Unrelated toxin–antitoxin systems cooperate to induce persistence

Rick A. Fasani and Michael A. Savageau

Department of Biomedical Engineering and Microbiology Graduate Group, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA

Persisters are drug-tolerant bacteria that account for the majority of bacterial infections. They are not mutants, rather, they are slow-growing cells in an otherwise normally growing population. It is known that the frequency of persisters in a population is correlated with the number of toxin–antitoxin systems in the organism. Our previous work provided a mechanistic link between the two by showing how multiple toxin–antitoxin systems, which are present in nearly all bacteria, can cooperate to induce bistable toxin concentrations that result in a heterogeneous population of slow- and fast-growing cells. As such, the slow-growing persisters are a bet-hedging subpopulation maintained under normal conditions. For technical reasons, the model assumed that the kinetic parameters of the various toxin–antitoxin systems in the cell are identical, but experimental data indicate that they differ, sometimes dramatically. Thus, a critical question remains: whether toxin–antitoxin systems from the diverse families, often found together in a cell, with significantly different kinetics, can cooperate in a similar manner. Here, we characterize the interaction of toxin–antitoxin systems from many families that are unrelated and kinetically diverse, and identify the essential determinant for their cooperation. The generic architecture of toxin–antitoxin systems provides the potential for bistability, and our results show that even when they do not exhibit bistability alone, unrelated systems can be coupled by the growth rate to create a strongly bistable, hysteretic switch between normal (fast-growing) and persistent (slow-growing) states. Different combinations of kinetic parameters can produce similar toxic switching thresholds, and the proximity of the thresholds is the primary determinant of bistability. Stochastic fluctuations can spontaneously switch all of the toxin–antitoxin systems in a cell at once. The spontaneous switch creates a heterogeneous population of growing and non-growing cells, typical of persisters, that exist under normal conditions, rather than only as an induced response. The frequency of persisters in the population can be tuned for a particular environmental niche by mixing and matching unrelated systems via mutation, horizontal gene transfer and selection.

1. Introduction

Persisters have traditionally been identified as subpopulations of cells that survive antibiotic treatment via epigenetic means. They were first recognized while treating Staphylococcus with penicillin [1] and were later identified as the source of multidrug tolerance in biofilms [2], making them responsible for 65% to 80% of bacterial infections [3,4]. Persisters have been implicated in the stubborn Pseudomonas aeruginosa infections to which most cystic fibrosis patients eventually succumb [5], as well as the oral Candida albicans infections common in cancer patients [6]. They may also explain the recurrence of Mycobacterium tuberculosis infections, responsible for 1.6 million deaths each year [7].

Persistence is not the result of a genetic mutation, but rather of a heterogeneous population. Modern single-cell studies have confirmed that persisters are rare, slowly growing cells [8], and that slowly growing cells are less susceptible to antibiotics [9]. More recent evidence suggests that slow growth is not necessary for nor a guarantee of persistence, but still increases the likelihood [10]. The mechanisms

© 2015 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.
that provide the antibiotic tolerance are not fully understood, but one common path to persistence appears to be through the pervasive and varied toxin–antitoxin systems [11]. Toxin–antitoxin systems are genetic modules, commonly found in free-living bacteria, that generally consist of two co-produced and co-regulated components: a relatively stable toxin that inhibits cell growth and a more labile antitoxin that specifically neutralizes the toxin. Stress can upregulate the proteases that degrade the antitoxin, thereby freeing the toxin [12]. The overexpression of toxin can slow growth [13–19] and confer multidrug tolerance [18,20–22]. Conversely, multiple toxin–antitoxin systems are upregulated in persistor-enriched samples [21,23]. In fact, the first gene tied to persistence was hipA [24], later identified as the toxic half of a toxin–antitoxin pair.

Toxin–antitoxin systems are found on the chromosomes and plasmids of most bacterial species and strains—the Escherichia coli K-12 genome boasts at least 36 [25] and the M. tuberculosis genome contains 88, more than any other human pathogen [26]. Yet, despite a growing understanding of the mechanisms underlying toxin–antitoxin systems, several important questions remain unanswered. What are their functions and how does each contribute to different cellular phenotypes or fates [27]? Why are there multiple types and apparently redundant systems in a single cell [28]? What is their coordinating signal [29]? Our previous work [30] answered some of these questions by forming a general model of the common type II toxin–antitoxin systems that target protein synthesis, and comparing the model behaviour to existing experimental results. Previous treatments addressed various aspects of toxin–antitoxin systems or persistent populations [8,31–36], whereas our analysis was the first to encompass them all, including molecular mechanisms of regulation, stochastic fluctuations, variable growth, and population dynamics, and to do so over a broad range of parameter values. We were able to describe a connection between the molecular mechanisms of toxin–antitoxin systems, the cooperation among systems to produce bistable expression, and the slow growth that is commonly associated with the persistent phenotype.

Our previous results [30] confirmed and explained genetic experiments [29] that revealed a characteristic and important relationship between the number of toxin–antitoxin genetic cassettes and the frequency of persisters that survive antibiotic treatment. Furthermore, these results suggested that although the specifics may vary, toxin–antitoxin systems are potentially bistable and can create a hysteretic switch between normal and persistent states. A bistable system can exhibit one of two stable behaviors under the same conditions, and it has become apparent that bistable genetic regulatory networks, when operating in noisy, fluctuating environments, can lead to heterogeneous populations of cells. This can be seen in Bacillus subtilis genetic competence, spore formation, and swimming or chaining, as well as the persistent phenotype studied here [37]. We previously showed [29] how toxin–antitoxin systems that do not exhibit bistability alone can be coupled to produce the same effect. Moreover, the total number of toxin–antitoxin systems in a cell tunes the frequency of persisters, using the growth rate as the coordinating signal.

For tractability in treating the large numbers of toxin–antitoxin systems, our previous analysis considered the systems kinetically identical, even if their specific mechanisms differed. Indeed, the fact that the toxins inhibit protein production via diverse molecular methods and targets creates a multihit mechanism that is necessary for cooperativity, but not always sufficient—it can be readily shown that two systems with randomly chosen, differing kinetics do not always cooperate. Moreover, critical questions remain: whether or not toxin–antitoxin systems from different families—with similar motifs but dramatically different kinetics—cooperate, and if so, what the essential factors are that determine their cooperation.

Here, we remove the key restriction that provided tractability in the original model and account for multiple distinct toxin–antitoxin systems. We mix and match systems with kinetic parameters that vary over an order of magnitude, reflecting the parameter ranges that have been measured across several well-studied toxin–antitoxin families. The results extend our past findings, as well as offer new insight. In particular, multiple unrelated toxin–antitoxin systems cooperate via the growth rate, particularly in the presence of stochastic fluctuations, to robustly increase the size of the bistable region and the frequency of persisters. Furthermore, the size of the bistable region not only depends on the parameters of each system, but critically on their relative switching thresholds—the value of the stimulus at the inflection point of the toxin induction characteristic. As such, different toxin–antitoxin systems can be mixed and matched to provide the variation on which selection can act to tune the persister frequency for a given environmental niche.

2. Methods

2.1. Single-system model

Toxin–antitoxin systems use different protein structures and mechanisms, yet are consistent in overall architecture. Here, we describe a generic model for a large class of toxin–antitoxin systems: type II systems that target protein synthesis. Figure 1a depicts the common species and their interactions, and figure 1b represents a generic model for toxin–antitoxin systems, where \( T \) and \( A \) are the concentrations of free antitoxin and toxin, respectively. \( T \) can represent either monomeric or dimeric toxin, provided the toxin completely folds and dimerizes at physiological concentrations. We assume that the synthesis and degradation of the polycistronic message, \( M \), as well as the formation and dissociation of the toxin–antitoxin complexes, \( C \) and \( D \), are relatively fast compared with the rest of the system. As such, the dynamics of a single toxin–antitoxin system can be described by the following system of conventional differential equations:

\[
\frac{dA}{dt} = \beta A + \frac{1}{1 + T^n/K_T^A} + (\frac{\Delta A^2}{K_{P_A}^2}Y^2 + \frac{\Delta^2 A^2}{K_{P_A}^2 K_T^A}) - \frac{\mu_{\text{max}}}{1 + T^n/K_T^A} A - \lambda_A A + \xi_A
\]

(2.1)

and

\[
\frac{dT}{dt} = \beta T + \frac{1}{1 + T^n/K_T^T} + (\frac{\Delta T^2}{K_{P_T}^2}Y^2 + \frac{\Delta^2 T^2}{K_{P_T}^2 K_T^T}) - \frac{\mu_{\text{max}}}{1 + T^n/K_T^T} T - \lambda_T T + \xi_T,
\]

(2.2)

or by differential algebraic equations in the generalized mass action form [38]:

\[
\frac{dA}{dt} = \sigma_{\text{max}} A X^{-1} Y^{-1} - \mu_{\text{max}} A X^{-1} - \lambda_A A + \xi_A,
\]

(2.3)

\[
\frac{dT}{dt} = \alpha_{\text{max}} X^{-1} Y^{-1} - \mu_{\text{max}} T X^{-1} - \lambda_T T + \xi_T,
\]

(2.4)
The toxin enhances repression, in some cases via a bridging mechanism. Free antitoxin is relatively labile and degraded by various proteases (not shown). Free toxin usually inhibits some aspect of global translation. (b) M, mRNA; A, antitoxin; T, toxin; C, antitoxin bound to one toxin; D, antitoxin bound to two toxins. A and T are both translated from the polycistronic message M. Antitoxin alone or in complex—A, C or D—autorepresses transcription. Free toxin T inhibits translation, including its own. All species are degraded, and diluted by cellular growth. The degradation of A increases with proteolytic activity (not shown), and dilution is slowed when T inhibits global translation and growth (not shown). Either toxin—T₁ or T₂—can inhibit translation in both systems as well as the overall growth rate, which slows the dilution of all species.

\[
X = 1 + \frac{T^n}{K_T^n} \quad (2.5)
\]

and

\[
Y = 1 + \frac{A^2}{K_{AI}^2} + \left( \frac{2A^2T}{K_{AI}^2,K_T^2} \right)^p + \frac{A^2T^2}{K_{AI}^2,K_T^2}. \quad (2.6)
\]

Equations (2.3)–(2.6) comprise a simplified model that retains the essential features of the version used in our previous work [30]. In both equations (2.3) and (2.4), the first and only positive term on the right-hand side represents protein synthesis. The multiplicative factor \( \sigma \) represents the translational coupling between toxin and antitoxin, and \( a_{\text{max}} \) is the maximum rate of protein synthesis. The second term represents protein loss due to dilution, where \( \mu_{\text{max}} \) is the maximum growth rate constant. The third term in each equation represents protein loss due to active degradation, where \( \lambda_A \) is the degradation rate constant of antitoxin and \( \lambda_T \) is the degradation rate constant of toxin. Both protein synthesis and cellular growth are slowed by the toxic inhibition of translation \( X^{-1} \) [14,17], which is defined by equation (2.5). The impact of the free toxin \( X^{-1} \) follows a Hill equation with Hill number \( n \) and a concentration for half-maximal activity \( K_T \). Similarly, the independent, unitless parameter \( \beta \) reflects an external change that simultaneously lowers protein synthesis and cellular growth, or \( a_{\text{max}} \) and \( \mu_{\text{max}} \). Protein synthesis is also proportional to the fraction of unbound promoter, \( Y^{-1} \), defined by equation (2.6). In most well-studied toxin–antitoxin systems, the operators are dissimilar, with one dominant, high-affinity site [39–44]. In this model, we ignore the weaker sites or, for the cooperatively binding complex, treat the sites in aggregate. The second term represents the dimerization of the antitoxin on the surface of the promoter, with a concentration of half-maximal binding at \( K_{AI} \). The third term represents the cooperative binding of the complex C to the promoter, with Hill number \( p \) and a dissociation constant \( K_{IC} \) that represents increased affinity. The complex C forms when a toxin binds the antitoxin dimer at one of two independent sites with a dissociation constant \( K_{IT} \). The single toxin can bind to either site, and hence binds with twice the affinity, or \( 2/K_{IT} \). The fourth term assumes that the complex D binds the operator with the same affinity as the bare dimer, or \( K_{DI} \). The complex D forms when two toxins bind the antitoxin dimer with overall affinity \( 1/K_{IT} \).

Studies over the past decade have confirmed the importance of stochastic fluctuations, or noise, in gene expression and genetic networks [45]. Noise is added to the model by augmenting equations (2.3) and (2.4) with additive noise terms, producing stochastic differential equations in Langevin form. \( \xi_t \) is a white noise term with zero mean, \( \langle \xi_t(t) \rangle = 0 \), and \( \delta \)-autocorrelation, \( \langle \xi_t(s) \xi_t(t) \rangle = d \delta(t - s) \), where \( d \) is proportional the strength of...
the perturbation. Together, equations (2.3)–(2.6) form a tractable, generic model of toxin–antitoxin regulation.

### 2.2. Coupled-systems model

Experimental evidence indicates that one toxin–antitoxin system can trigger another, unrelated toxin–antitoxin system [46–48]. Our previous work suggested an indirect method of accomplishing such coupling [30]. If one system alone increases its own toxin concentration in response to a specific stress, and consequently slows protein production and growth, then another identical system may respond to the change in growth just as it would respond to a decrease in β. In this work, we extended the model to include a second, unrelated toxin–antitoxin system, as illustrated in figure 1. The extended model is described by equations (2.7)–(2.13), in which equations (2.7)–(2.9) form the first system and equations (2.10)–(2.12) describe the second. Both subsets of equations mirror equations (2.3), (2.4) and (2.6) of the single-system model. The maximum growth rate constant \( \mu_{\text{max}} \) and the parameter \( \beta \) are the same for both systems, whereas the other parameters can take on separate values. The systems are coupled by equation (2.13), where either toxin can affect both systems by attenuating global protein production and growth.

### Parameter values

Apart from noise, the single-system model contains 11 parameters, which we estimated in previous work, based on the published data for six of the best-studied toxin–antitoxin systems: \( \text{kis-kid} \) (penciK), \( \text{ccdAB} \), \( \text{mazEF} \), \( \text{phd-loc} \), \( \text{relBE} \) and \( \text{yefM-yoeB} \) [30]. Table 1 includes the previously estimated parameter values. For technical reasons, in our previous work, we treated the kinetics of all toxin–antitoxin systems as identical. However, the published data indicate that toxin–antitoxin kinetics vary between families, sometimes over multiple orders of magnitude. Here, in addition to the original parameter estimates, we consider alternative sets of parameters, or alternative designs, that include various fold changes to the estimated values. The additional parameters are listed in table 1. The alternative sets of parameters vary over a broad range, but are not random. They were deliberately chosen for reasons we will make clear later in the text.

### 2.4. Computational procedures

We constructed and analysed the system design space using the Design Space Toolbox for Matlab v. 1.0 [52]. We simulated the deterministic model with the Matlab stiff solver, ode15s. We simulated the stochastic model with our own implementation of the Euler–Maruyama method [53] in Matlab. All tests were performed using Matlab v. 7.8 (R2009a).

### 3. Results

#### 3.1. Coupled cooperativity

The interaction of unrelated systems can be modelled by setting the parameters of equations (2.7)–(2.13) to represent two systems with different kinetics. Figure 2 shows the interaction of nearly identical systems where the difference in a single parameter value creates varying degrees of separation in their switching thresholds, or the value of the stimulus at the inflection point of the toxin induction characteristic. Figure 2r depicts the toxin profile of the first system, with the originally estimated parameter values, in response to changing \( \Delta_{\text{A}1} \), which is commonly correlated with changing proteolysis and stress. Note that the system does not exhibit hysteretic bistability, whereas a nearly identical system in our previous work did. In our previous model, the toxin concentrations for

<table>
<thead>
<tr>
<th>parameter</th>
<th>S1</th>
<th>S2</th>
<th>S3</th>
<th>S4</th>
<th>S5</th>
<th>S6</th>
<th>S7</th>
</tr>
</thead>
<tbody>
<tr>
<td>( t_x ) (min)</td>
<td>60</td>
<td>120</td>
<td>240</td>
<td>60</td>
<td>120</td>
<td>240</td>
<td>120</td>
</tr>
<tr>
<td>( t_f ) (h)</td>
<td>48</td>
<td>96</td>
<td>192</td>
<td>48</td>
<td>96</td>
<td>192</td>
<td>96</td>
</tr>
<tr>
<td>( \sigma )</td>
<td>10</td>
<td>10</td>
<td>2.5</td>
<td>2.5</td>
<td>10</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>( \alpha_{\text{max}} ) (nM min(^{-1}))</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>4</td>
<td>1</td>
<td>0.25</td>
<td>1</td>
</tr>
<tr>
<td>( K_{\mu} ) (nM)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>400</td>
<td>100</td>
<td>25</td>
<td>100</td>
</tr>
<tr>
<td>( K_{\mu} ) (μM)</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>10</td>
<td>1</td>
<td>0.1</td>
</tr>
<tr>
<td>( K_{\mu} ) (nM)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>2.5</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>( K_{\mu} ) (nM)</td>
<td>10</td>
<td>10</td>
<td>10</td>
<td>2.5</td>
<td>10</td>
<td>2.5</td>
<td>10</td>
</tr>
</tbody>
</table>

The known toxins target their own unique steps in translation [49] and it can be shown that their combined impact on translation is often multiplicative. For example, RatA blocks the association of the ribosomal subunits [50], whereas MazF cleaves mRNA [51]—decreasing both the concentration of functional ribosomes by half and the concentration of mRNA by half would reduce translation initiation fourfold. The coupled model nearly doubles the number of variables and parameters, but the same computational methods can be applied.
Figure 2. Toxin induction profiles for coupled and uncoupled systems. Steady-state toxin concentrations, $T_1$ and $T_2$, as a function of the changing rate constant for antitoxin degradation in both systems $\lambda_{A_1}$. An increase in $\lambda_{A_1}$ is commonly associated with an increase in stress. Changes in $\lambda_{A_1}$ are measured as a fold change from the normal values $\lambda_{A_1}^0$. The proteolytic activity changes $\lambda_{A_2}$ in both systems simultaneously, i.e. $\lambda_{A_2}/\lambda_{A_1}^0 = \lambda_{A_2}/\lambda_{A_1}^0$. (a) Uncoupled reference system, using the originally estimated parameter values (dashed black). (b) Uncoupled additional systems using the same parameter values, except varying the normal rates of antitoxin degradation: $\lambda_{A_1}^0 = 64\lambda_{A_1}^0$ (blue), $8\lambda_{A_1}^0$ (light blue), $\lambda_{A_1}^0$ (green), $1/8\lambda_{A_1}^0$ (pink) and $1/64\lambda_{A_1}^0$ (red). The toxic thresholds, or the values of the stimulus at the inflection points of the toxin induction characteristics, are evenly spaced at approximately $-3.7$ (blue), $-0.7$ (light blue), $2.3$ (green) and $5.3$ (pink). (c,d) Steady-state toxin concentrations (dashed colour) when the reference system (a, dashed black) is coupled with each of the additional systems (b, solid colour). (c) The coupled reference system and (d) the coupled additional systems when $\lambda_{A_2}^0 = 64\lambda_{A_1}^0$ (dashed blue), $8\lambda_{A_1}^0$ (dashed light blue), $\lambda_{A_1}^0$ (dashed green), $1/8\lambda_{A_1}^0$ (dashed pink) and $1/64\lambda_{A_1}^0$ (dashed red).

3.2. Design spaces

The model described by equations (2.3)–(2.6) is capable of exhibiting a rich phenotypic repertoire, and the individual phenotypes can be effectively enumerated and analysed within the system design space, as has been shown with other biological systems [54–56]. In design space, the behaviour of each phenotype is described by a dominant subsystem with a single, analytically defined steady state [52]. Note that equations (2.3) and (2.4) each have one positive term on the right-hand side, excluding noise, and multiple negative terms, while equations (2.5) and (2.6) each use multiple terms to define $X$ and $Y$. Biologically, each term represents a process. For a given set of parameter values, one negative term or one defining term in each equation may be larger than the others, or dominate. If the smaller terms, or processes, are ignored, the behaviour of the remaining subsystem can be analysed using well-known techniques [57]. There are 32 possible cases, or combinations of dominant terms, and each case represents a potentially unique phenotype. In the figures that follow, each case will be depicted as a region in design space. Table 2 lists all of the regions that are visible in this work, along with a description of their dominant terms and some properties of their resulting phenotypes.
same behaviour should be expected over a wide range: if such ultrasensitivity. Furthermore, figure 3 indicates that the systems that do not exhibit hysteresis can still exhibit a space can overestimate the size of the hysteresis region, in figure 2 Hysteresis, or bistability, is not evident in the toxin profile of tic transition between two stable steady states—high and low work showed that the intermediate regions represent a hysteretic region of multiple phenotypes, or steady states. Our previous the growth rate falls. Notably, the system passes through a and transcription is completely derepressed in the fourth region. Antitoxin loss and toxin loss are found in equations (2.3) and (2.4), respectively. When the first negative term in either equation dominates, the principal form of loss is dilution, but when the second negative term dominates, the loss is primarily due to active degradation. Toxic inhibition hinges on equation (2.5), and when the first term dominates, the toxic inhibition of protein production is considered low; when the second term dominates, or the toxin concentration rises above $K_h$, toxic inhibition is considered high. Transcriptional repression is described by equation (2.6), and when the first term dominates, the promoter is unbound, but when the third term dominates, the toxin–antitoxin complex acts as a strong antirepressor by cooperatively binding the promoter. Each region exhibits a distinct phenotypic behaviour, including the toxin concentration, growth rate and system stability, which are shown here. More detailed descriptions of the relevant regions and their significance can be found in the text.

**Table 2.** Summary of characteristics for the regions in design space. As described in the text, each of the numbered regions represents a distinct case in which a different combination of terms creates a dominant subsystem that can be more effectively analysed. Here, the dominant terms, or processes, are described in each region. Antitoxin loss and toxin loss are found in equations (2.3) and (2.4), respectively. When the first negative term in either equation dominates, the principal form of loss is dilution, but when the second negative term dominates, the loss is primarily due to active degradation. Toxic inhibition hinges on equation (2.5), and when the first term dominates, the toxic inhibition of protein production is considered low; when the second term dominates, or the toxin concentration rises above $K_h$, toxic inhibition is considered high. Transcriptional repression is described by equation (2.6), and when the first term dominates, the promoter is unbound, but when the third term dominates, the toxin–antitoxin complex acts as a strong antirepressor by cooperatively binding the promoter. Each region exhibits a distinct phenotypic behaviour, including the toxin concentration, growth rate and system stability, which are shown here. More detailed descriptions of the relevant regions and their significance can be found in the text.

<table>
<thead>
<tr>
<th>region</th>
<th>dominant terms</th>
<th>system phenotype</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>antitoxin</td>
<td>toxin</td>
</tr>
<tr>
<td></td>
<td>loss</td>
<td>loss</td>
</tr>
<tr>
<td>1</td>
<td>dilution</td>
<td>dilution</td>
</tr>
<tr>
<td>3</td>
<td>dilution</td>
<td>dilution</td>
</tr>
<tr>
<td>7</td>
<td>dilution</td>
<td>dilution</td>
</tr>
<tr>
<td>17</td>
<td>degradation</td>
<td>dilution</td>
</tr>
<tr>
<td>19</td>
<td>degradation</td>
<td>dilution</td>
</tr>
<tr>
<td>21</td>
<td>degradation</td>
<td>dilution</td>
</tr>
<tr>
<td>23</td>
<td>degradation</td>
<td>dilution</td>
</tr>
<tr>
<td>27</td>
<td>degradation</td>
<td>degradation</td>
</tr>
<tr>
<td>29</td>
<td>degradation</td>
<td>degradation</td>
</tr>
<tr>
<td>31</td>
<td>degradation</td>
<td>degradation</td>
</tr>
</tbody>
</table>

Figure 3 depicts the design space over a wide range of parameter values. The normal operating point—defined by the originally estimated parameters—resides within Region 3, where dilution dominates in equations (2.3) and (2.4), free toxin is below its $K_a$, and transcription is mostly repressed by the cooperative complex, or the third term, in equation (2.6). The result is a phenotype with normal, or relatively fast, growth. Under stressful conditions, increased proteolytic activity increases the active degradation of the antitoxin, which is represented by moving to the right in figure 3. If the antitoxin half-life is reduced eightfold ($\lambda_A/\lambda_A^0 = 8$), the system moves from Region 3 to 21, where active degradation dominates the first equation, dilution still dominates the second equation, free toxin is above its $K_a$ in the third equation, and transcription is completely derepressed in the fourth equation. In other words, the toxin concentration rises and the growth rate falls. Notably, the system passes through a region of multiple phenotypes, or steady states. Our previous work showed that the intermediate regions represent a hysteretic transition between two stable steady states—high and low toxin levels—with an intermediate, unstable steady state [30]. Hysteresis, or bistability, is not evident in the toxin profile of figure 2a, but we also noted in our previous work that design space can overestimate the size of the hysteretic region, in which systems that do not exhibit hysteresis can still exhibit a supralinear, or ultrasensitive, profile. Indeed, figure 2a displays such ultrasensitivity. Furthermore, figure 3 indicates that the same behaviour should be expected over a wide range: if $\beta$ is increased or decreased by twofold, an increase in $\lambda_A$ would also move the system from Region 3 to Region 21, albeit at a different value of $\lambda_A$. In fact, the boundaries are linear functions of the multiplicative parameter values [52], and therefore mathematically relate the changing threshold to the change in $\beta$. Simple observation indicates that halving $\beta$ would halve the change required of $\lambda_A$ to reach the threshold. In the same fashion, design space analysis can be used to explore the global behaviour of the nonlinear system over wide ranges of every parameter value—the impact of the other parameters is shown in the full set of design spaces depicted in figure 4. In nearly every case, a twofold increase or decrease in the parameter value does not change the system’s hysteretic response to increasing $\lambda_A$. As for the threshold of transition, increasing $\alpha_{\text{max}}$, $K_{P2}$ or $K_{H1}$ moves it closer to the
normal operating point, whereas increasing $\beta$, $\mu_{\text{max}}$, $\sigma$ or $K_T$ moves the threshold farther away. Decreasing any of those parameters has the opposite effect. Changing $\lambda_T$ or $K_P$ has little to no effect.

Decreasing $\beta$—a concomitant decrease in protein synthesis and growth—exhibits similar hysteretic behaviour, as shown in figure 5a. An eightfold decrease in $\beta$ moves the system from Region 3 to Region 29, which is similar to Region 21 save that active degradation dominates in equation (2.4). In other words, the toxin concentration is high and the growth rate is even lower than in Region 21. The transition is hysteretic, similar to the transition induced by increasing proteolytic activity [30]. Furthermore, figure 4b, d shows that it is the decrease in maximum growth rate $\mu_{\text{max}}$, not the decrease in maximum protein production $\alpha_{\text{max}}$, that induces the change in phenotype. As in figure 4, the design spaces of figure 5 indicate how each parameter affects the threshold of hysteresis when lowering $\beta$. Increasing $\lambda_T$, $\alpha_{\text{max}}$, $K_P$ or $K_I$ moves the threshold closer to the normal operating point; increasing $\mu_{\text{max}}$, $\sigma$ or $K_T$ moves the threshold farther away. Decreasing those parameters has the opposite effect. Changing $\lambda_T$ or $K_P$ has little to no effect.

### 3.3. Steady states

In each case, the dominant subsystem can be analytically solved for a single steady state [52]. Of particular interest are the steady-state concentrations at the endpoints of the hysteretic switch, in Regions 3, 21 and 29. The concentrations, in terms of the independent parameters, are described by equations (3.1)–(3.6).

\[
A = \left( \frac{K_T^2 K_P^2 \alpha_{\text{max}} \sigma}{4 \mu_{\text{max}}} \right)^{1/7} \tag{3.1}
\]

\[
T = \left( \frac{K_T^2 K_P^2 \alpha_{\text{max}}}{4 \mu_{\text{max}} \sigma^2} \right)^{1/7} \tag{3.2}
\]

\[
A = \left( \frac{\beta K_T^2 \mu_{\text{max}} \sigma}{\alpha_{\text{max}} \lambda_A} \right)^{1/3} \tag{3.3}
\]

\[
T = \left( \frac{\alpha_{\text{max}}}{\mu_{\text{max}}} \right)^{1/3} \tag{3.4}
\]

\[
A = \left( \frac{\beta K_T^2 \alpha_{\text{max}} \sigma \lambda_A^2}{\lambda_T} \right)^{1/3} \tag{3.5}
\]

\[
T = \left( \frac{\beta K_T^2 \alpha_{\text{max}}}{\lambda_T} \right)^{1/3} \tag{3.6}
\]
Equations (3.1) and (3.2) indicate that the steady-state concentrations in Region 3 are not dependent on \( \lambda_s \). However, as previously shown, an eightfold increase in \( \lambda_s \) shifts the system to Region 21, where equation (3.4) reveals that the toxin concentration is higher, but again independent of \( \lambda_s \). Together, equations (3.2) and (3.4) define the low and high toxin levels of a hysteretic switch that is sensitive to increased proteolytic activity. Similarly, an eightfold decrease in \( \beta \) shifts the system from Region 3 to Region 29, where the steady-state solutions are described by equations (3.5) and (3.6). Interestingly, the equations show that the concentrations remain dependent on \( \beta \), so a further decrease in \( \beta \) will continue to decrease the toxin and antitoxin, albeit relatively slowly. Regardless, equations (3.2) and (3.6) define the low and high toxin concentrations of a hysteretic switch that is sensitive to the rate of protein production, or growth.

3.4. Alternative designs

The characteristics of the hysteretic switch depend on the parameter values, and the parameter values vary among toxin–antitoxin systems. The normal operating point is a representative estimate of measured parameter values in several well-studied systems [30] and is listed as design S1 in table 1. The table also describes pseudo-random alternative designs, or parameter sets, where each set includes multiple fold changes to the parameters of the normal operating point. Although none of the alternative designs represents a specific toxin–antitoxin system, the variations are typical of those seen in experimentally characterized systems [30]. For example, the antitoxin YefM has a measured half-life of 60 min [58], the value at the normal operating point, but Phd has a half-life of 120 min [59], a value used in alternative designs S2, S5 and S7. Furthermore, the promoter dissociation constant \( K_{phd} \) of Phd [39,60,61] is at least an order of magnitude lower than other well-studied systems [31,41,43,44,62–65], similar to the 10-fold decrease in alternative design S7. The translational coupling factor in relBE is 10 [66], the normal operating value, but for kis-kid the measured coupling factor is 2 [67], similar to alternative designs S3 and S4. The toxin and antitoxin typically bind tightly, but measurements of the dissociation constant \( K_{IT} \) vary: some estimates for ccdAB, maxEF and relBE are lower than the normal operating value of 100 nM [66,68,69], as is the value in S6, whereas the measured value for yefM-geoB is 400 nM [58], the same value used in S4. Thus, the variation among the alternative designs is typical of the variation among toxin–antitoxin systems.
The panoply of design spaces in figures 4 and 5, together with equations (3.2), (3.4) and (3.6), describe the influence of each parameter on the relevant characteristics of the switch: the low and high toxin concentrations and the threshold of transition between them. The information can be used to predict the behaviour of the alternative designs. In fact, the alternative designs were chosen to exhibit specific variations in the toxic profiles, as shown in figure 6. Alternative designs S1, S2 and S3 were chosen to vary the threshold of toxic switching while maintaining approximately the same low and high toxin concentrations, which can be seen in figure 6a–c. Alternative designs S4, S5 and S6 were chosen to change both the toxic range and the toxic threshold, which can be seen in figure 6d–f. Designs S2, S5 and S7 were chosen to exhibit identical toxic profiles, despite their kinetic differences. Furthermore, figure 6 shows that none of the toxic profiles exhibits bistability, or more than one steady state along the curve, which is a result of the relatively low toxic impact. However, the systems exhibit switch-like behaviour between two states and behave as designed, illustrating the potential power of design space to inform the forward engineering of biological systems.

3.5. Heterogeneous cooperativity

Figure 7 describes the effect of coupling the distinctly different alternative designs. The first system is set to the values for alternative design S7, and paired with each of the other alternative designs. \( \lambda_{A,1} \) is increased with no corresponding change in \( \lambda_{A,2} \), which models an increase in proteolytic activity that only affects the first system. The resulting steady-state toxin concentrations are shown in figure 7a,b. In every case, the stressed induction of the first system indirectly triggers the second system. Furthermore, in several cases, the systems together produce hysteretic behaviour, whereas none was observed for either system alone (figure 6). Figure 7b indicates that the cooperative effect depends on the relative toxic thresholds, but not the relative toxic ranges. The results are similar when \( \lambda_{A,2} \) is increased alone, or when \( \lambda_{A,1} \) and \( \lambda_{A,2} \) are increased together, as if both systems respond to the same protease (not shown). Figure 7c,d shows similar hysteretic behaviour when global growth and protein production are decreased via \( \beta \). Decreasing \( \mu_{\text{max}} \) alone also exhibits hysteresis (not shown). We showed in previous work that multiple identical toxin–antitoxin systems can cooperate to create hysteresis and increase the size of the bistable region [30]. Here, we show that extremely different toxin–antitoxin systems also cooperate when their toxic thresholds are similar, regardless of their toxic range. Similarly, the cooperation is robust to large changes in several parameter values, although the effect varies based on the relative toxic thresholds. As such, the mixing and matching of diverse systems via mutation, horizontal gene transfer and selection can generate a variety of switching profiles, of which some are optimal for a particular environmental niche.

3.6. Stochastic switching

Any member of a population that operates within the bistable region should tend towards either the low or high toxin concentration, creating a distinctly bimodal population. Furthermore, our previous work showed that stochastic fluctuations can spontaneously switch cells from one state to the other [30]. However, those simulations assumed identical...
systems with correlated noise. In this work, we used the coupled model of equations (2.7)–(2.13), and stochastically simulated the unrelated systems with uncorrelated noise terms. We poised an entire population at the low toxin concentration, which corresponds to normal growth, and then added noise. Figure 8\textit{a},\textit{b} shows that the toxin concentrations of the two systems fluctuated independently, generally near the low toxin concentration at which they were poised. The systems infrequently and spontaneously switched to fluctuate about the high steady state and appeared to switch together. As a result, figure 8\textit{c} shows that individual members of the population infrequently and spontaneously switched to the slow growth, persistent state and, in at least one case, subsequently recovered. The transitions are infrequent, as is to be expected given that persister frequency in a population can be as low as $10^{-6}$ [24]. Numerous additional simulations exhibited the same behaviour, and mirror the stochastic behaviour we thoroughly tested and described in our previous work [30]. Here, the results confirm that the bistability introduced by unrelated toxin–antitoxin systems can also give rise to a dynamically changing subpopulation of persisters.

\section{Discussion}

Previous experimental work has shown that the frequency of persisters in a population is correlated with the number of toxin–antitoxin systems in the organism [29]. Our previous work provided a mechanistic link between the two. Using a simplified model where every toxin–antitoxin system was kinetically identical, we showed that the frequency of persisters is related to the width of the bistable region, which is in turn dependent on the number of toxin–antitoxin systems in the cell [30]. By extending the previous work, we are now able to show that the same result can be accomplished using a more realistic model that includes toxin–antitoxin systems with dramatically different kinetics. The bistable switch can be driven by proteolysis of the antitoxin, stochastic fluctuation or a change in the growth rate. The presence of bistability and the switching threshold itself are dependent on the other parameters of the system. Even if a single system lacks bistability, multiple systems can be coordinated by the growth rate to produce the same effect. Other work has shown a similar link between growth and bistability, where modulating the growth rate can create an implicit feedback loop, bistability and a heterogeneous population [70]. In our past work, the toxin–antitoxin systems were kinetically identical. The results here indicate that kinetically different systems—even systems that are dramatically different—can be coupled to produce bistability, and that the strength of the effect, or the size of the bistable region, is dependent on the relative switching thresholds of the individual systems.

Experimental work has shown that sufficient stress can induce persistence [71–73], which can also be seen in our model (figure 2). However, our model also shows how, under relatively normal conditions, stochastic fluctuations can spontaneously transition toxin–antitoxin systems to the persistent state. Even when the stochastic fluctuations in each system are uncorrelated, the toxin–antitoxin systems in a cell switch together, coupled by the growth rate. The spontaneous, coordinated switch to persistence describes how toxin–antitoxin systems can give rise to a bimodal population of normal and persistent cells. We predict that a persistent subpopulation would always be present, even under normal conditions, as a
bet-hedging strategy to survive a catastrophic event. Indeed, there is some experimental evidence suggesting that this might be the case [8]. Interestingly, the switch back to the normal state is also spontaneous, and therefore the time spent in persistence is variable and unpredictable. It is possible that the emergence from persistence is just as important, if not more important, than entering it—persisters that emerge too early, before an environmental stress abates, will still perish. However, remaining dormant indefinitely is not a viable strategy either. A persistent subpopulation that survives catastrophe may employ a bet-hedging strategy of its own: individual members might randomly revert to normal and thrive if the stress is gone; if not, the remaining persistent population carries on.

Persisters frequency in a wild-type population of *E. coli* is typically between $10^{-6}$ and $10^{-5}$ [24], but can vary by species, strain and environment—the frequency in a biofilm of *P. aeruginosa* may be as high as $10^{-2}$ [74]. Experimental results show that the frequency can be altered dramatically by varying the overall number of toxin–antitoxin systems in the cell [29]. Our model confirms those results and suggests that the effect is dependent on the average toxic impact of each system [30]. The average toxic impact of each system, in turn, is a function of their relative toxic thresholds. In other words, adding toxin–antitoxin systems with similar switching thresholds increases the persister frequency more than systems with relatively different thresholds. These results offer an explanation for the abundance, variation and apparent redundancy of unrelated toxin–antitoxin systems: the heterogeneous systems can be mixed and matched by mutation and horizontal gene transfer, creating populations with varying persister frequencies, each fit for a particular environmental niche.

**Authors’ contributions.** R.A.F. and M.A.S. designed research; R.A.F. performed research; R.A.F. contributed new analytic tools; R.A.F. and M.A.S. analysed data; and R.A.F. and M.A.S. wrote the paper.

**Competing interests.** We declare we have no competing interests.

**Funding.** This work was supported in part by a grant from the US Public Health Service (RO1-GM30054) to M.A.S., and by an Earl C. Anthony Fellowship to R.A.F.

**Acknowledgements.** We thank Dean A. Tolla, Jason G. Lomnitz, Stephen W. Barthold and Mitchell Singer for discussions and suggestions.

**Figure 8.** Stochastic simulation with spontaneous switching. Fifty thousand members of a population, with coupled alternative designs S1 and S7, were poised at the low-toxin steady state for $\lambda_t/\lambda_F = 2.5$ and simulated with noise ($\sigma = 0.2$). A random selection of 100 members (light grey) and two simulations of interest (dark grey) are compared with simulations without noise at the high steady-state toxin concentration (red) or the low toxin concentration (green). The concentrations of the first toxin (a) and second toxin (b) are different, but they switch at approximately the same time. (c) The growth rate is a function of both toxins, and spontaneously switches between normal (fast-growing) and persistent (slow-growing) states.


