor and M. Goulian, *PNAS*, 2003, **100**, 691–696.
- 8 R. B. Hoyle, D. Avitabile and A. M. Kierzek, *PLoS Comput. Biol.*, 2012, **8**, e1002396.
- 9 O. A. Igoshin, R. Alves and M. A. Savageau, *Mol. Microbiol.*, 2008, **68**, 1196–1215.
- 10 C. M. Barbieri, T. Wu and A. M. Stock, *J. Mol. Biol.*, 2013, **425**, 1612–1626.
- 11 T. R. Maarleveld, B. G. Olivier and F. J. Bruggeman, *PLoS One*, 2013, **8**, e79345.
- 12 E. Couturier and E. P. Rocha, *Mol. Microbiol.*, 2006, **59**, 1506–1518.
- 13 C. Tan, P. Marguet and L. You, *Nat. Chem. Biol.*, 2009, **5**, 842–848.
- 14 I. Golding, J. Paulsson, S. M. Zawilski and E. C. Cox, *Cell*, 2005, **123**, 1025–1036.
- 15 A. M. Kierzek, L. Zhou and B. L. Wanner, *Mol. BioSyst.*, 2010, **6**, 531–542.
- 16 L. Zhou, G. Grgori, J. M. Blackman, J. P. Robinson and B. L. Wanner, *J. Integr. Bioinform.*, 2005, **2**, 11.
- 17 S. Mukherji and A. van Oudenaarden, *Nat. Rev. Genet.*, 2009, **10**, 859–871.
- 18 A. S. Khalil and J. J. Collins, *Nat. Rev. Genet.*, 2010, **11**, 367–379.