MATHEMATICAL APPROACHES TO INVESTIGATE ENERGY DEPENDENCE IN GENETIC PATHWAYS INVOLVED WITH ANTIMICROBIAL RESISTANCE

by

RYAN DAVID THOMAS KERR

A thesis submitted to
University of Birmingham
for the degree of
DOCTOR OF PHILOSOPHY

School of Mathematics
College of Engineering and Physical Sciences
University of Birmingham
May 2022

UNIVERSITY OF
BIRMINGHAM

University of Birmingham Research Archive

e-theses repository

This unpublished thesis/dissertation is copyright of the author and/or third parties. The intellectual property rights of the author or third parties in respect of this work are as defined by The Copyright Designs and Patents Act 1988 or as modified by any successor legislation.

Any use made of information contained in this thesis/dissertation must be in accordance with that legislation and must be properly acknowledged. Further distribution or reproduction in any format is prohibited without the permission of the copyright holder.

Abstract

Antimicrobial resistance is becoming increasingly common and is causing a global health crisis. Microorganisms have developed numerous sophisticated resistance mechanisms contributing to the reduced efficacy of antimicrobials. Combined with slow development of new antimicrobial drugs (including antibiotics) the possibility of a post-antimicrobial era is growing. To combat this global issue, we need to develop new treatments for microbial infections, and engineering novel targets could be a solution.

Genetically identical cells can display significantly different phenotypes in the same environment. Developing our understanding of how bacterial cells support and generate different phenotypes without changes to their DNA is critical in combatting antimicrobial resistance, as some cell fate decisions (such as generating persister cells) enable cells to tolerate high concentrations of antibiotics and outlast antibiotic treatment. Consequently, persister cells, and other antimicrobial resistance phenotypes, reduce the efficacy of antibiotics, and increase the difficulty of eradicating infections.

To do this, we first need to better understand how cells in an isogenic population generate different phenotypes. Functionally vital cell fate decisions from a range of phenotypic choices are made by regulatory networks, the dynamics of which rely on gene expression and hence depend on cellular energy availability, particularly ATP levels. However, despite pronounced cell-to-cell ATP differences observed across biological systems, the influence of energy availability on regulatory network dynamics is often overlooked as modulating cellular decision-making and phenotypic variation, limiting our knowledge of how energy budgets affect cell behaviour.

This study will introduce and analyse the behaviour of gene regulatory networks regulating cellular decisions when cell-to-cell ATP variability is considered. Results for a simple generalisable network will be presented, and then the model will be extended to

develop a more scalable and general modelling framework for analysing the effects of energy availability within a clonal population on the dynamics of regulatory networks. This framework is then applied to a gene regulatory network underlying a clinically-important antimicrobial resistance mechanism.

An alternate mathematical modelling investigation into antimicrobial resistance will then be introduced. In this approach, a framework called HyperTraPS is applied to resistance profiles of *Pseudomonas aeruginosa* samples extracted from patients being treated in hospitals located in Mexico. Results from this framework allow the inference of possible pathways leading to the acquisition of resistance to a set of antibiotics commonly used in the treatment of this clinically-relevant bacterial species.

Throughout this study we will demonstrate that energetic differences between cells may be an important consideration to help explain observed variability in cellular decision-making across biological systems, including those decisions promoting resistance, and provide a novel target for future therapeutics.

ACKNOWLEDGEMENTS

I must start by thanking my supervisors, Sara Jabbari and Iain Johnston, for their consistent support and enthusiasm throughout my PhD. It truly has been an amazing experience working with you both, and I have learnt from you in ways that will help me throughout my entire career. I would also like to thank the Wellcome Trust and University of Birmingham for their funding, and the opportunity to be part of the antimicrobials and antimicrobial resistance doctoral training programme. I have been enrolled onto the training programme with a great cohort, who have been supportive and helpful whenever it was required, and will be life-long friends.

I have been fortunate to have supportive friends and family around me during this period of my life. In particular, my parents Anne and Chris, sister Gabriella, girlfriend Jessica and her family, Tracey, Richard and Lucy. During these last four years I have lost my grandparents Ann and David, who were always proud of what I achieved, and I miss them dearly. It is to all of these people that this work is dedicated, and I am eternally grateful for their support.

Publications

Intracellular energy variability modulates cellular decision-making capacity. Scientific Reports, 9(1), 1-12 (2019)

Kerr, R., Jabbari, S., & Johnston, I. G.

Contribution: Performed all simulations, analysis and wrote the manuscript.

Dynamic Boolean modelling reveals the influence of energy supply on bacterial efflux pump expression. Journal of the Royal Society Interface, 19(186)

Kerr, R., Jabbari, S., Blair J. B., & Johnston, I. G.

Contribution: Performed all simulations, analysis and wrote the manuscript.

CONTENTS

1 Introduction **1**

1.1 Antimicrobial resistance 1

1.2 Existing mathematical contributions to combatting AMR 4

1.2.1 Mathematical modelling of mechanisms that reduce antibiotic efficacy 4

1.2.2 Mathematical modelling of efflux pumps 6

1.2.3 Mathematical models of gene regulatory networks controlling efflux pump genes 8

1.2.4 Data-driven studies of AMR evolution 10

1.3 ATP and cellular free energy 12

1.3.1 The molecular currency of the cell and its influence on cellular processes 12

1.3.2 Measuring single-cell ATP 12

1.3.3 Cell-to-cell variability in ATP 13

1.4 Thesis outline 14

2 Intracellular Energy Variability Modulates Cellular Decision-Making Capacity **16**

2.1 Introduction 16

2.2 Methods 19

2.2.1 Regulatory architecture 19

2.2.2 Energy dependence 22

2.3 Results 25

2.3.1 Intracellular energy availability modulates cellular decision-making landscapes . . . 26
2.3.2 Modelling the energy parameter through a linear function . . . 38
2.3.3 Cooperativity effects on system behaviour . . . 39
2.3.4 Regulatory protein binding strength can limit system behaviour . . 39
2.3.5 External ATP effects on decision-making landscapes . . . 45
2.3.6 Stable steady-state transitions . . . 46
2.4 Experimentally exploring the degree of ATP variability in real *E. coli* cells 49
2.5 Discussion . . . 54

3 Dynamic Boolean modelling reveals the influence of energy supply on bacterial efflux pump expression 59
3.1 Heterogeneity in efflux pump expression . . . 60
3.2 Methods . . . 62
3.2.1 Literature-based model construction accounting for energy availability 62
3.2.2 Wiring diagrams . . . 65
3.2.3 Initial conditions . . . 68
3.2.4 Modulation of transition rates allows energy variability to be captured in GRN models . . . 68
3.2.5 Node update function . . . 70
3.2.6 Node update methods . . . 71
3.2.7 Methods to analyse the Boolean network . . . 72
3.3 Results . . . 74
3.3.1 Efflux network models reproduce experimentally observed behaviour in *E. coli* and *Salmonella* . . . 74
3.3.2 Stress-free efflux network behaviour with a maximal cellular energy budget . . . 82
3.3.3 Efflux network behaviour with a maximal cellular energy budget after exposure to a stressor . . . 93

3.3.4 Efflux network behaviour is modulated by available cellular energy . 96
3.3.5 Energy levels influence heterogeneity in individual expression level dynamics 97
3.3.6 *E. coli* and *Salmonella* efflux networks share functional similarity . 99
3.4 Discussion 101

4 Identifying evolutionary pathways of antibiotic resistance in *P. aeruginosa* using Bayesian modelling 106
4.1 Introduction 106
4.2 Methods 108
4.2.1 Methodology overview 108
4.2.2 Application of Bayesian inference framework to *P. aeruginosa* data 112
4.3 Posterior orderings for the Test and Resistance Datasets 115
4.4 Discussion 119

5 Conclusions 122

List of References 127

LIST OF FIGURES

1.1 A timeline of key events: Year of clinical deployment and reported resistance for antibiotic classes. 2

2.1 Energy dependence in model genetic architectures. 20
2.2 Model relationship between energy availability and gene expression rates. . 23
2.3 Increasing energy levels support more distinct attractors. 27
2.4 Attractor basins with increasing energy availability A^* and maximum expression level from conditional promoter. 28
2.5 Attractor basins with increasing energy availability A^* and constitutive expression level. 29
2.6 Attractor basins with increasing energy availability A^* and degradation level. 30
2.7 Re-entrant behaviour in decision making options. 31
2.8 Destabilisation of central attractor in re-entrant behaviour. 32
2.9 Decision-making capacity depends on cellular energy and dynamic parameters. 33
2.10 Decision-making landscapes depend on energy availability and dynamic rate parameters. 35
2.11 Re-entrant behaviour showing destabilisation of the 'near-zero' attractor. . 36
2.12 Re-entrant behaviour showing destabilisation of the central attractor. . . . 37
2.13 Linear energy function displays similar qualitative decision-making landscapes to the default case. 40

2.14 Reduced Hill coefficient value ($n = 3$) displays similar qualitative decision-making landscapes to the default value ($n = 4$). 41
2.15 Hill coefficient $n = 2$ displays simplified qualitative decision-making landscapes to the default value ($n = 4$). 42
2.16 Non-cooperative binding limits the decision-making landscape. 43
2.17 Strength of binding limits the decision-making landscape. 44
2.18 Model relationship between energy availability and gene expression rates with external ATP influences. 46
2.19 Decision-making landscape variability is observed when considering external ATP influences. 47
2.20 The ability to transition between extreme states becomes harder as energy availability increases. 48
2.21 Sample mounting preparation protocol for microscopy. 50
2.22 Segmentation of microscopy images. 52
2.23 Clonal *E. coli* populations display cell-to-cell heterogeneity in a stress-free environment. 53

3.1 Boolean network modelling of gene regulatory network architectures controlling efflux pump gene expression. 63
3.2 Schematic representation of network wiring diagrams. 67
3.3 Mean expression dynamics for *E. coli* from the unstressed initial condition in response to a short or long stress period. 83
3.4 *E. coli* mean expression dynamics from the stressed initial condition. . . . 84
3.5 *E. coli* expression level coefficient of variation. 85
3.6 Mean expression dynamics for *Salmonella* from the unstressed initial condition in response to a short or long stressor duration. 86
3.7 *Salmonella* mean expression dynamics from the stressed initial condition. . 87
3.8 *Salmonella* expression level coefficient of variation. 88

3.9 *E. coli* displays significant heterogeneity in *marRA* gene expression in the absence of stress. 89
3.10 Simulations predict *E. coli* expresses *acrAB* in heterogeneous pulses in the absence of stress. 90
3.11 *Salmonella* displays significant heterogeneity in *ramA* gene expression in the absence of stress. 91
3.12 Simulations predict *Salmonella* expresses *acrAB* in heterogeneous pulses in the absence of stress. 92
3.13 *E. coli* transition dynamics depend on cellular energy and the presence of stress. 95
3.14 Energy availability and stress modulate *E. coli acrAB* pulse variability. . . 98
3.15 *Salmonella* displays analogous coarse-grained qualitative behaviour compared to *E. coli*. 100

4.1 Coarse-grained illustration of the HyperTraPS pipeline. 108
4.2 Antibiotic resistance profiles of 92 *P. aeruginosa* isolates. 113
4.3 Inferred pathways of antibiotic resistance acquisition in *P. aeruginosa* isolates. 117
4.4 Inferred acquisition order for drug resistance in *P. aeruginosa* clinical isolates. 118

LIST OF TABLES

3.1 *E. coli* node initial values. 68

3.2 *Salmonella* node initial values. 68

3.3 Theoretical model reproduces realistic behaviour in *E. coli* and *Salmonella*. 75

3.4 Summary of energy availability effects on regulatory network behaviour. . . 102

4.1 Antibiotics grouped by antibiotic class. 112

CHAPTER 1

INTRODUCTION

1.1 Antimicrobial resistance

Antimicrobial resistance (AMR) is a huge global health challenge, contributing to over 700,000 deaths worldwide [208]. AMR is when a microorganism is able to circumvent antimicrobial treatment through intrinsic, acquired, or adaptive mechanisms (for example, bacteria can become resistant to antibiotics). It has been estimated that the number of deaths caused by antimicrobial-resistant bacteria, if left neglected, could surpass deaths attributed to cancer by 2050 [209]. The accelerating issue of AMR resulted in the World Health Organization (WHO), in coordination with the Food and Agriculture Organization of the United Nations, endorsing a global action plan to help tackle it in 2015 [210].

The 'Golden Age' of antibiotic discovery, between the 1940s and 1960s, led to a healthy production of effective antibiotics for patient use, providing enough options to still effectively treat resistant infections [295] (see Fig. 1.1). However, the development of each new antibacterial drug was followed by the detection of resistance to it [295] (shown in Fig. 1.1). Although the detection of resistance to an antibiotic does not render it universally ineffective, it reduces the available treatments for a patient. Since these years of rich antibiotic discovery, the pipeline of new antibiotic discoveries has slowed down and the spread of AMR has accelerated, creating a progressive and growing threat to public health worldwide.

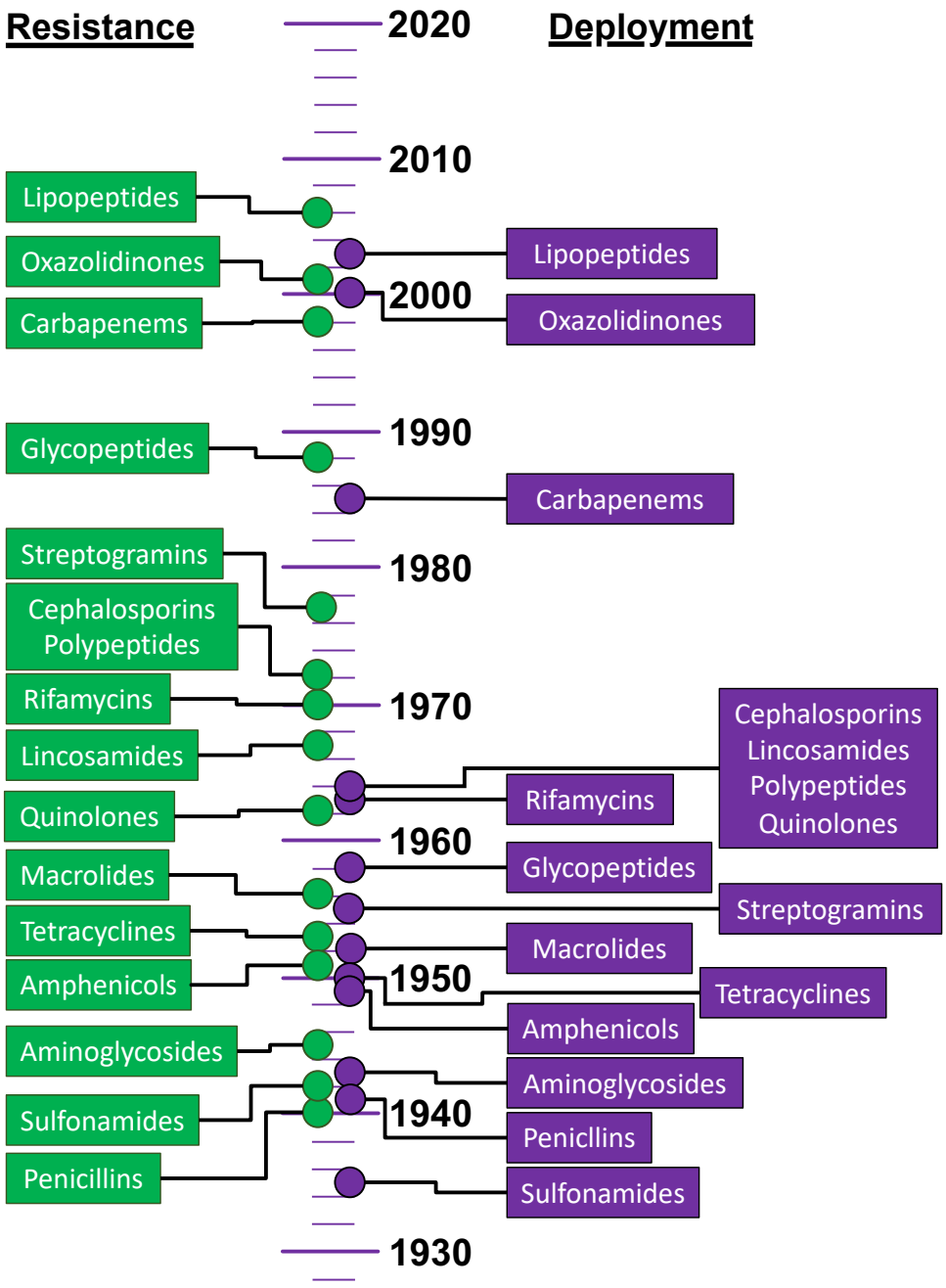


Figure 1.1: **A timeline of key events: Year of clinical deployment and reported resistance for antibiotic classes.** Timeline displaying the approximate year of clinical deployment (purple) and observed resistance (green) for several antibiotic classes; based on information from [68, 83, 94, 140, 249, 287, 303]. Although new generations of current antibiotic classes are still engineered, only a few new classes have been discovered since the mid-1960s. Notes: (1) The observation of penicillin resistance pre-dated clinical deployment [3]; (2) In the presence of conflicting data, the earliest date was selected.

Mathematical modelling is an underused and powerful tool that can aid in the fight against AMR. Traditionally, studies aimed at tackling AMR have been purely empirical, using a "reductionist" approach and focussing on particular elements within resistance mechanisms or pathways. Employing theoretical models can provide insights into complex AMR processes that would otherwise be significantly more difficult, or even impossible, to investigate experimentally. By considering a resistance system as a whole, theoretical frameworks can analyse how the system operates, rather than considering effects of a reduced subset of elements within the system. Mathematical analyses can therefore complement traditional experimental biology, providing potential sources of new targets by exploiting the knowledge of the system as a whole [238, 239, 272]. Integrating theoretical and experimental research as a combined strategy may offer an accelerated approach to combat AMR.

To understand how mathematical modelling can help tackle AMR, we briefly survey several resistance mechanisms, providing some context and focussing down on particular modelling-related case studies. Contained within this is a detailed discussion on efflux pumps, a mechanism that confers intrinsic (inherent properties of the bacterium) and acquired (a previously sensitive bacterium acquires a resistance mechanism) resistance to antibiotics, and is investigated later in this thesis. Following this, a summary of data-driven studies of AMR evolution is provided, exhibiting the range and volume of AMR-related data that is accessible, and how it can be employed within computational frameworks. The chapter is concluded by examining the dependence of ATP (the molecular energy currency of a cell) in regulatory processes, the variability of single-cell ATP levels within a genetically identical population, and the potential consideration of ATP as a novel AMR target.

1.2 Existing mathematical contributions to combatting AMR

Theoretical modelling can help the development pipeline of new antimicrobial treatments by advising on optimal drug doses and dose regimen [97, 159, 176, 271], investigating the probability of resistant mutants emerging in a bacterial population when exposed to various dosing regimens [19, 105, 202] and analysing the effectiveness of novel antimicrobial treatments [238, 239, 272]. Results from these models can provide an explicit theoretical basis to advise on the development and evaluation of new or current treatments. This could, for example, predict the conditions under which a treatment might be expected to eradicate the infection, or if a combination of multiple antibiotics is a more effective treatment strategy. An accurate framework could also analyse the compound modifications that make a substance more potent, or evaluate different drug doses, significantly quicker than testing experimentally. Employing theoretical models could accelerate the current drug pipeline, a requirement to tackle the current spread of AMR.

1.2.1 Mathematical modelling of mechanisms that reduce antibiotic efficacy

Mechanisms of AMR in bacteria generally fall into three major categories: target modification, drug inactivation and drug transport. Combatting each mechanism involved with reducing antibiotic efficacy is important in the battle against AMR. In this section we will present mathematical models of a selection of AMR mechanisms. The discussed mechanisms of antimicrobial resistance are in no part suggested to be of more importance than those not reviewed (spontaneous mutations or modification of drug targets, for example [33]), and mathematical models of efflux pumps will be discussed in the section that follows.

Numerous resistance mechanisms can be transferred between bacteria through horizontal gene transfer (HGT), a process of acquisition of new genetic material from an

external source or by taking up free fragments of DNA, often in the form of plasmids. HGT plays a crucial role in the development of multidrug resistance (MDR) in bacteria [84], and contributes significantly to the rapid spread of resistance in clinical environments [143]. The accumulated antibiotic resistance genes (ARGs) via HGT may be involved in different resistance mechanisms, increasing the number of AMR mechanisms a bacterium can utilize, and the severity of resistance. Mathematical models have previously investigated HGT processes in natural bacterial communities, assisting the development of monitoring strategies [86, 192, 259]. Gama *et al.* [86] developed an ordinary differential equation (ODE) model analysing plasmid maintenance in a bacterial population when there are multiple interacting plasmids. Simulations of the model predict the conditions required for plasmid survival in bacterial populations, for example, if they provide a fitness advantage. It is important to understand how interference between plasmids determine their stability in bacterial populations, as this can help guide the development of novel treatments tackling AMR through HGT.

Decreased permeability of the bacterial outer membrane reduces the influx of antibiotics [82, 196, 289]. Limiting antibiotic entry into the bacterial cell in combination with efflux of the drug is a well-established mechanism of intrinsic antibiotic resistance. This combination limits the interaction of the drug with its intracellular target, reducing efficacy. Mathematical modelling has quantitatively described a passive influx process in bacteria, estimating the optimal permeability coefficient to ensure passive permeation into replicating bacterial cells [190]. This coefficient can advise the development of novel antibacterial agents by advancing the understanding of drug penetration into Gram-negative and Gram-positive bacteria. This framework can also be adapted to quantitatively analyze the interplay between passive inwardly directed permeation and active efflux [190]. Quantitative modelling can be used alongside experimental research to guide drug development or discovery, advising on compound properties to more effectively penetrate the cell and reach the target site.

Bacterial persistence is a transient tolerance to external stress, including lethal doses

of antibiotics, by a small fraction (roughly 10^{-5} to 10^{-3} %) of an otherwise susceptible bacterial population [135, 149]. Persister tolerance to antibiotics is different to resistance as a persister cell is genetically identical to a susceptible cell in the population. Persisters can stochastically switch to the wild-type phenotype, subsequently growing a new population. This attribute is why they are considered a leading factor in the recalcitrance of chronic infections [80], including oral thrush [136], lung infections in cystic fibrosis patients [182] and tuberculosis [317]. Although there is currently no general mechanism for generating the persister phenotype, numerous mechanisms have been identified, including toxin/antitoxin modules [58, 66, 128, 284], SOS response [308], (p)ppGpp [64, 91, 134, 257] and more recently, an ATP/energy-related mechanism [56, 153, 184, 251, 315]; specifically, results suggest low energy is sufficient for tolerance [56, 251, 315]. Persister-related mathematical models have investigated toxin-antitoxin (TA) modules [89, 157, 244], the (p)ppGpp biochemical network [8] and population level dynamics [21, 42, 144]. Investigating a proposed mechanism of persister formation through a mathematical framework facilitates the investigation of a system as a whole, helping experimental investigations of persister cells by predicting key elements that significantly modulate the level of persisters [89, 157, 244]. From mathematical studies of TA module regulation, for example, it was proposed that increasing the toxin level above a threshold results in bacterial dormancy [89, 244]; this prediction was later validated experimentally [244]. Mathematical frameworks can help facilitate the detection of key elements that could be used as anti-persister targets.

1.2.2 Mathematical modelling of efflux pumps

Efflux pumps are protein structures that are located in the cytoplasmic membrane. These protein structures have been observed in numerous Gram-positive and Gram-negative bacterial species, including *Streptococcus pneumoniae* [92], *Acinetobacter baumannii* [1], *Pseudomonas aeruginosa* [151], *Klebsiella pneumoniae* [160], *Haemophilus influenzae* [246] and *E. coli* [90]. Efflux pumps are able to expel antibiotics from the intracellular to the

extracellular environment of a bacterium, decreasing cellular drug accumulation. This function, enabling the expulsion of drug-specific compounds [197, 224] or multiple unrelated substances [199, 224], is why they are a mechanism of drug resistance. Efflux pumps that excrete a wide range of compounds can result in multidrug resistance (MDR) [146, 198].

Numerous mathematical models studying the effect of efflux pumps on bacterial drug resistance have captured this mechanism through systems of ODEs [37, 158, 220, 305, 314]. The models often aim to describe the change in intracellular antibiotic concentration by considering efflux pumps as a mechanism of expelling drug compounds. Establishing the relationship between certain input parameters and model outcomes within mathematical models, by perturbing model parameters, can provide an insight into the elements that have a larger effect on resistance; theoretically computing this is significantly quicker than performing the equivalent experiments. Analysing this information can complement experimental biology by providing potential new drug targets. Accurate models can also be used to extrapolate beyond situations for which experimental data is not available.

Perez *et al.* [220] applied an ODE framework to model the major facilitator superfamily efflux pump, TetB, in the Gram-negative bacterium *E. coli* (each cell has an inner and outer membrane). The major facilitator superfamily efflux pumps span one membrane, and TetB is responsible for resistance to tetracycline antibiotics. The model considered substrate diffusion across both the inner and outer cell membranes, with TetB effluxing the substrate from the cytoplasm to the periplasm. The study calculated the steady states of the system of equations for the concentration of substrate in the cytoplasm, to then determine an equation relating growth rate to the substrate concentration in the media surrounding the *E. coli* cells. By applying the model to fitness data of strains expressing a chromosomal copy of *tet(B)* at a wide range of tetracycline concentrations, binding constants of the substrate to the transporter as well as a parameter relating to pump efficiency could be estimated. We highlight this study as an example of applying a mathematical framework to experimental data to predict parameters that would otherwise be

more difficult to calculate experimentally. The prediction of parameter values or relationships between model parameters may help advise the development of adjuvants that block, or reduce, the efflux of antibiotics, increasing the potency of antibiotic treatments.

Bootsma *et al.* [37] presented a model to analyse non-inherited antibiotic resistance levels in bacteria through different means of physiological adaptation, including efflux pumps. The model described efflux pump density, intracellular antibiotic concentration and antibiotic MIC-value dynamics to observe the development of antibiotic resistance caused by physiological adaptations (without genetic changes). This theoretical framework predicted the time it takes (after stopping exposure to an antibiotic) for a cell's MIC-value to decrease to a susceptible level [37]. This data can be used to advise on the optimal dosage regimen that increases the efficacy of the drug and reduces the risk of developing resistance.

1.2.3 Mathematical models of gene regulatory networks controlling efflux pump genes

Diverse resistance mechanisms are under the control of a gene, or signalling, regulatory network. Each regulatory network consists of a group of genes, proteins, enzymes etc. which interact to modulate the expression of target genes. Gene expression variability leads to cell-to-cell variations in mRNA and protein levels in genetically identical cells and can drive the generation of diverse phenotypes and strategies for survival [34, 59, 85, 123, 212]. This phenomena, where a cell can realise different cell fates, is known as cellular decision-making, and provides bacteria with a strategy to overcome environmental stressors by generating phenotypic diversity, for example, the generation of persister cells [244] and sporulation [223]. Mathematical frameworks have previously been applied to model important regulatory networks involved in drug resistance [46, 74, 88, 244]. By understanding the dynamics of these networks, mathematical models can determine the influential factors in these cellular processes and the effects when these processes are dysregulated. Further, regulatory models can be extended to observe the effect of new

or hypothetical elements, a task that can be more difficult to accomplish in experimental research.

The expression of efflux pumps is under the control of various transcriptional regulators which, together with the genes encoding them, form a gene regulatory network (GRN) modulating the expression of efflux pump genes [10, 32, 201, 264]. Efflux pump GRNs are complex, yet those that are known are still considered to be incomplete. Additionally, individual components within a network may perform other functional roles within a bacterium. For example, the multidrug resistance regulators MarA, SoxS, Rob, and RamA in *Salmonella enterica* serovar Typhimurium also, either directly or indirectly, regulate flagellar gene expression and motility [276]. In *E. coli*, the RND pump regulator MarA has been shown to also regulate lipid trafficking and DNA repair [252].

Mathematical modelling of efflux pump GRNs enables a deeper insight into the behaviour of efflux pumps that is otherwise difficult, or impossible, to investigate experimentally. Each regulatory network is often determined through genome sequencing, computational predictions, and experiments [124, 142, 172, 177], and are often part of a larger, inter-connected regulatory network. One such example is Youlden *et al.* [314], where an ODE model of the GRN governing the expression of efflux genes *acrAB acrEF* in *Salmonella* is presented, and a complex asymptotic analysis is used to tackle the lack of quantitative information on the resulting large parameter set. They also calculated asymptotic approximations of the steady states, for both wild-type and *ramR* mutant (a bacterium with an alteration to the DNA sequence of *ramR*) cases, allowing an insight into the dominant behaviours that control the GRN. Together, with identifying parameter groupings that have the most effect on the expression of efflux genes, the asymptotic analysis can provide potential targets that can be tested experimentally in order to combat efflux related MDR.

The role of efflux pumps in MDR, pathogenesis and biofilms, makes them an important therapeutic target. To combat efflux pump related AMR, molecules that inhibit efflux pump expression have been considered as potential therapeutic agents. Efflux pump

inhibitors (EPIs) are small compounds that can work in conjunction with antibiotics that would otherwise not be effective against bacterial pathogens [189, 219, 265, 311, 319]. Developing our understanding of regulatory pathways involved with antibiotic efflux, may enable novel targets and treatments to be developed.

1.2.4 Data-driven studies of AMR evolution

The development of faster genome sequencing and phenotypic susceptibility testing across microbiology has enabled increasing volumes of real-time data to be accessible, aiding earlier and more accurate diagnoses of resistant bacterial infections [48, 65, 76]. Previous work has produced significant volumes of genomic data that can be utilized in studies of AMR, including genetic mutation data, the evolutionary relatedness of clinical isolates through phylogenomic analysis, and the presence of acquired antimicrobial resistance genes [17, 43, 93, 137, 232, 234]. Lam *et al.* [137] developed a tool for analysing genomes of *K. pneumoniae* (and its associated species complex) and determine the relatedness of the isolates. The ability of this tool to rapidly and accurately genotype large-scale clinical isolate data, to identify clinically relevant features including genetic mutations and the prevalence of AMR genes, provides a framework that can supply a detailed evolutionary insight into a clinically relevant bacterial species, and support genomic surveillance including highlighting global trends in *K. pneumoniae* AMR.

A study by Casali *et al.* [43] conducted whole-genome sequencing of 1,000 patient isolates from Russia to determine the transmissibility and prevalence of drug-resistant tuberculosis in a population. Whole-genome sequencing enabled the tracking of evolutionary mechanisms that promote the development and transmission of drug resistance within pathogen populations by identifying loci that are subject to frequent mutation within the isolates, and polymorphisms that could contribute to resistant phenotypes. Following this, a maximum-likelihood phylogeny was reconstructed, presenting the evolutionary relationships among the isolates, which shows the genetic diversity in the isolates and the global lineage of the isolates. The phylogenetically linked data can also be har-

nessed to learn about the dynamics of evolution and disease progression [120, 121, 307], including the evolution of MDR in tuberculosis [95].

Developing computational frameworks that can utilize the abundance of real-time data to predict the likelihood of evolving resistance to individual antibiotics, can be useful tools for clinicians. These tools could suggest appropriate antibiotics based on current isolate data, supporting clinical decisions, and complementing current diagnostic methodologies. This could enable advanced, real-time treatment for patients, improving their treatment and accelerating recovery. Advancing diagnostic frameworks, that involve both computational and experimental developments, could be a powerful tool in combatting antimicrobial resistant infections.

Of particular importance is the development of frameworks that can be applied to bacterial strains stated to be of critical priority for the research and development of new and effective antibiotic treatments by the WHO. Included in the critical priority list is *Pseudomonas aeruginosa*, a Gram-negative, opportunistic, bacterial pathogen associated with a broad spectrum of infections in humans. Developing new and effective antibiotic treatments, and reducing AMR of this bacterium is crucial, due to its intrinsic resistance to several classes of antibiotics, the ability to develop resistance to last-line therapeutic options [267], and the ability of this microorganism to persist within hospitals (contaminating floors, bed rails and sinks [50]). Utlizing the large volumes of phenotypic and genotypic data of AMR bacteria (such as *P. aeruginosa* [93]) within computational frameworks may improve patient outcomes, and decrease the spread of infections caused by multidrug-resistant organisms, through data-driven decisions on prescribing therapeutics.

1.3 ATP and cellular free energy

1.3.1 The molecular currency of the cell and its influence on cellular processes

Numerous mechanisms promoting the resistance of antimicrobials are controlled by an underlying GRN. The regulation of gene expression within each regulatory network relies on several processes, including transcription, translation and post-translational modifications. Transcription and translation have been shown to contain a core dependence on ATP [62, 122], the principal molecule for storing and transferring energy in cells, and considered the "molecular currency" in all forms of living cells [98, 114]. Therefore, there exists a core ATP dependence in the dynamics of gene expression and gene regulatory networks.

The ATP-dependence within the core processes required for gene expression and regulation asserts this simple molecule as a possible target and area of exploration to tackle bacterial infections, including those that are resistant to antibiotics. Previous work has proposed the development and use of novel or pre-existing compounds targeting ATP-related machinery, including the enzyme producing ATP in bacteria, called the ATP synthase [52, 138, 141]. These approaches would be required to selectively target pathogenic bacteria without harming non-pathogenic bacteria and human cells. In theory, this could eradicate a bacterial infection without toxicity to human cells. Collectively, targeting ATP is a plausible and novel opportunity for treating both susceptible and resistant microbial infections. ATP's role in cells is significant, as survival is impossible without it. Aiming to develop novel therapeutics focussed on targeting ATP may have far-reaching effects in the future of medicine.

1.3.2 Measuring single-cell ATP

The continuous development of methods to quantify single-cell ATP concentrations, and further understanding intracellular ATP variability is essential. The development of

genetically-encoded biosensors in the last two decades has transformed the analyses of important intra- and extracellular metabolites. Indicators that quantitatively measure intracellular metabolites involved with cellular energy, such as ATP, at a single-cell level have only been developed in recent years. This new technology has generated numerous sensors for use in both prokaryotic and eukaryotic cells, advancing the understanding of heterogeneity of metabolites in genetically-identical cells.

The genetically-encoded protein-based sensors are a relatively non-invasive way to accurately measure intracellular ATP in living cells. Prior to the development of the indicators called ATeams [111], conventional ATP quantification methods contained several limitations, including pH sensitivity, or outputs provided the averaged ATP level of an ensemble of cells. Since the development of ATeams, other novel ATP biosensors have been engineered, including QUEEN [309], BTeam [313] and ARSeNL [179]. Additional biosensors have been based on previous iterations of proven ATP indicators, improving upon indicator limitations or targeting alternative metabolites [38, 154, 185, 280]. Continuing to develop novel, or re-engineered, biosensors will improve the direct, quantitative measurements of intracellular ATP dynamics.

1.3.3 Cell-to-cell variability in ATP

The application of ATP biosensors has demonstrated that identical cells within a population can display substantial variability in ATP levels. In prokaryotes, clonal populations of *E. coli* [309] and *S. aureus* [315] have displayed significant single-cell ATP variability. The quantified range of intracellular ATP can cover almost an order of magnitude, varying between 0.32-2.76 mM in *E. coli* [309], for example. Further, an *E. coli* population contained heterogeneous cell-to-cell ATP levels in each observed growth phase [242], indicating it is not sub-populations at different stages in the growth cycle leading to the population level ATP variability. Instead, it has been suggested variability in ATP levels is a result of stochastic fluctuations in the abundance of energy-generating components [166, 309]. Intracellular ATP variability has also been observed in other branches of life, including

yeast [269, 270], plants [63] and human cell lines [154, 235, 313]. Continuing to improve the current ability of biosensors will allow the magnitude of cell-to-cell heterogeneity to become clearer.

The core ATP-dependence within the fundamental processes required for gene expression and regulation, and the substantially different levels of available energy (linked to the level of intracellular ATP) suggest the dynamics, and thus potentially the outputs, of regulatory networks involved with resistance may vary between cells due to these diverse ATP levels.

1.4 Thesis outline

In Chapter 2, we introduce a framework to investigate the role of intracellular energy in decision-making capacity, which is the process of a cell supporting different numbers of functionally important fates, without genetic modifications. The motivation, and application, of the framework considers both prokaryotic and eukaryotic cell types, allowing an investigation into the potential role of intracellular energy in decision-making capacity more generally.

Chapter 3 develops a more scalable framework to explore the influence of ATP variability on GRN dynamics. This framework, using a coarse-grained Boolean network modelling approach, modified to integrate ATP-dependence within GRN processes, is applied to two GRNs controlling the expression of efflux pump genes in *E. coli* and *Salmonella*. Using these powerful tools, this thesis aims to advance the current understanding of the roles and influences of single-cell ATP variability on cellular decision-making and AMR-related mechanism behaviour.

Finally, in Chapter 4, we explore an alternative method to investigate the global AMR issue. We apply a Bayesian inference methodology to predict pathways leading to AMR through the acquisition of resistance to several antibiotics in the opportunistic pathogen *P. aeruginosa*. By applying this inference tool, we are able to investigate the development of

antimicrobial resistance, offering a different insight into the evolution of AMR in bacteria, which could allow patient treatment based on the current antimicrobial resistance profile status of the clinical strain causing the resistant infection.

Overall, this research expands the current knowledge of how ATP dynamics modulate a decision-making mechanism at the single-cell level. Additionally, the theoretical frameworks are used to identify how energy availability modulates regulatory network behaviour, and advocate further investigation into the development of ATP-related antimicrobial therapies as novel AMR treatments.

CHAPTER 2

INTRACELLULAR ENERGY VARIABILITY MODULATES CELLULAR DECISION-MAKING CAPACITY

Cells generate phenotypic diversity both during development and in response to stressful and changing environments, aiding survival. For example, the generation of persister cells in a genetically identical bacterial population can reduce the efficacy of antimicrobials, and developing multicellular organisms make decisions to follow different developmental pathways, and hence towards one of a diverse range of possible phenotypes. Furthering our understanding of how a clonal population displays intrinsic phenotypic heterogeneity may lead to the development of novel methods to regulate this behaviour. In this chapter, we consider a mathematical model of a highly generalisable, ATP-dependent, decision-making regulatory network, and observe how cell-to-cell ATP variability modulates the set of phenotypes a cell can support.

The theoretical work in this chapter is taken from research published by Kerr *et al.* [131].

2.1 Introduction

Biological cells are faced with many decisions during their existence. Genetically identical single cells in a population choose different phenotypic strategies for survival. Across

organisms, gene expression variability leads to cell-to-cell variations in mRNA and protein levels in genetically identical cells and can drive the generation of diverse phenotypes and strategies for survival [34, 59, 85, 123, 212]. The process of a cell assuming different functionally important fates from a range of phenotypic possibilities in response to or in anticipation of extracellular change, without genetic modifications, is known as a cellular decision [39].

In multicellular organisms, phenotypic heterogeneity has been observed in a diverse range of cell types [22] for a wide range of cellular decisions, from seed germination [118, 181, 279] through the famous example of hematopoietic cell differentiation [45, 70], to mosaic development of retinal cells in the *Drosophila* eye [306]. Waddington's famous 'epigenetic landscape' [292] pictures these developmental decisions as bifurcating channels that a developmental 'ball' can roll down to select different possible cell fate decisions; bifurcations in the landscape correspond to multistability, where a cell can support distinct, differentiated cell fates. These repeated differentiation decisions allow, for example, human pluripotent stem cells to differentiate into all cell types in the human body [214, 273], while modern technology allows differentiated cells to move back 'up' the epigenetic landscape [214, 268]. These cellular decisions are central to development and knowledge of their dynamics offers useful applications in medicine and fundamental biology [165].

Single-celled organisms also embrace the advantages of diverse cell behaviours. Cells in their natural environment have to deal with the challenges presented by changes in extracellular conditions. These may include temperature changes, pH variability, nutrient limitation or, in some cases, the presence of antibiotics. To overcome such environmental fluctuations, genetically homogenous cells can generate phenotypic diversity in order to increase the probability that some members of the population will survive [110, 147, 180]. This variability may enable a cell to support multiple, distinct phenotypes [4, 262, 306] potentially helping a cell handle different environmental conditions; the resulting population heterogeneity may even increase the overall fitness of the species [4, 255].

In microbiology, diversity and cell decision-making underlie a range of biologically

and medically important behaviours. Large-scale studies have revealed the generation of phenotypic diversity in yeast through noisy gene expression, potentially benefitting the population in varying and often stressful environments [34, 188]. The decision for bacteria to sporulate into robust endospores is a survival mechanism used by diverse genera including *Bacillus* and *Clostridium* [223], contributing to foodborne disease and food spoilage [9, 304]. The decision to become a persister cell (a phenotype more tolerant to external stress, including antibiotics), is made in several bacterial species [21, 53, 56, 135] and can have a dramatic impact on the efficacy of treatments [80, 149, 182, 317].

Many of the mechanisms behind cellular decision making in eukaryotes and prokaryotes remain poorly understood, challenging our knowledge of fundamental biology and our ability to design cellular interventions to suit human purposes. Regulatory networks, representing the interactions between genes that govern these decisions, are often used to summarise our knowledge [125]. Typically, edges in a schematic network illustrate processes such as transcription and translation and nodes represent genes. However, this coarse-graining can often omit a substantial amount of important detail. In particular, the fact that the processes represented by these edges are energy dependent (Fig. 2.1(a)-(b)) is rarely considered. Transcription and translation require a substantial ATP budget [62, 122], so there exists a core energy dependence in the dynamics of gene regulatory networks.

This energy dependence is important because different cells, particularly in microbiology, can have substantially different levels of available energy, potentially affecting the decisions supported by a given cell. Energy variability has been observed within genetically-identical *Escherichia coli* bacteria cells in a population, where absolute concentrations of intracellular ATP were spread over at least half an order of magnitude, in a skewed distribution around 1.54 ± 1.22mM [309]. Substantial intracellular ATP variability has also been observed in other branches of life, including yeast [270], HeLa cells [313] and plants [63]. Clearly, the dynamics, and thus potentially the outputs, of cellular decision-making networks may vary between cells due to these diverse energy levels.

Previous work has shown that energy availability can modulate the stability of decision-making circuits in eukaryotic cells [119], supported by experimental observations linking, for example, mitochondrial content and quality (supporting high, stable ATP supply) to less differentiated cell outcomes [126, 248, 273]. However, the broader influence of energy variability on the behaviour of these circuits remains to be revealed, and a theoretical understanding of this potentially important modulator of cellular decision-making remains absent. Here, we seek to address this challenge using a theoretical approach that both provides highly generalisable understanding and circumvents the experimental difficulties in interrogating these complex systems. We use a general regulatory network to model a genetic architecture observed as a decision-making motif across organisms. By including the effects of energy variability on physiological processes, we explore how the behaviour of the system will change when cell-to-cell variability exists in intracellular energy supply, and what effects result on the ability of cells to make diverse decisions.

2.2 Methods

2.2.1 Regulatory architecture

We consider a well-known cellular decision-making architecture consisting of two genes, their respective protein products, and the regulation occurring through their interactions [109, 119]. To study a general model of cellular decision-making we designate the genes 'Gene 1' and 'Gene 2', as displayed in Fig. 2.1(b). Protein 1 and Protein 2 (expression levels x_1 and x_2) are expressed by Gene 1 and Gene 2 respectively, which then apply feedback control to the system through self-activation and cross-repression; both proteins also degrade. We define a cell fate as the level of these proteins at steady state.

We construct the governing ordinary differential equations (ODEs) with guidance from the literature [109]. We coarse-grain gene expression dynamics, avoiding an explicit representation of mRNA levels, and focussing on the levels of the protein products, correlating

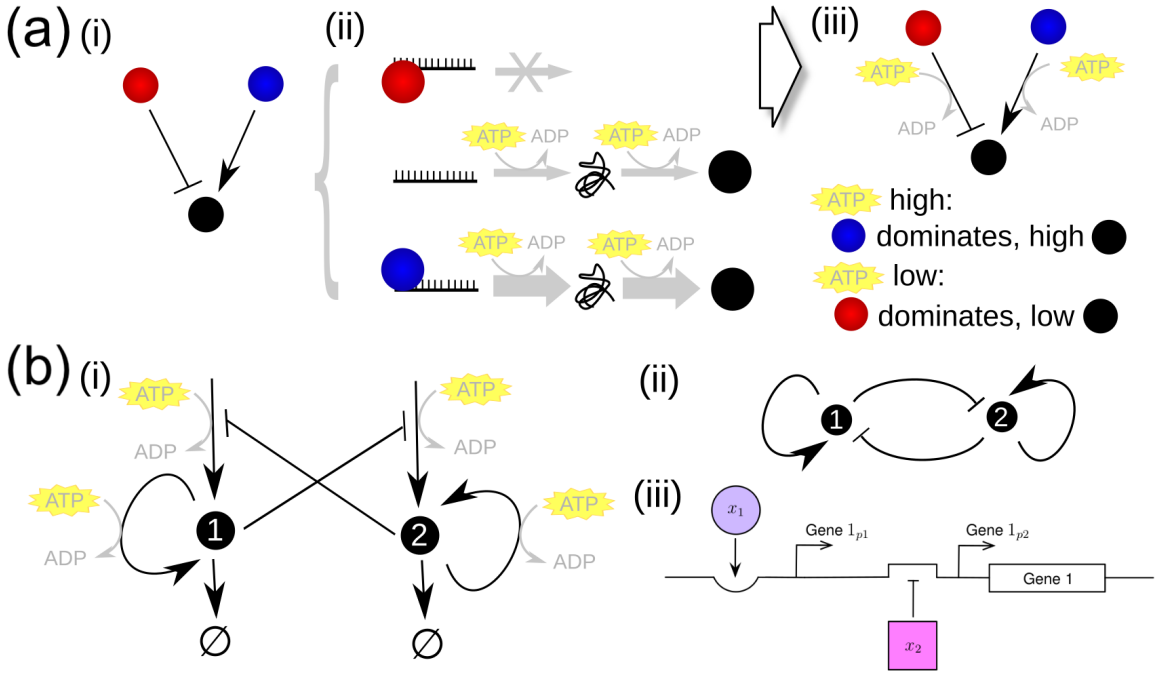


Figure 2.1: **Energy dependence in model genetic architectures.** **(a)** Energy dependence in gene regulation. In an example schematic regulatory network (i), genes have positive and negative regulatory interactions. The expression of these genes, and thus their regulatory interactions, rely on transcription and translation, processes with substantial ATP requirements (ii). The dynamics, behaviour, and states supported by a regulatory network are thus expected to be ATP-dependent (iii). In this example, when ATP levels are high, activation by a positive promoter (blue) may lead to high levels of expression of a downstream target (black). However, when ATP levels are low, there may be insufficient energy to support this increased expression level, and the system's behaviour is instead dominated by a repressive interaction (red). **(b)** The regulatory network for the general system considered in our model. The system involves two genes which self-activate and cross-repress, with both sets of processes being ATP-dependent (i). The energy dependence is often neglected in shorthand network representations of gene regulatory networks (ii). (iii) shows a possible arrangement of regulatory regions for one gene in the system.

with the expression level of their respective gene in our model. The expression of each gene has contributions from self-activation (via the gene's product activating a conditional promoter) and from a constitutive promoter (which can be repressed by the other gene's product), Fig. 2.1(b). For Gene i, the maximum expression level of the conditional promoter for Gene i is a_i, and the basal expression level of the constitutive promoter is b_i. Each protein product is degraded with rate k_i. Overall, we then obtain Equations (2.2.1)-(2.2.2) with initial conditions (2.2.3), previously studied by Huang et al. [109], with $x_{1,0}, x_{2,0} \in \mathbb{R}$ being some initial levels of proteins x_1 and x_2.

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = a_1 \frac{x_1^n}{\theta_{a_1}^n + x_1^n} + b_1 \frac{\theta_{b_1}^n}{\theta_{b_1}^n + x_2^n} - k_1 x_1, \tag{2.2.1}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = a_2 \frac{x_2^n}{\theta_{a_2}^n + x_2^n} + b_2 \frac{\theta_{b_2}^n}{\theta_{b_2}^n + x_1^n} - k_2 x_2, \tag{2.2.2}$$

$$x_1(0) = x_{1,0}, \ x_2(0) = x_{2,0}. \tag{2.2.3}$$

The additive terms in Equations (2.2.1)-(2.2.2) reflect, from left to right, self-activation up to a maximum level of a_i, cross-repression down from a basal level of b_i, and degradation. The interaction processes are represented by Hill functions, with n and θ parameters determining the steepness and inflection point of the sigmoidal curves, respectively; θ_{a_i}, θ_{b_i} can be interpreted as the dissociation constants of the activator and inhibitor regulatory proteins to the promoter regions, respectively.

Huang *et al.* suggest a default example parameter set [109], with $a_i = b_i = k_i = 1, \theta_{a_i} = \theta_{b_i} = 0.5, n = 4$ $(i = 1, 2)$. These parameters give rise to a tristable system, where, depending on initial conditions, the steady state can take one of three values. Interpreting these protein level values as cell fates, this tristability corresponds to three possible cell fates for an organism and therefore, for example, a wider decision-making landscape compared to a monostable system.

2.2.2 Energy dependence

Each step in transcription and translation requires energy in the form of ATP (Fig. 2.1(a)), so we enforce that the rates of the corresponding gene expression processes in our model are dependent on an ATP concentration parameter. To this end, we first transform the parameters a_i and b_i to $a_i \to \lambda a_i$ and $b_i \to \lambda b_i$ $(i = 1, 2)$ with λ being a function of intracellular ATP level. For simplicity and generality, we ignore the possible energy dependence of degradation as being of lower magnitude than these constructive processes. We will usually consider λ values between 0 and 1, reflecting the modulation of transcription rates from a theoretical maximum value (when ATP supply is not rate-limiting) to zero (in the absence of any available ATP supply), but consider relaxation of this scale later.

The rates of biophysical processes depend on corresponding free energy availability, a complicated quantity that for gene expression strongly depends on the cellular ATP/ADP ratio, as well as other metabolic factors. Typical ATP/ADP ratio values have been shown to vary between 2-3 in mouse pancreatic β-cells [200], and 19 ± 4 in wild-type *E. coli* cells [116]; variability has also been observed within and between different plant cell components [260]. To model this complex quantity and make it transferable across different biological species, we use a parameter A^* to reflect a ratio of a cell's free energy availability to a maximum possible value. This sets the upper and lower limits of A^* to be 0 and 1 for each cell. By assuming a constant sum of ATP and ADP concentrations within an *E. coli* cell, as observed by Jensen and Michelson [116] for wild-type and *atp* mutant cells, an increase in ATP concentration corresponds to an increase in free energy availability by changing the ATP/ADP ratio, and allows values of A^* to be mapped to, for example, specific ATP concentrations for a given biological system of interest (Fig. 2.2).

To model the dependence of transcription rate λ on energy availability A^*, we follow das Neves *et al.* [62], who found a sigmoidal relationship between the total transcription rate in a cell and ATP concentration. As we are primarily concerned with the ATP

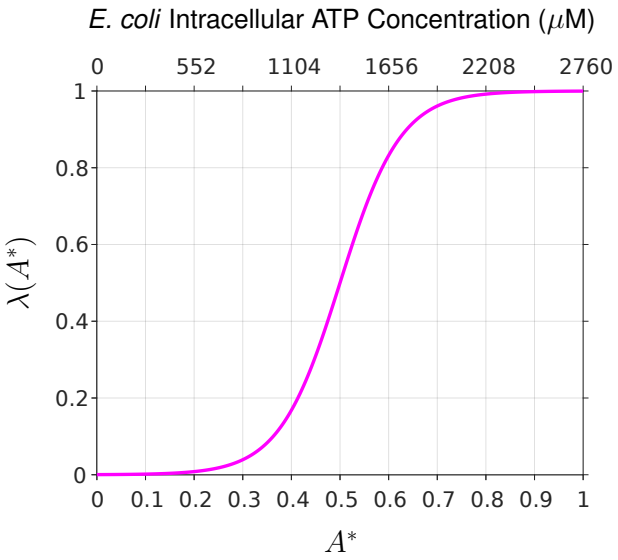


Figure 2.2: **Model relationship between energy availability and gene expression rates.** Sigmoidal curve given by Equation (2.2.4), displaying limiting λ for small A^* and maximal λ as A^* approaches 1. Secondary x-axis (top x-axis) provides a biological example of the intracellular ATP concentration of *E. coli* with respect to the maximum ATP concentration using known upper and lower bounds [309]. In this example, the lower bound of *E. coli* intracellular ATP concentration is 320 μM, corresponding to small λ.

dependence of gene expression, we modelled λ with a sigmoidal curve, yielding monotonically increasing rates as energy increases (adoption of a linear rather than a sigmoidal relationship between scaled energy and dynamic rates did not change our qualitative observations, as displayed in Section 2.3.2). Specifically, we use Equation (2.2.4), where $s_1 = 16, s_2 = -8$ are chosen to produce the monotonically increasing curve for $\lambda(A^*)$ in Fig. 2.2.

$$\lambda(A^*) = \frac{1}{1 + e^{-(s_1 A^* + s_2)}}. \tag{2.2.4}$$

For biological reference, Fig. 2.2 displays the known range of intracellular ATP concentration of *E. coli* (1.54 ± 1.22 mM [309]). At the lower bound of ATP concentration, 320 μM, λ is small, equating to a low maximum expression level. For the upper bound, 2760 μM, λ is at its maximum value and represents maximal expression level. Concentrations below 320 μM are outside the detected *E. coli* ATP range [309], so we take this to be the critical intracellular energy to continue as a living cell and therefore may represent

non-living or dying cells. Overall we thus obtain:

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = \lambda(A^*)a_1\frac{x_1^n}{\theta_{a_1}^n + x_1^n} + \lambda(A^*)b_1\frac{\theta_{b_1}^n}{\theta_{b_1}^n + x_2^n} - k_1x_1, \tag{2.2.5}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = \lambda(A^*)a_2\frac{x_2^n}{\theta_{a_2}^n + x_2^n} + \lambda(A^*)b_2\frac{\theta_{b_2}^n}{\theta_{b_2}^n + x_1^n} - k_2x_2. \tag{2.2.6}$$

Time-dependent and steady state numerical solutions of Equations (2.2.5)-(2.2.6), with Equation (2.2.4), were generated in Python 3.6.5 using the *integrate.odeint* and *optimize.root* functions within the Scipy module; *integrate.odeint* is a stiff or non-stiff solver for a system of first-order ordinary differential equations which implements a Backward Differentiation Formula (stiff) or predictor-corrector (non-stiff) method, and *optimize.root* solves the system of nonlinear equations using a modification of the Levenberg-Marquardt algorithm. Prescribed initial conditions covered a grid of x_1 and x_2 at equally spaced points when calculating numerical solutions; bifurcation diagrams and heatmaps were constructed using the numerical solutions. In the case where multiple degenerate solutions are found, attractor quantity was confirmed using both Maple (2019) and Matlab (R2018a; Mathworks). All scripts used for this study are openly accessible through `https://github.com/kerr-mathbio/energy-variability-decision-making.git`.

For each parameter set, the stable steady states of the system correspond to cell fates. We do not consider unstable or metastable steady states to represent cell fates as small perturbations in expression due to noise will move the system towards a stable point; we consider both unstable and metastable steady states as 'functionally unstable' in our study. If multiple stable states exist, the level of proteins x_1 and x_2 at each steady state determine the different phenotypes a cell can generate, as shown by Huang *et al.* [109]; in their study, high x_1 and low x_2 corresponded to a progenitor cell differentiating into an erythroid cell, rather than a myeloid cell.

As we are modelling the observed interactions in Fig. 2.1(b) rather than biochemical interactions, the analysis of our results considers the qualitative behaviour rather than

observing quantitative results.

From this point, we set $a = a_1 = a_2$, $b = b_1 = b_2$, $k = k_1 = k_2$, $\theta_a = \theta_{a_1} = \theta_{a_2}$, $\theta_b = \theta_{b_1} = \theta_{b_2}$ and each section will explicitly state the fixed and varied parameters for the work that follows.

2.3 Results

Our model system generally exhibits dynamic behaviour that, starting from some initial condition $(x_{1,0}, x_{2,0})$, converges to a particular steady state characterised by values (x_1, x_2). A limited number of these steady states exist for a given parameterisation of the system; these are *attractors*, so called because initial conditions are 'attracted' towards these stable states. A range of initial conditions will converge to the same attractor; this range of $(x_{1,0}, x_{2,0})$ values is the *attractor basin* of that attractor.

We view distinct attractors, with distinct patterns of protein levels, as distinct cell fates. The attractor basin of each corresponds to the range of cell states that will return to that cell fate. This range makes each fate robust to fluctuations: for example, a small change in the level of one protein in one attractor state will likely still fall within that attractor's basin, so the system will return there.

In the following sections, we consider the number and properties of distinct cell fates (attractors) supported by the system at different energy levels. As we increase energy levels we often observe *bifurcation*, where one stable state transitions to two new states (increasing the number of options for cellular decision making). We also consider the "difficulty" of switching between attractors (moving from one cell fate to another); in biology this can be accomplished through noisy or controlled external influences [215, 217, 274].

2.3.1 Intracellular energy availability modulates cellular decision-making landscapes

We first sought to understand how increasing energy availability changes the number of decision-making options available to the cell. To this end, we explored the bifurcation behaviour of the steady-state protein levels in our model as we changed energy availability. As our parameterisation for now imposes a symmetric structure on the phase space of steady-state protein levels, we begin by simply using the diversity of state level x_2 levels to illustrate the range of attractors present in the system for given energy levels. An example bifurcation diagram for default values and varying A^* is displayed in Fig. 2.3, with inset figures displaying time-dependent trajectories towards stable attractors for fixed A^* values. Throughout this section, θ_a, θ_b and n are fixed at their default values, and parameters a, b and k are varied.

We explored the emergence of different attractors (hence, the emergence of different cell fates that can be decided between) with energy level A^* under a variety of different parameterisations. A clear general trend emerged, whereby the number of distinct attractors supported by the system increases with increasing energy availability (Figs. 2.4-2.6). There is thus an increased diversity of stable protein states, and hence an increased number of options for cellular decisions, if energy levels are higher. The separation of these attractors also increased with increasing energy, with the attractor basin associated with the intermediate attractor becoming wider, in agreement with [109, 119]. This suggests a biological stabilisation of states characterised by intermediate protein levels (see Fig. 2.4(b) for example).

The quantitative properties of these attractor states vary with model parameters. Increased promoter activity a and b has the effect of increasing the separation of protein levels in distinct states, likely due to a simple elevation of maximum level. Correspondingly, increased degradation activity k has the opposite effect, diminishing the differences between distinct states (Fig. 2.6). In particular, high values of k prevent the emergence of tristability, limiting the system to bistability even at high energy levels (Fig. 2.6(c)-(d)).

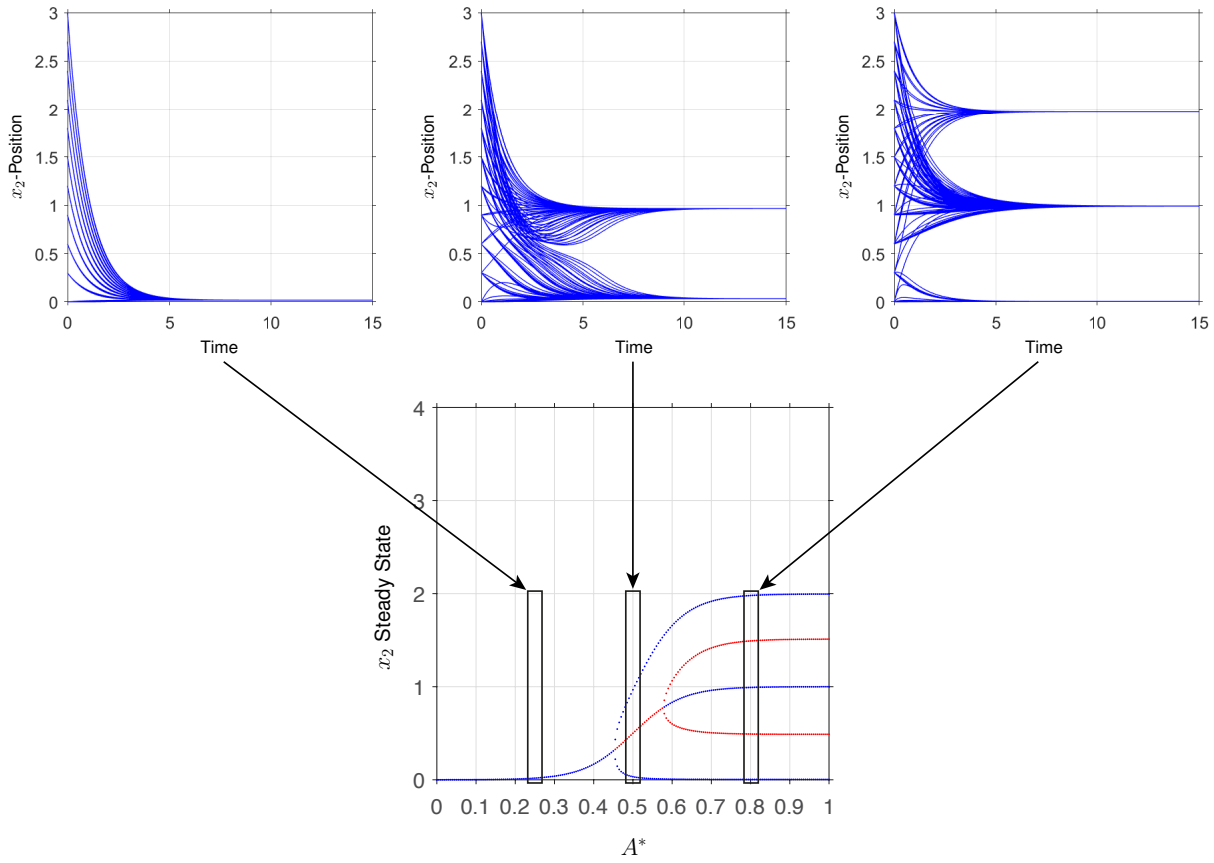


Figure 2.3: **Increasing energy levels support more distinct attractors.** Bifurcation diagram for the default parameter set, displaying stable steady states (blue circles) and unstable steady states (red circles) for $A^* \in [0, 1]$ in steps of size 5×10^{-3}. Insets show behaviour of time-dependent simulations with multiple initial conditions at specified A^* values. Time-dependent examples are displayed for $A^* = 0.25, 0.5$ and 0.8 to show examples of the behaviour included in the bifurcation diagram.

The bifurcation dynamics display some subtle variation in different cases. For example, at high levels of a (the maximum expression level of the conditional promoter), increasing energy availability drives the system through re-entrant behaviour, where the number of attractors (decision options) runs from 1 to 3 then down to 2 before returning to 3 (see Figs. 2.7 and 2.8 for detailed views of this behaviour). We shall see below (for example in Figs. 2.9 and 2.10) that similar behaviour is also observed in other regions of parameter space.

Generally, a single stable branch exists below an A^* threshold (Figs. 2.4-2.6). Biologically this implies that a single cell fate exists below an intracellular energy level, restricting the decision-making ability of a cell under this architecture. Once intracellu-

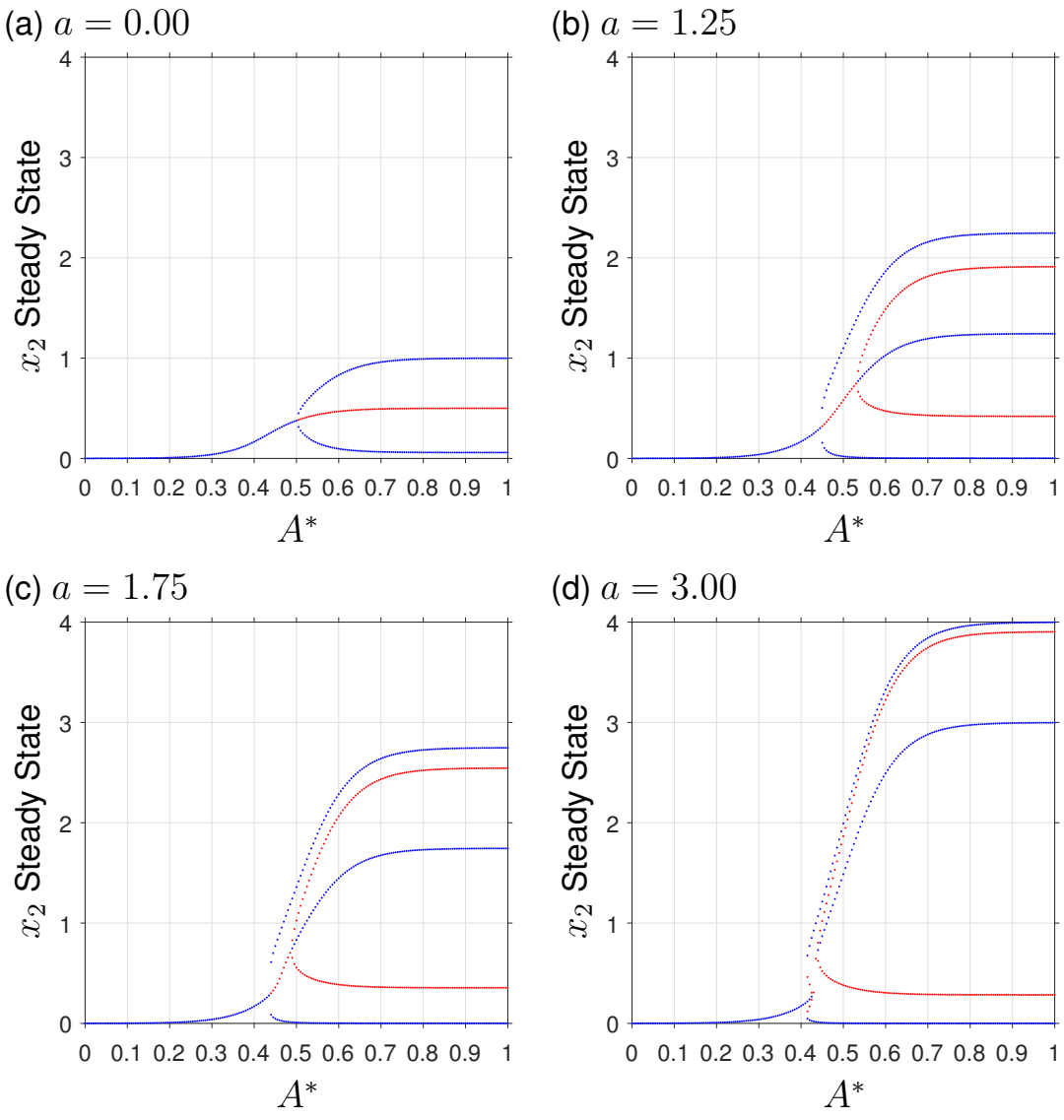


Figure 2.4: **Attractor basins with increasing energy availability A^* and maximum expression level from conditional promoter.** Panels (a)-(d) display bifurcation diagrams for increasing values of the maximum conditional promoter expression level a. Plots display x_2 values for stable steady states (blue circles) and unstable steady states (red circles) over $A^* \in [0, 1]$ in steps of size 5×10^{-3}. Throughout (a)-(d), all parameters, except a and A^*, are fixed at their default values.

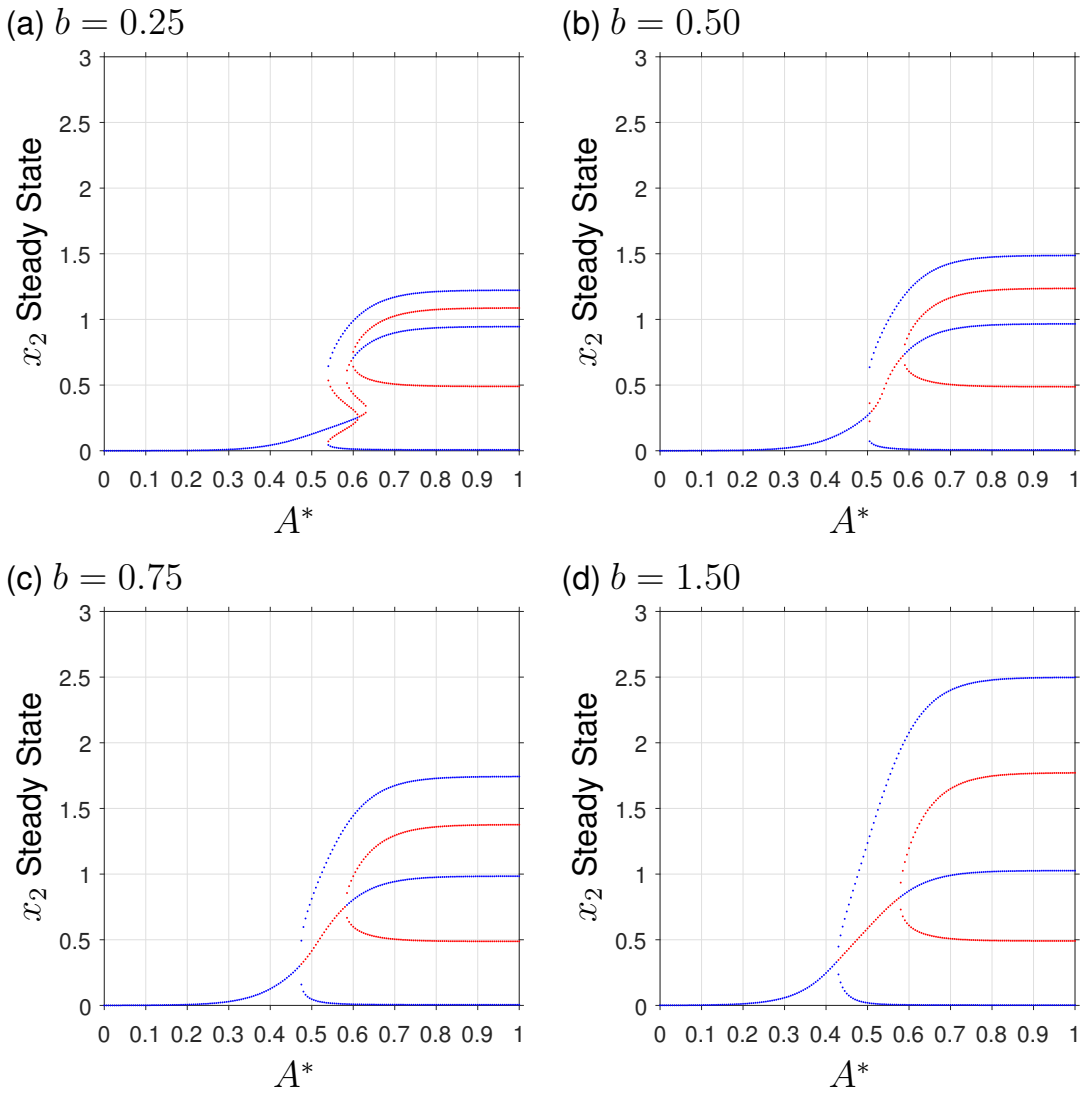


Figure 2.5: **Attractor basins with increasing energy availability A^* and constitutive expression level.** Panels (a)-(d) display bifurcation diagrams for increasing constitutive promoter expression b. Plots display x_2 values for stable steady states (blue circles) and unstable steady states (red circles) over $A^* \in [0, 1]$ in steps of size 5×10^{-3}. Throughout (a)-(d), all parameters, except b and A^*, are fixed at their default values.

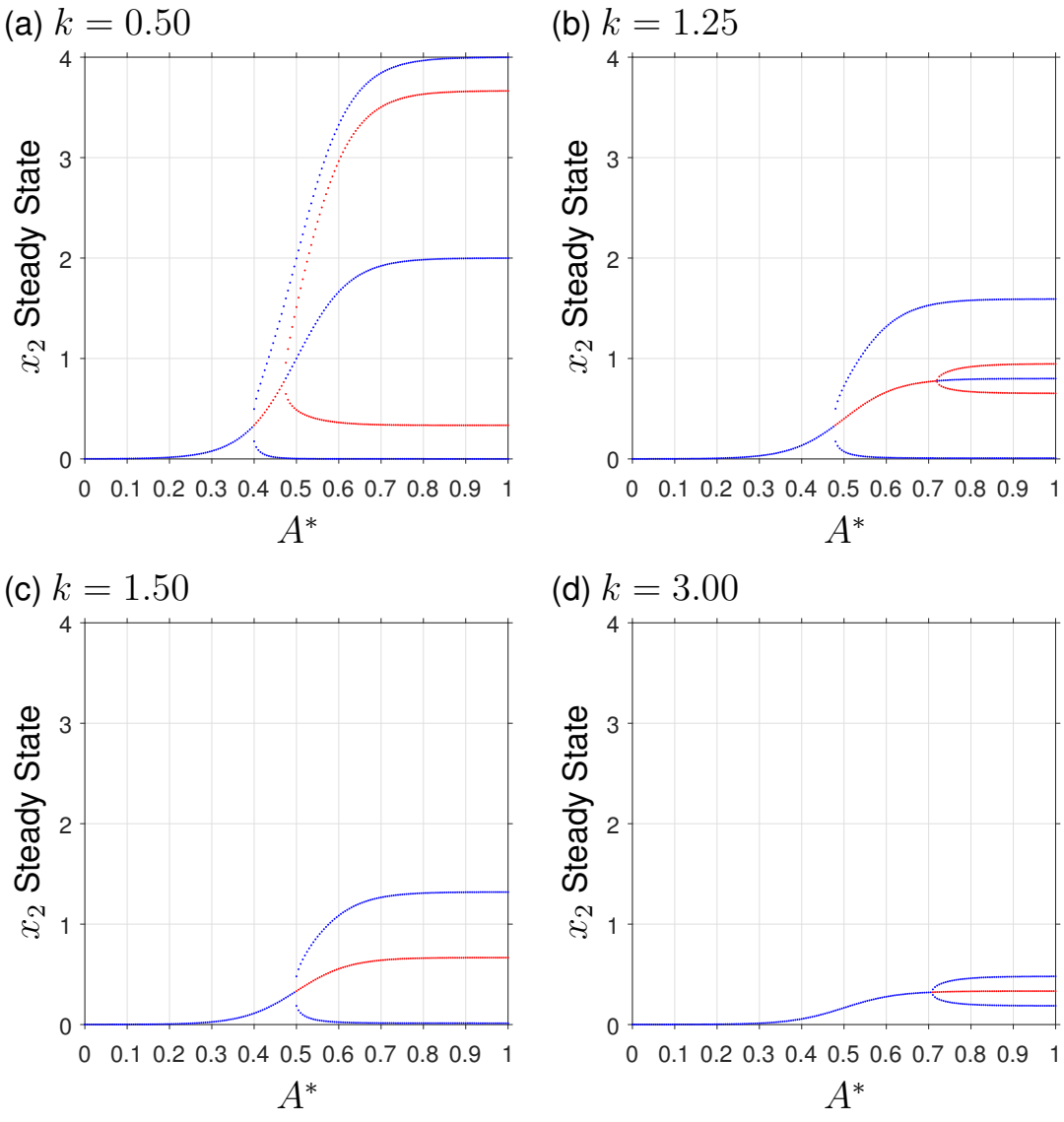


Figure 2.6: **Attractor basins with increasing energy availability A^* and degradation level.** Panels (a)-(d) display bifurcation diagrams for increasing values of degradation level k. Plots display x_2 values for stable steady states (blue circles) and unstable steady states (red circles) over $A^* \in [0, 1]$ in steps of size 5×10^{-3}. Throughout (a)-(d), all parameters, except k and A^*, are fixed at their default values.

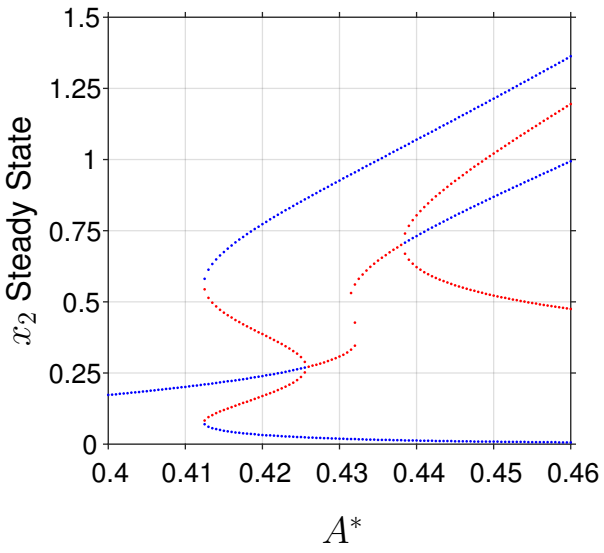


Figure 2.7: **Re-entrant behaviour in decision making options.** Close-up view of Fig. 2.4(d) over $A^* \in [0.4, 0.46]$ in steps of size 5×10^{-4}. The figure displays stable steady states (blue circles) and unstable steady states (red circles). The number of stable steady states transitions from 1 to 3, then 3 to 2 and finally 2 to 3 in a small region of A^*.

lar energy availability exceeds a threshold, multiple steady states are present, enabling cellular decision-making. All figures demonstrate that, for our decision-making architecture, an increase in A^* generally has the effect of accumulating stable branches in the bifurcation diagrams. This means a broader decision-making landscape for a cell when it has high intracellular energy availability. If degradation, k, is too high the central stable attractor does not exist and the system is limited to bistability. Therefore, the decision-making capabilities of a cell would be reduced if the degradation level is sufficiently high. Re-entrant stable steady state behaviour exists for certain parameter sets, suggesting a potential optimum intracellular energy range for phenotypic diversity. This may give a cell an optimal range of intracellular energy levels to support an increased variety in cell phenotypes, enabling superior adaption to stochastic extracellular environmental changes. This behaviour could also have negative consequences for a cell due to phenotypes becoming unsupported with slight variances in energy, forcing a cell to establish a different phenotype which, due to being a time dependent process, may not be performed quick enough to tolerate an environmental change.

In Fig. 2.9 we show this 'phase portrait' of attractor landscapes for varying energy

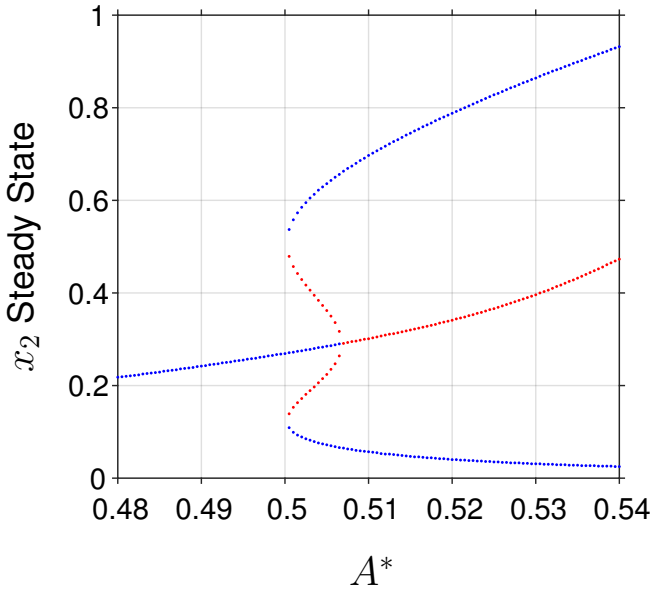


Figure 2.8: **Destabilisation of central attractor in re-entrant behaviour.** Close-up view of Fig. 2.5(b) for $A^* \in [0.48, 0.54]$ in steps of size 5×10^{-4}. The figure displays stable steady states (blue circles) and unstable steady states (red circles). The number of stable steady states transitions from 1 to 3, then 3 to 2 in the small region of A^*.

A^* and constitutive expression b, along with example structures of the corresponding attractor basins. The general observation of increasing number of fate options with energy holds throughout. There is a separatrix line in this phase portrait between regions supporting a single attractor basin and those supporting more than one. For a given dynamic parameter, this separatrix is crossed as energy availability increases, then further increase has the effect of expanding the attractor basin associated with the state where both genes are expressed symmetrically (as in Johnston *et al.* [119]).

What is the physical intuition behind these structural changes with energy and dynamic parameters? Start with the low activity, low energy case. Here there is simply not enough cellular 'production capacity' – the ability to express genes – to allow dynamic signalling interactions between gene products. The system is limited to low levels of expression for both products, favouring a single, low-expression state.

As we now increase b, the basal expression rate increases, and the high coefficient associated with the cross-repression term induces very high antagonism – any balance between the gene products is highly unstable. The system invests all available energy

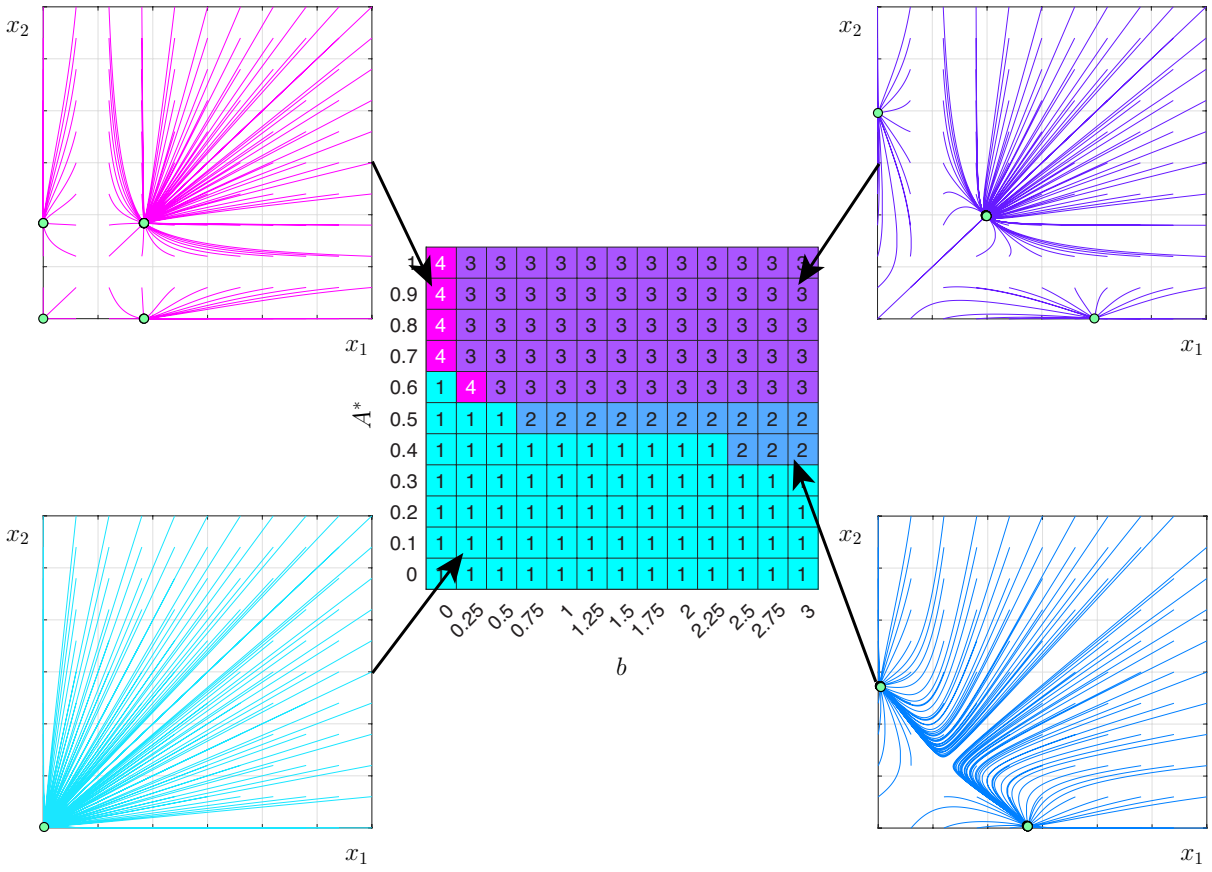


Figure 2.9: **Decision-making capacity depends on cellular energy and dynamic parameters.** Heatmap for $a = 1$, $b \in [0, 3]$ and $A^* \in [0, 1]$, with all remaining parameters fixed at their default values. Inset figures present examples of the attractor landscape with solution trajectories (coloured) and stable steady states (green circles with black circumference). The 4 categories are, when viewed in portrait orientation, 1 stable steady state (bottom left, turquoise), 2 stable steady states (bottom right, blue), 3 stable steady states (top right, purple) and 4 stable steady states (top left, pink).

getting to either one state or another. The low energy levels and correspondingly low expression levels mean that 'small number' effects are important: any small perturbations from symmetric expression will be amplified to favour one gene over the other.

In the alternative case, where we have low basal expression but high energy, we have substantial 'production capacity' that can now support an energetically more demanding intermediate state. Higher expression levels mean that fluctuations can be buffered and intermediate states are more stable. Here, there is constant tension between genes, requiring high energy levels; the fixed states are also still supported. In order to favour either gene product we need enough initial protein to jump-start the system (otherwise we reach the trivial zero or near-zero attractor).

This latter case remains true in the final case of high energy and high basal expression, except here the zero or near-zero attractor is unstable because from here, the substantial 'production capacity' means that some protein can always be produced, which is likely to drive the system towards one of the non-trivial attractors.

We next explored this behaviour varying both promoter levels a and b together (Fig. 2.10). At low values of the conditional promoter activity ($a \leq 0.5$) the system is restricted to mono- or bistability for all values of b and a single stable attractor exists below an A^* threshold. High values of a support the greater diversity of stable states previously seen in Fig. 2.9, however the attractor landscape changes minimally for higher a (Fig. 2.10(e)-(f)). For high a and some values of b, we again see the re-entrant behaviour from the bifurcation diagrams (Figs. 2.7 and 2.8), with the 4-attractor state supported only at some intermediate values of energy availability. Illustrations of the nature of these re-entrant steady states are given in Figs. 2.11 and 2.12, for increasing values of A^* and b respectively.

This re-entrant behaviour for increasing values of A^* or b (Figs. 2.7-2.12) merits further analysis. The reduction from 4 to 3 stable states with increasing energy is likely due to the destabilisation of the near-zero attractor (Fig. 2.11), where expression levels of both genes is low. Above an energetic threshold, a baseline of production will be constantly

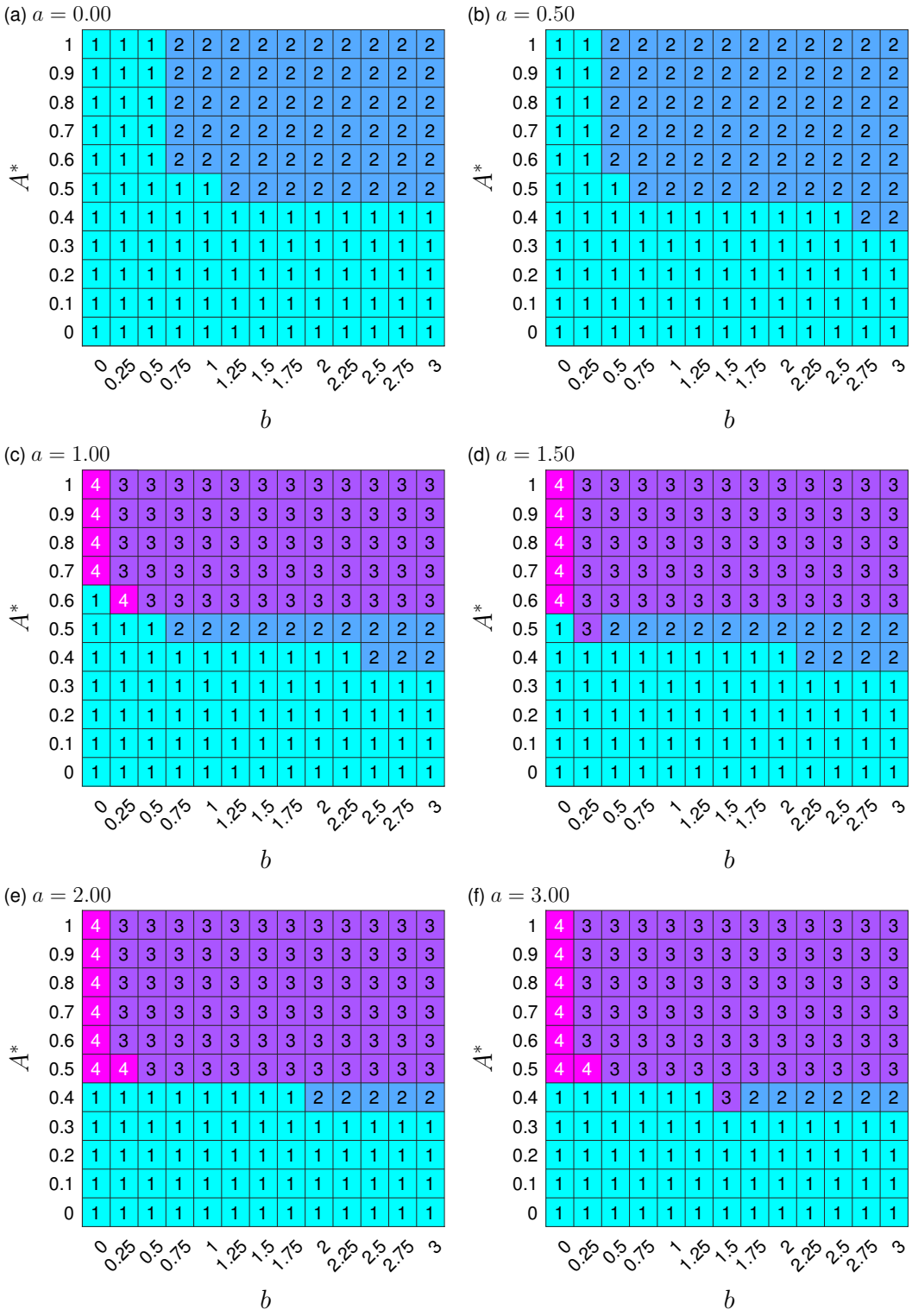


Figure 2.10: **Decision-making landscapes depend on energy availability and dynamic rate parameters.** Panels (a)-(f) display heatmaps for 6 increasing values of a, when $n = 4$. Each panel exhibits the number of stable steady states for combinations of $b \in [0, 3]$ and $A^* \in [0, 1]$. Throughout (a)-(f), all parameters, except a, b and A^*, are fixed at their default values.

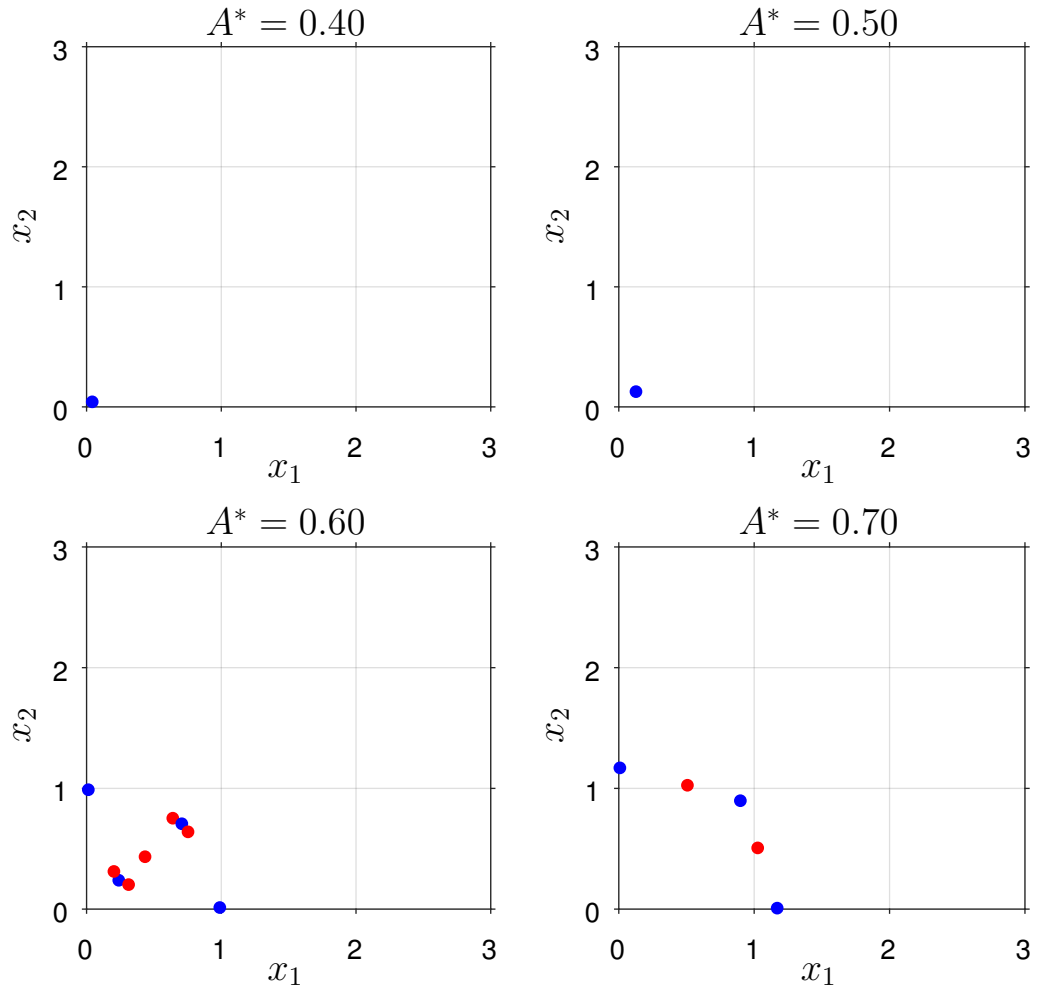


Figure 2.11: **Re-entrant behaviour showing destabilisation of the 'near-zero' attractor.** Array displays stable (blue) and unstable (red) states for $a = 1$, $b = 0.25$ and $A^* \in [0.4, 0.7]$ in Fig. 2.10(c). Increasing energy develops 4 stable attractors which decreases to 3 as A^* increases. All remaining parameters are at their default value.

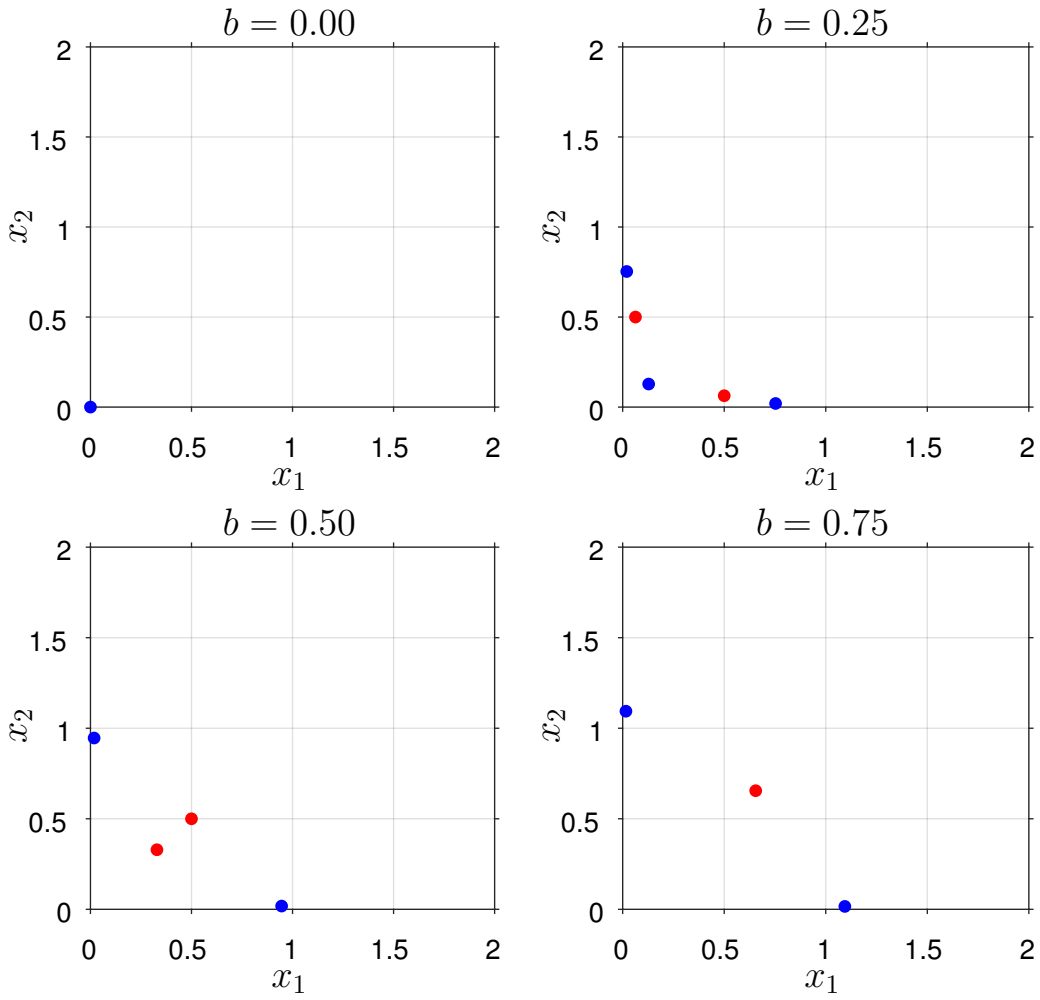


Figure 2.12: **Re-entrant behaviour showing destabilisation of the central attractor.** Array displays stable (blue) and unstable (red) states for $a = 1.5$, $A^* = 0.5$ and $b \in [0, 0.75]$ in Fig. 2.10(d). Increasing energy develops 3 stable attractors from the initial single stable state which decreases to 2 as b increases. All remaining parameters are at their default value.

occurring, and if the activity of the conditional promoter (a) is high enough, any small increases in protein level can rapidly self-amplify and drive the system to an attractor with higher expression of protein products. The 4-attractor system will then only be supported at low b and sufficiently low energies, as we observe. When the number of attractors (decision options) runs from 1 to 3 then down to 2 before returning to 3, the central stable attractor is destabilised as energy increases (Fig. 2.12). This is likely due to a higher contribution from the constitutively expressed promoter as a proportion of the overall expression when energy increases, enhancing the antagonism, increasing the overall expression of both proteins and therefore driving the system to one state or the other as the central state cannot be sustained.

In summary, when a is small (Fig. 2.10(a)-(b)), the system is limited to mono- or bistability, which may equate to a limited diversity in phenotypic possibilities for a cell. Increasing a lowers the separatrix and decreases the A^* threshold for tristability. Biologically, this suggests that for our decision-making architecture, as the maximum expression level of the promoter under activator control increases, less energy is required for an expansion in a cell's decision-making landscape. Independent of the value of a and b, a singular stable steady state always exists at low energy, in theory limiting a cell to generating a singular phenotype.

2.3.2 Modelling the energy parameter through a linear function

The parameter λ was initially modelled by the sigmoidal curve in Equation (2.2.4). Modelling the ATP-dependent modulatory term through a linear relationship, $\lambda = A^*$ (Fig. 2.13), whilst fixing θ_a, θ_b, n and k at their default value, shows qualitatively similar behaviour compared to when $\lambda(A^*)$ is sigmoidal (Fig. 2.10). In contrast we observe that significantly lower A^* thresholds are required for multiple stable steady states, increased parameter sets with 4 stable attractors and there exist more parameter sets with re-entrant behaviour.

The choice of function to model λ therefore has an effect on the qualitative behaviour,

but the number of stable attractors still increases as A^* increases. Biologically this again displays an increased decision-making landscape for a cell as intracellular energy availability increases, as observed for sigmoidal λ.

2.3.3 Cooperativity effects on system behaviour

We next asked how these relationships between attractor basin structure (number of decision options) and energy depends on the cooperative nature of our model interactions (represented by the Hill coefficient n in our governing equations). We found that decreasing the Hill coefficient to $n = 3$ (Fig. 2.14) led to some minor rearrangements of the phase portrait, decreasing to $n = 2$ (Fig. 2.15) does not display any re-entrant behaviour, while the quasi-linear case of $n = 1$ (Fig. 2.16) removed any diversity in attractor basin structure for the parameters, and A^* values, we consider.

Therefore, if there is noncooperative binding ($n = 1$) of the proteins, the number of stable states is limited to one (Fig. 2.16). This would relate to no decision-making capability for any cell, independent of intracellular energy level or gene expression levels, in the regimes we are considering. In contrast, if cooperativity is reduced from $n = 4$ (default value, Fig. 2.10) to $n = 3$ or $n = 2$ (Fig. 2.14 or 2.15 respectively), the observed behaviour is qualitatively similar, suggesting that a limited degree of nonlinearity in cellular interactions is sufficient to support decision-making and that energy dependence remains important even in these limited cases.

2.3.4 Regulatory protein binding strength can limit system behaviour

To understand how the binding strength of regulatory proteins affects the decision-making abilities of a cell, the influence of Hill parameters θ_a and θ_b on attractor basin structure was also analysed (Fig. 2.17); θ_a and θ_b can be interpreted as the dissociation constants of the activator and inhibitor regulatory proteins to the promoter regions, respectively.

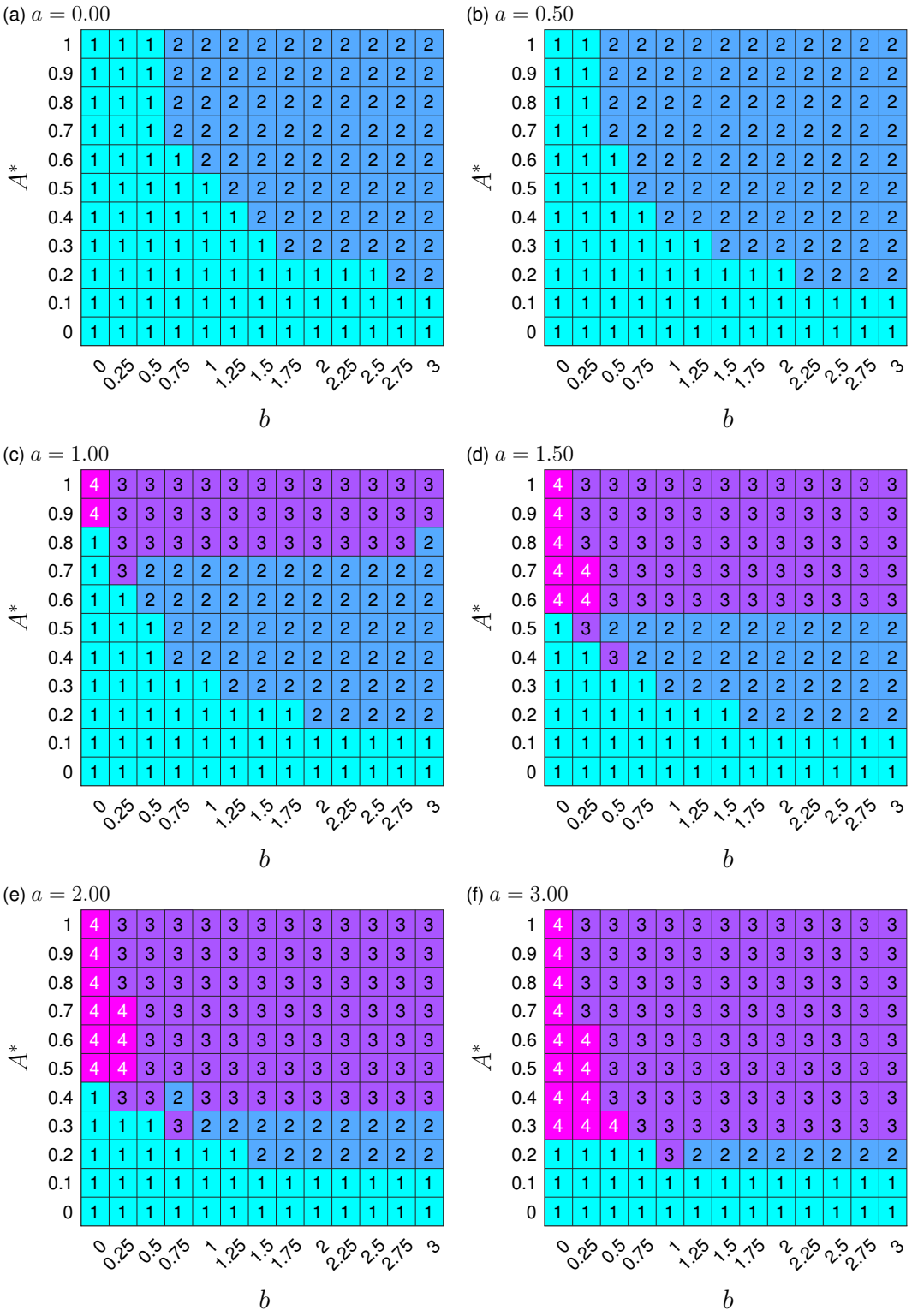


Figure 2.13: **Linear energy function displays similar qualitative decision-making landscapes to the default case.** Panels (a)-(f) display heatmaps for 6 increasing values of a, when $\lambda = A^*$. Each panel exhibits the number of stable steady states for combinations of $b \in [0, 3]$ and $A^* \in [0, 1]$ with all remaining parameters fixed at their default values.

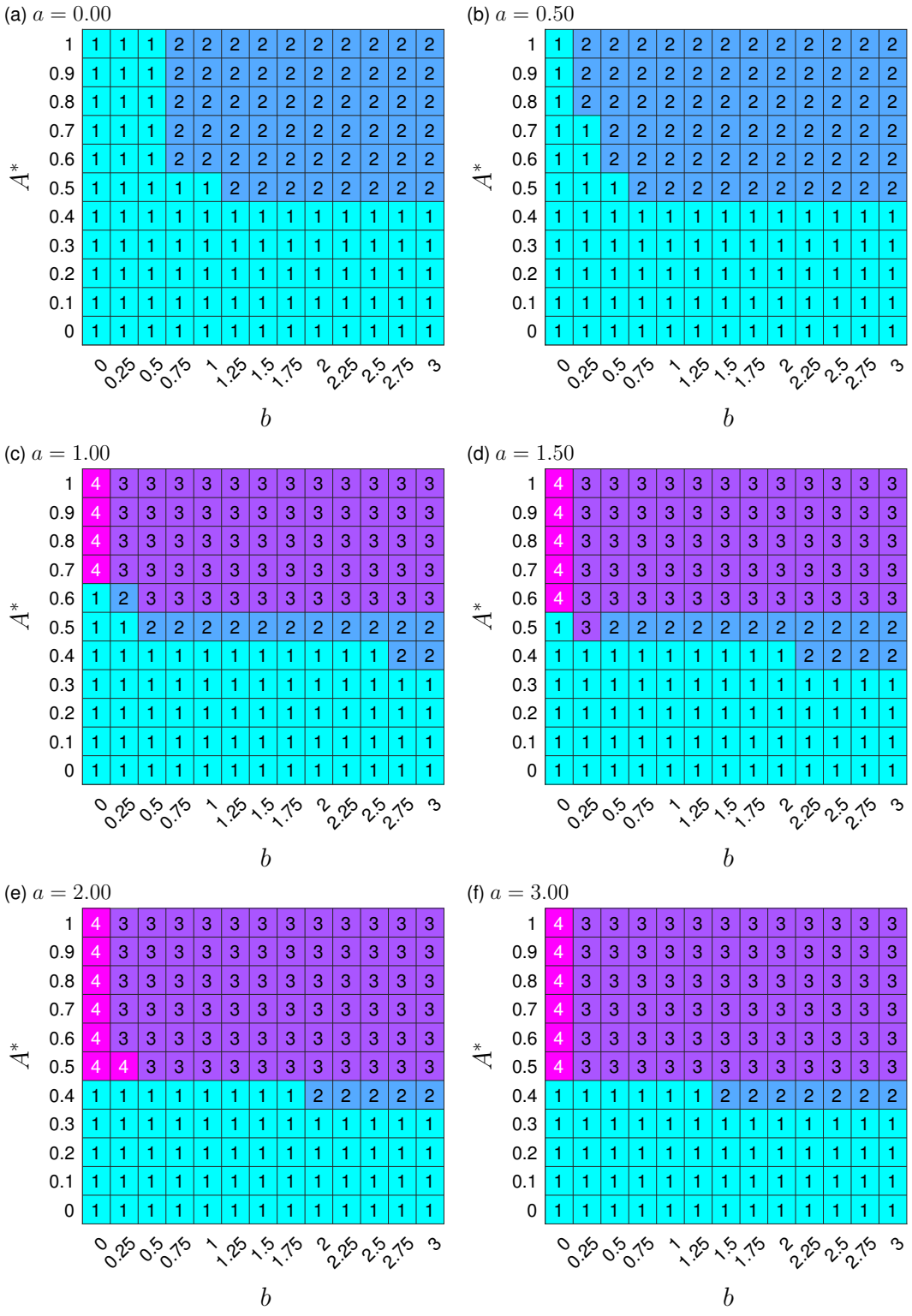


Figure 2.14: **Reduced Hill coefficient value ($n = 3$) displays similar qualitative decision-making landscapes to the default value ($n = 4$).** Panels (a)-(f) display heatmaps for 6 increasing values of a, when $n = 3$. Each panel exhibits the number of stable steady states for combinations of $b \in [0, 3]$ and $A^* \in [0, 1]$, with all remaining parameters fixed at their default values.

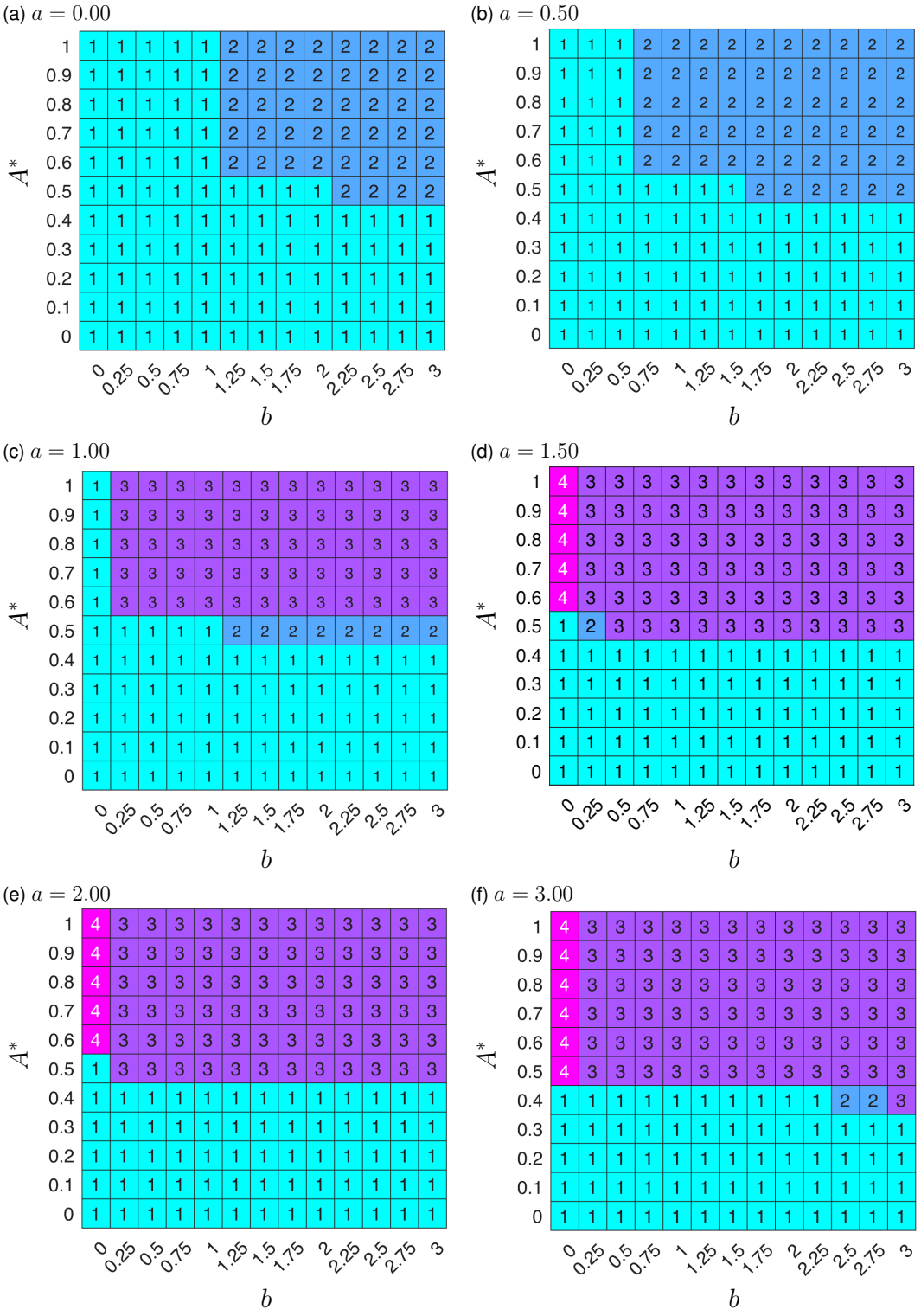


Figure 2.15: **Hill coefficient $n = 2$ displays simplified qualitative decision-making landscapes to the default value ($n = 4$).** Panels (a)-(f) display heatmaps for 6 increasing values of a, when $n = 2$. Each panel exhibits the number of stable steady states for combinations of $b \in [0, 3]$ and $A^* \in [0, 1]$, with all remaining parameters fixed at their default values.

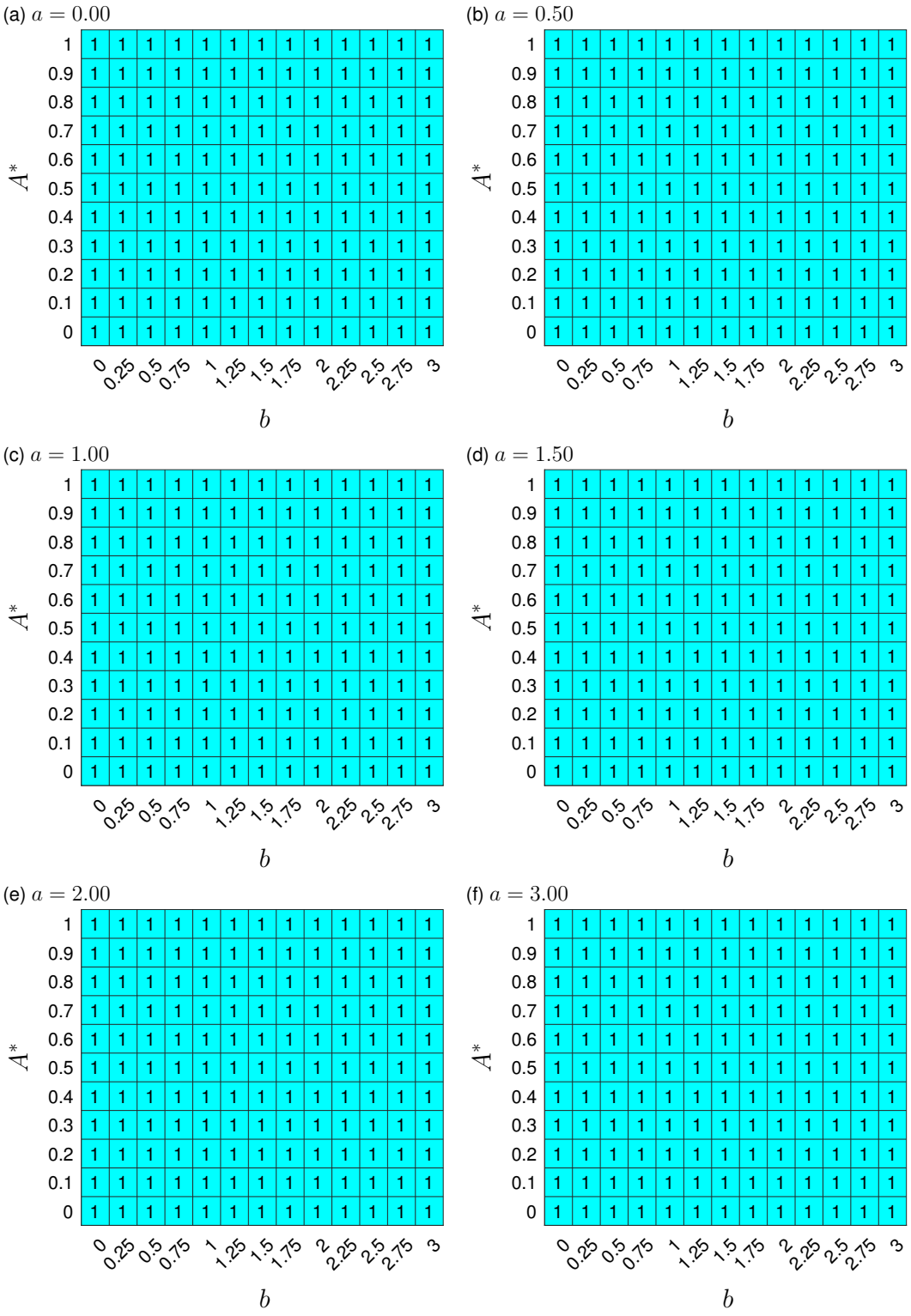


Figure 2.16: **Non-cooperative binding limits the decision-making landscape.** Panels (a)-(f) display heatmaps for 6 increasing values of a, when $n = 1$. Each panel displays the number of stable steady states when spanning $b \in [0, 3]$ and $A^* \in [0, 1]$, with all remaining parameters fixed at their default values.

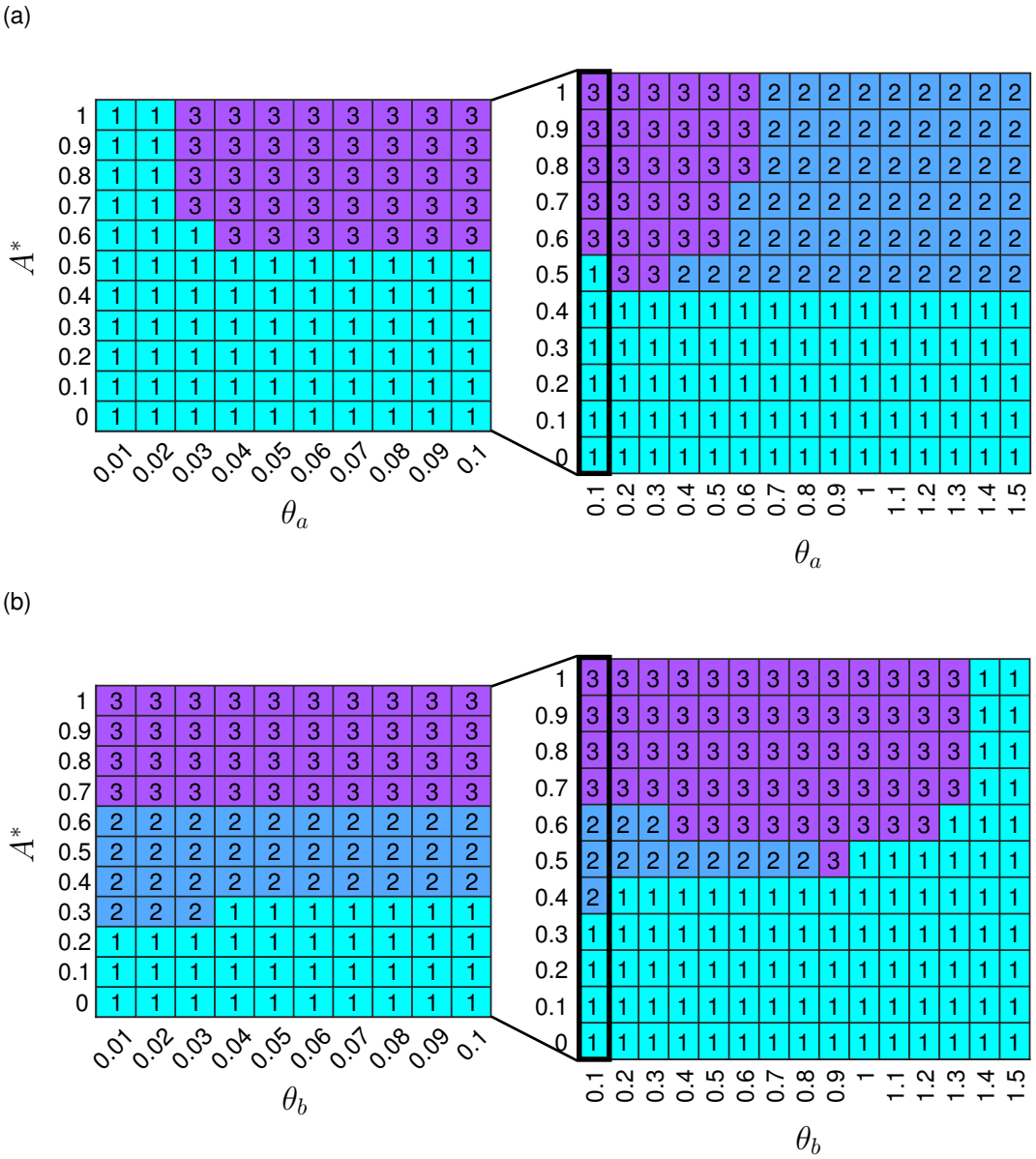


Figure 2.17: **Strength of binding limits the decision-making landscape.** Main panels display the number of stable states for (a) θ_a, (b) $\theta_b \in [0.1, 1.5]$ and $A^* \in [0, 1]$. Inset sub-panels display stable attractor behaviour for θ_a, $\theta_b \in [0.01, 0.1]$. Each panel contains colour coordinated number of stable steady states: 1 (turquoise); 2 (blue); 3 (purple). Remaining parameters are fixed at their default values. Behaviour displayed for $\theta_a = \theta_b = 1.5$ continues for larger values of each parameter.

Increasing θ_a (θ_b) corresponds to decreasing the binding affinity of the corresponding protein to the relevant promoter site, thus requiring more protein to achieve the equivalent levels of activation (inhibition).

Changing θ_a changes the A^* threshold required for tristability; there is a 'sweet spot' in θ_a that supports tristability at higher A^* values. This behaviour is not observed for θ_b; for low θ_b, single, double, and triple stable states are observed at different energy levels, while for high θ_b the system is limited to a single attractor.

Hence, the relationship between energy availability and decision-making capacity of the system depends differently on the interactions encoded by θ_a and θ_b (Fig. 2.17). There is an optimal range of θ_a for tristability, and outside of this range it is more probable for the system to be bistable, but it can also be limited to monostability. This suggests a cell requires a stronger activator binding strength to be less sensitive to changes in the number of activator proteins modulating expression levels. The optimal range for tristability is increased when parameter θ_b is varied. A cell would be less sensitive to changes in the number of inhibitor proteins modulating expression for both weak and strong binding strengths of the inhibitor protein. When the inhibitor binding strength becomes too weak the system is limited to a single stable state, removing the decision-making ability for a cell.

2.3.5 External ATP effects on decision-making landscapes

We next asked how the presence of external ATP sources (for example, from the environment, other cells, or a host cell in the case of a parasite) affected decision-making behaviour in our model. The effect of uptake of external ATP is to raise the baseline level of energy available for transcription [62]. We capture this by adopting a non-zero value of the sigmoid function (λ) at $A^* = 0$, and allowing its maximum to surpass 1 (Fig. 2.18).

This models an externally available ATP source for a cell. We found that external ATP supply shifted the profile of supported phenotypic states towards those with more stable phenotypes (Fig. 2.19), suggesting that external ATP enables the cell to enter

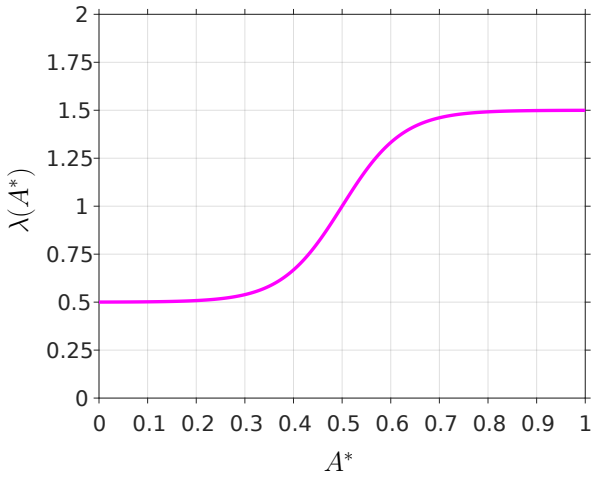


Figure 2.18: **Model relationship between energy availability and gene expression rates with external ATP influences.** Modified λ produced by vertically displacing the original sigmoid in Fig. 2.2.

a higher-decision state-space more easily (for example, compare Fig. 2.10(c) and Fig. 2.19(c)).

2.3.6 Stable steady-state transitions

The transition between phenotypes can be analysed by observing the 'difficulty' of moving between attractors at each energy level. Here, we consider the minimum work to transition from one extreme attractor state to another, by calculating the distance required to transition directly or indirectly between extreme attractor basins. This distance qualitatively corresponds to the magnitude of an intrinsic fluctuation, or an externally induced change, in gene expression levels required to induce a permanent shift in phenotype. Specifically, we compute the minimum of the following quantities: (i) the shortest distance in (x_1, x_2) space between the high-x_1 attractor and the high-x_2 basin; (ii) the sum of shortest distances between the high-x_1 attractor and an intermediate basin and the shortest distance from that intermediate attractor to the high-x_2 basin. In (ii), an indirect transition within the 3 or 4 attractor landscape can occur through a move into an intermediate attractor basin, relaxing towards the intermediate attractor point, and then a movement into the second extreme attractor basin. This distance is only calculated for landscapes with 2 or

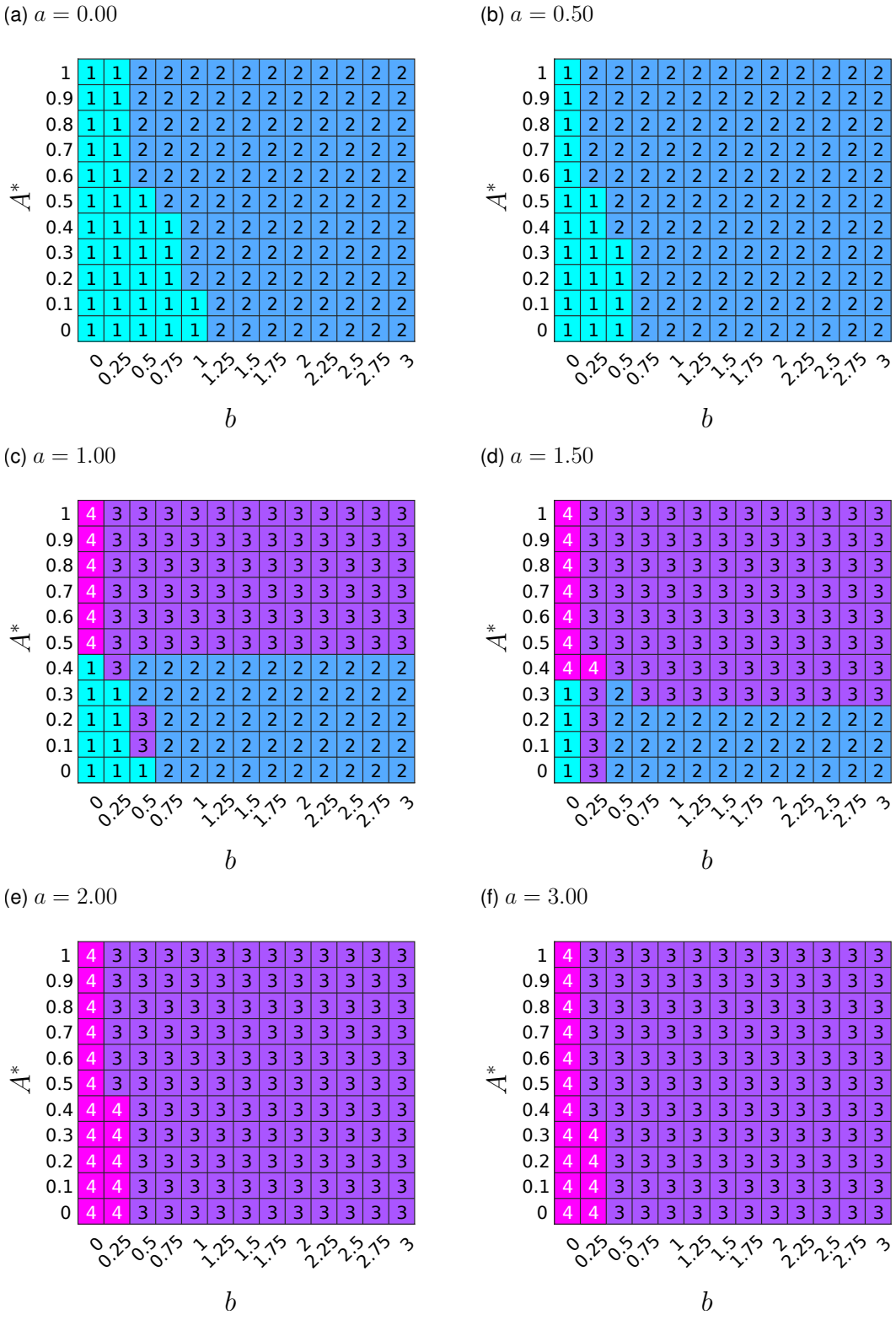


Figure 2.19: **Decision-making landscape variability is observed when considering external ATP influences.** Panels (a)-(f) display heatmaps for 6 increasing values of a, when $n = 4$, for the sigmoid in Fig, 2.18. Each panel exhibits the number of stable steady states for combinations of $b \in [0, 3]$ and $A^* \in [0, 1]$, with all remaining parameters fixed at their default values.

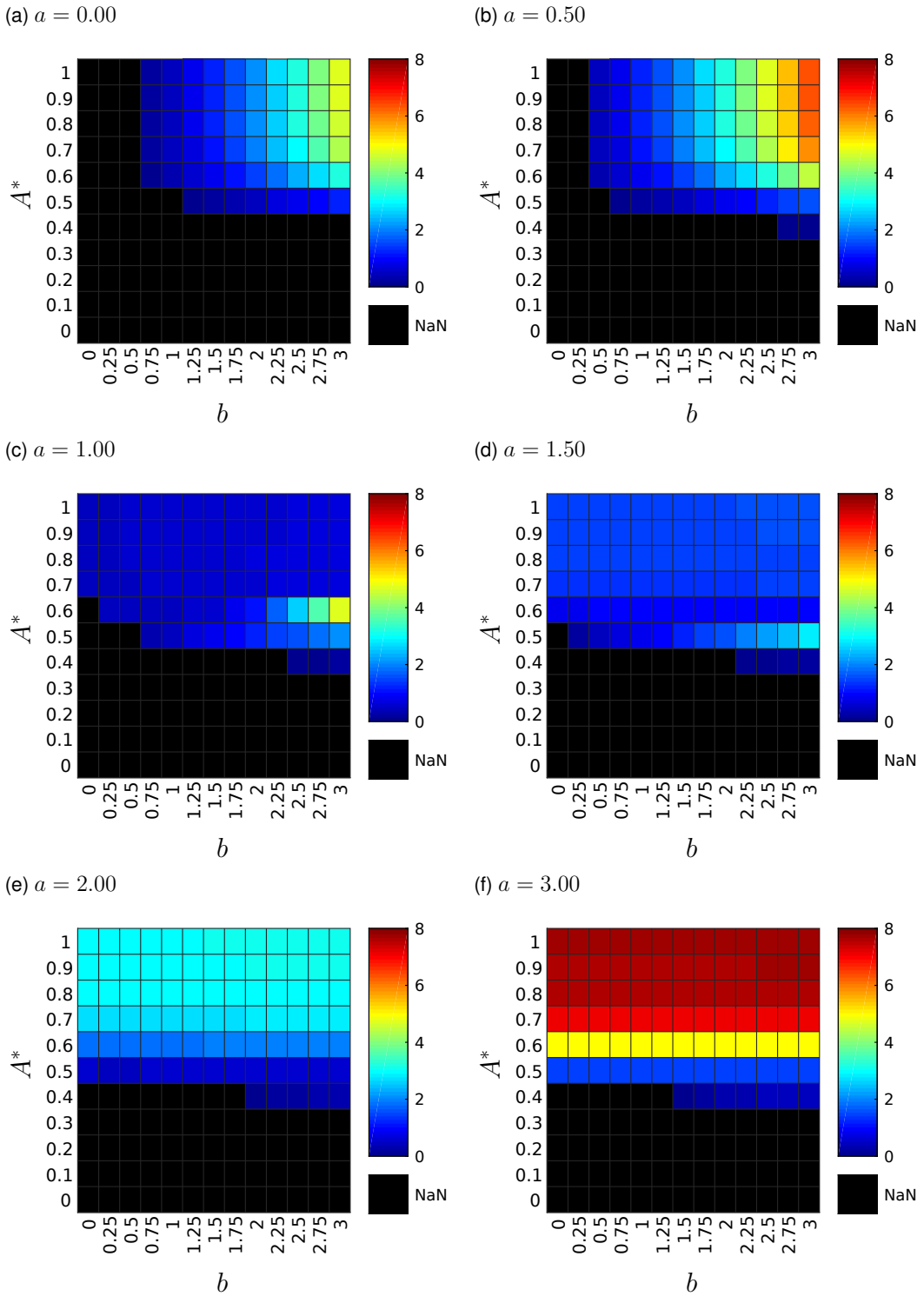


Figure 2.20: **The ability to transition between extreme states becomes harder as energy availability increases.** Panels (a)-(f) display heatmaps for 6 increasing values of a, when $n = 4$. Each panel exhibits the minimum distance between stable steady states, representing the 'difficulty' of extreme attractor transitions, for combinations of $b \in [0, 3]$ and $A^* \in [0, 1]$, with all remaining parameters fixed at their default values. The work required to transition ranges from low (dark blue) to high (dark red). For monostable landscapes the metric is not defined and is represented by black in each panel.

more attractors, as a monostable system cannot transition.

The easiest transitions (shortest distances) occur as soon as the landscapes transitions from monostable to multi-stable (Figs. 2.10 and 2.20), where extreme attractors are closest with little intermediate volume of phase space. The most difficult transitions (hence, most stable phenotypic states) appear for high values of the conditional promoter activity ($a = 3$) where attractors in the landscape are spread apart (see Fig. 2.20(f)). We consistently observe the most stable phenotypic states for large b, however, for intermediate a values (Fig. 2.20(c)-(d)) the transition difficulty at large b values is reduced compared to low and high a (Fig. 2.20(b) and (f) respectively).

Therefore, increasing energy availability both increases the number of phenotypic choices supported by our model architecture and stabilises the corresponding states (Figs. 2.10 and 2.20). For a bistable system, moderate A^* enables easier transitions, and an increase in energy availability shifts the transitions to be considerably more difficult, stabilising phenotypes. In a tristable landscape, the transition 'difficulty' is initially similar to the bistable case, but an increase in A^* does not alter the transition distance to match that of the bistable landscape, unless a is very high, where transitions become considerably more difficult.

2.4 Experimentally exploring the degree of ATP variability in real *E. coli* cells

Possible experimental tests, which would help validate model predictions, would involve jointly tracking bacterial ATP levels and a measurable phenotypic response in a clonal population prior to, during and post introduction of an antibiotic, using fluorescent reporters. To do this, we adapted an initial protocol to qualitatively monitor single-cell ATP concentrations in a clonal bacterial population. Our methodology followed that of Yaginuma *et al.* [309], who developed a genetically-encoded ratiometric fluorescent ATP indicator "QUEEN", constructed from circularly permuted Green Fluorescent Protein

(cpGFP) fused to an ATP-binding subunit of *Bacillus* PS3 F_0F_1 ATP synthase, suitable for quantifying ATP concentrations inside bacteria cells. This single fluorescent protein biosensor can be excited at two wavelengths, and the ratio between the two emission signals reports a ratiometric ATP level; the ratio does not depend on the amount of the reporter, eliminating errors due to variability in QUEEN-2m levels amongst bacterial cells [309].

We initially transformed *E. coli* JM109(DE3) (wild-type) cells with a plasmid containing the ATP indicator sequence (either the 2m or 7μ variant) and then cultured the cells on selective media (100 μg ml^{-1} ampicillin). The qualitative measurement of ATP concentration within single-cells was determined by calculating the relationship between the emission profile of QUEEN within cells following 458nm (458ex) and 488nm (488ex) excitation. In the original study, samples were excited by 405nm and 480nm, which was unavailable for our protocol. Although the optimal emission would be produced from 405ex and 480ex, our setup still allowed for a quantitative understanding of cell-to-cell variability in ATP within a population.

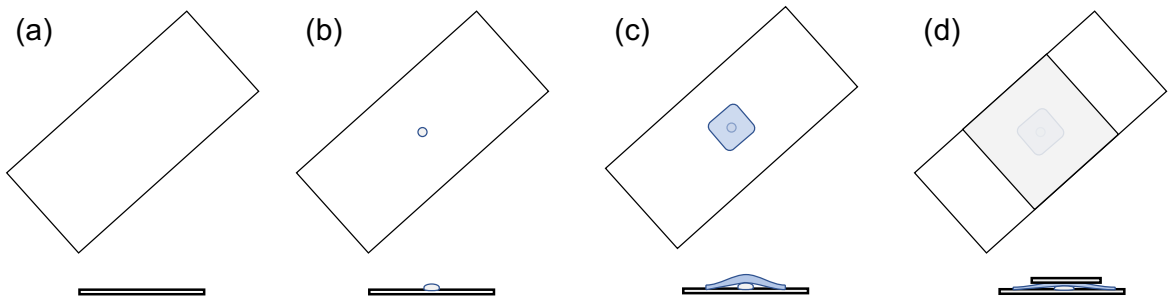


Figure 2.21: **Sample mounting preparation protocol for microscopy.** Schematic illustration of individual steps to prepare the imaging sample with coverslips, agarose pad and bacterial culture. A small volume of *E. coli* sample is placed onto a microscope slide (a)-(b), before the agarose pad (c) and cover slip (d) are placed on top. Samples are then imaged in this setup.

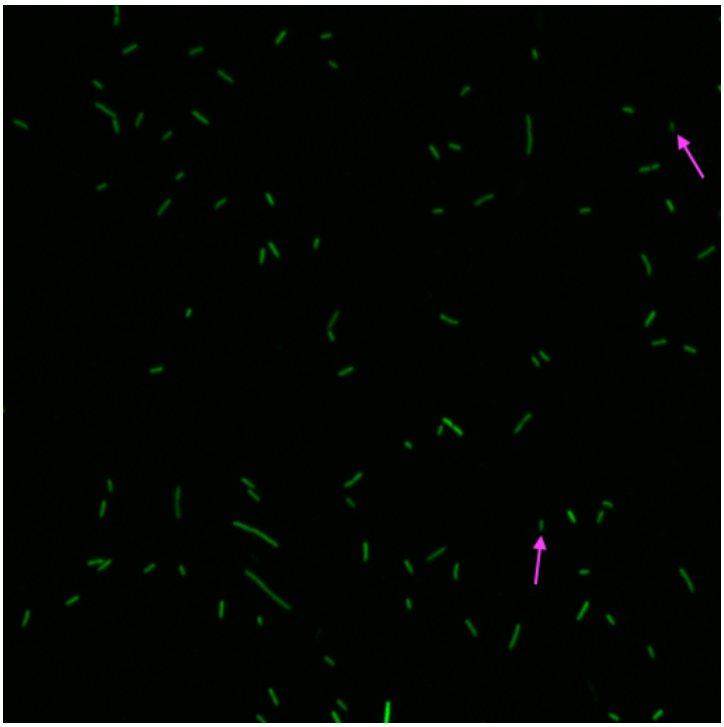
For the single-cell measurement of QUEEN, *E. coli*-2m cells were immobilised on top of a cover slip using a thin agarose pad (see Fig. 2.21) [254, 309]. This technique helps immobilize the cells during the imaging process, aiding in multi-colour fluorescence

microscopy experiments, in addition to lowering induced stress on the bacterial cells. A small amount of cell suspension (5-8 μl), grown to $OD_{600} = 0.2$, was used in each imaging sample. The immobilised cells were imaged with a Zeiss LSM 710 confocal microscope, and the emission profiles of the sample were exported for image analysis.

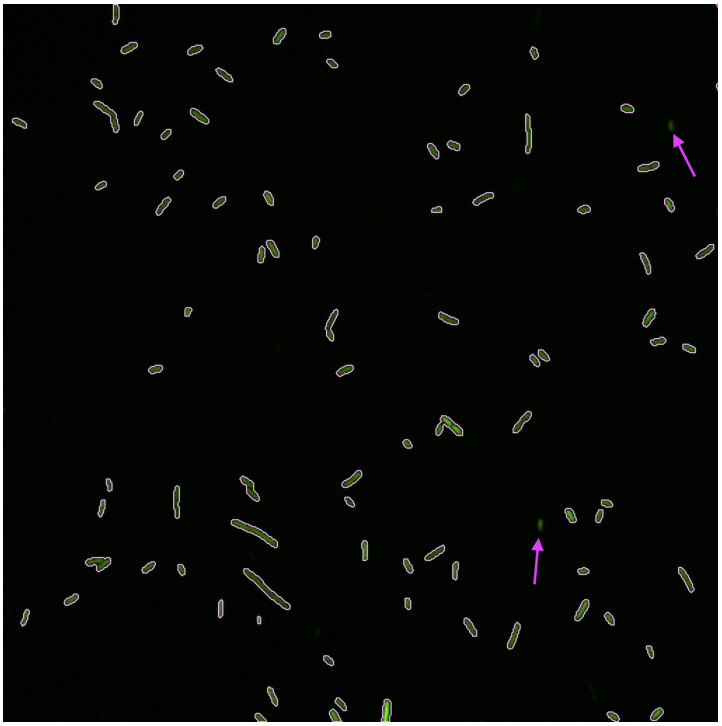
The acquired images (Fig. 2.22(a)) were analyzed with original Python scripts and processed using the following method. First, digital images captured from the microscope were grouped into 458ex and 488ex pairs, and the 488ex image segmented by applying morphological operations on the original image, determining cell boundaries (Fig. 2.22(b)). The boundaries allowed the areas of segmented cells (in number of pixels) to be calculated, with cell areas below a threshold removed for the subsequent calculations. For each cell, mean intensity of all pixels within the cell boundary was calculated (for both 458ex and 488ex images in each pair). Finally, single-cell 488ex/458ex ratio was calculated from the average emission intensity values for each cell.

Image analysis of *E. coli*-2m samples displayed a negatively skewed distribution with mean $\mu = 5.154$ and standard deviation $\sigma = 0.664$ (the output from Fig. 2.22(b) is displayed in Fig. 2.23(a)); cells displaying an average 458ex intensity below a threshold were removed from the final ratiometric calculation (see arrows in Fig. 2.22). From the distribution (Fig. 2.23(a)) it is also observed that the range of 488ex/458ex values covers cells which have twice the ratiometric value to other, genetically-identical cells. To ensure the behaviour contained within outputs was cell-to-cell ATP heterogeneity, and not an artefact of the assay, wild-type *E. coli* cells were imaged as a negative control, under the same conditions and protocol, and displayed negligible fluorescence,.

Further checks were considered within the *E. coli*-2m analysis. For the population of cells imaged, single-cell 488ex/458ex ratio was plotted against mean 458ex intensity (see Fig. 2.23(b)), and single-cell area (see Fig. 2.23(c)). These additional correlation checks were performed to tighten the confidence that the data was displaying a real effect of cell-to-cell ATP heterogeneity and not an artefact of, for example, cells with larger areas displaying an increased ratio (correlating to ATP level). To confirm our data was



(a)



(b)

Figure 2.22: **Segmentation of microscopy images.** Microscopy image of an *E. coli*-2m cell sample, displaying the emission profile after excitation at 488nm (a). Emission intensity was labelled in green for improved image analysis. Segmentation software provided the boundaries of recognised cells within the sample (b). Individual cells with an average 458ex intensity below a threshold are removed (pink arrows).

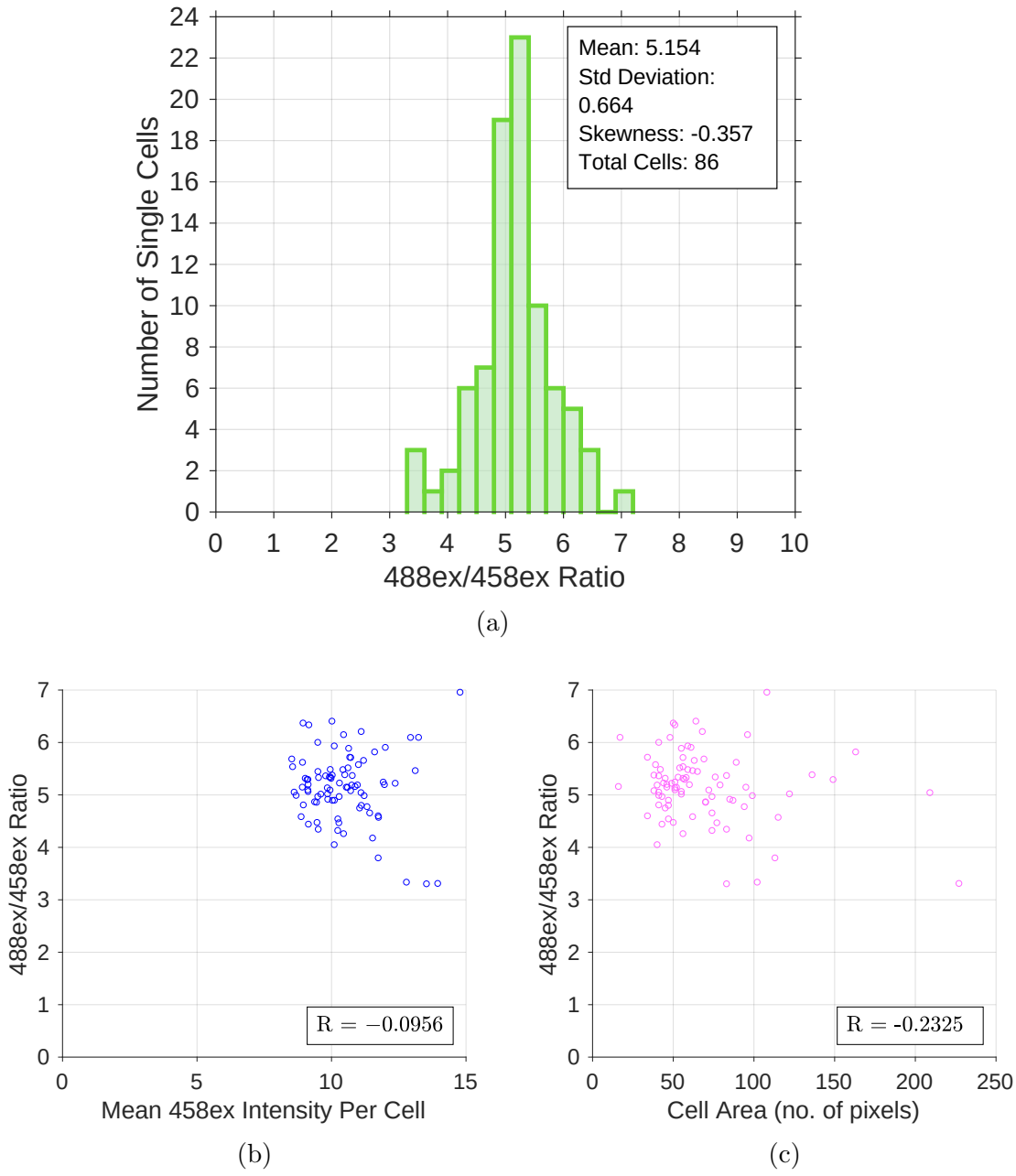


Figure 2.23: **Clonal *E. coli* populations display cell-to-cell heterogeneity in a stress-free environment.** (a) Negatively skewed distribution of 488ex/458ex ratio of *E. coli* QUEEN-2m cells in liquid culture, and scatter plots displaying 488ex/458ex ratio against (b) mean 458ex intensity, and (c) cell size of individual *E. coli* cells. In (b)-(c), correlation coefficient, R, is calculated to quantify the strength of the relationship between the two pieces of data.

displaying cell-to-cell ATP heterogeneity, we calculated the Pearson correlation coefficient, R, measuring how strong the linear relationship is between two variables. In both cases, there was a weak linear relationship (Fig. 2.23(b)-(c)). Together, our results qualitatively agreed with the original study [309] in the degree of cell-to-cell heterogeneity in ATP levels (see Fig. 2.23(a)).

Our original objective aimed to follow the imaging and ratiometric analysis with mapping single-cell ATP ratio levels to absolute intracellular ATP concentrations, similar to the original research [309]. Once the protocol was refined, the intention was to jointly track bacterial ATP levels and persister cell development in a clonal population prior to, during and after exposure to an antibiotic. After developing our protocol monitoring qualitative cell-to-cell ATP heterogeneity, the subsequent experimentation was disrupted due to COVID-19. Although we were unable to complete this research, a similar project was recently completed by Manuse *et al.* [166]. The group observed a general low energy mechanism of persister formation, strengthening previous reports of a link between ATP level and persister formation in bacteria [251, 298, 315], and supporting our theoretical prediction that energy availability modulates cellular-decision making (see Discussion).

2.5 Discussion

We have used a simple and general regulatory model, representative of the core of several known decision-making motifs [81, 108, 109, 205, 318], to explore how the decision-making capacity of a cell depends on the energy available to fuel the processes involved in this decision making. Our theory predicts that, across a broad range of biological contexts, differences in cellular energy levels will cause differences in the dynamics and stable outcomes of cellular decision-making, via modulation of the expression rates of interacting regulatory genes. More precisely, increased ATP levels both support the ability to transition into multiple stable states, and increase the separation between the attractors as energy increases, stabilising the decisions that can be made (Figs. 2.4-2.6). Concurrently,

without a sufficient energy supply, decision-making circuits lose the capacity to select different phenotypes (Figs. 2.4-2.6 and 2.10). Higher energy levels allow, and stabilise, separate and intermediate states.

Our findings – that the fate choices supported by decision-making architectures depend on available energy – exists independently of the other metabolic requirements of given precursor or descendant cell types. It is therefore predicted the influence of ATP variability on the dynamics of regulatory motifs can cause this general dependence, which exists in addition to specific metabolic budgets of the cell types in question.

The architecture we consider (Fig. 2.1(b)) can describe the behaviour of many naturally-occurring decision-making circuits, and our findings that energy availability influence this behaviour thus provide testable predictions for this class of systems. Accordingly, our preliminary experimental design (see Section 2.4) can be developed to readily test our theoretical predictions. Jointly tracking single-cell ATP concentrations within a clonal bacterial population, together with the activity of an intrinsic AMR mechanism (such as efflux pumps or persisters), before, during, and after antibiotic exposure, will allow new insights into the effects of energy availability on AMR, and cell-to-cell ATP heterogeneity in a bacterial population in different environments. More generally, our hypothesis can be tested by, for example, jointly tracking prokaryotic phenotype or eukaryotic fate decisions in a culture exposed to the same stimulus but titrated with different external concentrations of ATP. Our prediction that phenotypic decisions depend on ATP is already supported by several independent experimental observations, discussed below.

The observation that different levels of ATP may induce different cellular behaviours comes as increasing cell heterogeneity in ATP levels is being experimentally characterised. From our own experimental work, we observed significant ATP variability in a clonal *E. coli* population (see Section 2.4), qualitatively agreeing with the degree of heterogeneity in previous *E. coli* studies [309]. The connection between energy variability and cellular decision-making in bacteria has also been discussed previously [56, 251], further supporting the hypothesis that energy diversity may modulate the number of regulatory states

supported by a bacterium, playing an important causal role in inducing variability in cellular decision-making within bacteria.

Setting $a_1 = a_2 = 0$ reduces the system to consist of a bistable switch with cross repression, known as a toggle-switch. This motif is found at the core of several known decision-making architectures throughout biology, with an extensive body of associated literature [36, 49, 152, 155, 221, 281, 288, 299]. Notably, this motif occurs in the λ-bacteriophage switch from lysogenic (integrates its genome into the host cell genome) to lytic (actively reproduces using a host cell to manufacture more phages) state in *E. coli*. The lysis-lysogen decision can generally be described as being determined by two genes, *cI* and *cro*, which mutually cross-repress through their products; the outcome of high *cI* (*cro*) expression is *cro* (*cI*) repression and *cI* (*cro*) transcription, resulting in the lysogenic (lysis) state [5, 16, 18, 228]. The switch from lysogenic to lytic state occurs when the *cI* product is cleaved by RecA, a result of the response to DNA damage, derepressing *cro*, increasing Cro production and turning off *cI* [5].

This motif consisting of two genes with feedback loops is found at the core of several known decision-making architectures throughout biology, but of course, it will not exist in isolation in real biological systems (although theoretical studies have underlined how analysis of its dynamics can be valuable in understanding its downstream physiological significance [109, 118, 119]). However, our hypothesised link between energy variability and decision-making dynamics is not restricted to this motif. Relaxing the toggle-swtich picture used here, links between energy variability and cell fate are anticipated in a broad variety of other cell decisions. A pertinent example already mentioned is the formation of bacterial persister phenotypes in the presence of antibiotics. One regulatory motif involved here is the *hipBA* module, which expresses HipA (toxin) and HipB (antitoxin), which can form a complex that represses their own expression. Rotem *et al.* [244] found that cells become dormant if the level of HipA exceeds a threshold, and as the level of HipA exceeded the threshold further, it determined the duration of dormancy. This is similar to our model and architecture where the protein levels, x_1 and x_2, feed back into the

system and establish the generated cell fate. Current experimental observations, in both *S. aureus* [56] and *E. coli* [251], suggest a low-energy mechanism of persister generation. The application of our generalisable framework to the *hipBA* module could provide theoretical support into whether intracellular energy availability is a factor involved in the persister decision.

ATP variability in eukaryotic cell types has also been observed. Intracellular ATP has been quantified to vary between 3.7-3.9mM in HeLa cells [313], 3-4mM in living yeast cells [270] and over a several-fold change occurs in different tissues of a plant [63]. The connection between energy variability and cellular decision-making in eukaryotic cells has also been discussed previously [14, 51, 78, 248, 273]. In particular, the relationship between pluripotent stem cell behaviour and energy metabolism has been extensively studied [273]. Culturing human embryonic stem cells (ESCs) in low oxygen (suppressing oxidative phosphorylation and thus reducing energy production) prevents differentiation [78]. Differentiation of pluripotent stem cells (PSCs) to cardiomyocytes requires a high oxidative phosphorylation capacity [51], and human PSCs maintain higher mitochondrial membrane potential ($\Delta\Psi_m$) than their differentiated descendants [14]. In mouse ESCs, cells with high $\Delta\Psi_m$ remained in a pluripotent state while cells with low $\Delta\Psi_m$ largely adopted a single differentiated state [248]. These observations all align with our prediction of reduced decision-making capacity at low energy levels due to the energy dependence of regulatory networks (although the different metabolic requirements of different cell types will also play a role in these observations). Together, this further supports the hypothesis that energy diversity may modulate the number of regulatory states supported by a cell, playing an important causal role in inducing variability in cellular decision-making.

We predict that ATP variability will influence fate decisions across this broad variety of systems, organisms and branches of life. Generally, the propensity to generate different cell fates will be intrinsically affected by ATP supply (in addition to extrinsic signals that may also be linked to nutrient and energy availability), and for bistable regulatory motifs, higher ATP levels will promote undifferentiated states from which more fates are possible.

The influence of noise in gene expression [215, 217, 274] on the behaviour and stability of this system will also be valuable to link to specific biological situations. Together, the inclusion of noise and additional genetic actors will help shed further light on how cells are able to support different cell fates (in our model, how the number of available attractor states are accomplished). Some subtle and re-entrant behaviour was observed in the bifurcation dynamics associated with transitions between different numbers of distinct stable states. If this behaviour is general across other architectures, it may provide a potential optimum intracellular energy range for phenotypic diversity, enabling a cell to have superior adaption to stochastic extracellular environmental change.

CHAPTER 3

DYNAMIC BOOLEAN MODELLING REVEALS THE INFLUENCE OF ENERGY SUPPLY ON BACTERIAL EFFLUX PUMP EXPRESSION

One key factor contributing to AMR is the ability of bacteria to export drugs through efflux pumps, which relies on the ATP-dependent expression and interaction of several controlling genes. Recent studies have shown there exists significant cell-to-cell ATP variability within a clonal bacterial population, but the contribution of intrinsic cell-to-cell ATP heterogeneity is generally overlooked in understanding efflux pump behaviour. Here, we could apply the energy-dependent ordinary differential equation (ODE) framework from Chapter 2. However, while ODE models can in principle be used for GRNs of arbitrary size, the associated parameter space rapidly increases with network size. In this chapter, we consider how ATP availability influences two gene regulatory networks (GRNs) controlling the expression of efflux pump genes in two different bacterial species, by applying scalable framework which can explore the influence of ATP variability on GRN dynamics and observe how differences in energy availability can cause pronounced downstream heterogeneity in efflux gene expression.

Work in this chapter is taken from research published by Kerr *et al.* [130].

3.1 Heterogeneity in efflux pump expression

Protein complexes called efflux pumps are an important mechanism for intrinsic (and acquired) resistance to antibiotics [173, 174, 225, 256, 285]. Efflux pumps are capable of transporting a range of substances from the cell compartment to the extracellular medium, reducing intracellular accumulation of substances including antibiotics, and thus conferring resistance [198, 199, 226]. Their central role in facilitating clinically relevant AMR makes efflux pumps an attractive drug target, with much research directed at efflux pump inhibitors as a novel therapeutic to overcome efflux-mediated resistance [156, 168, 187, 219, 265].

A current mystery in the activity of efflux pumps is the causes and effects of cell-to-cell heterogeneity in their expression. Recent data demonstrates efflux pump gene expression is not homogeneous throughout the population – instead, bacterial populations display significant cell-to-cell variability [73, 213, 247]. Heterogeneity in efflux activity has been suggested to have further downstream influence on AMR. One study, by Sánchez-Romero and Casadesús [247], reported a correlation between bacteria with high efflux pump activity in the absence of antibiotic compounds and increased resistance to antibiotics. Additionally, a study by El Meouche and Dunlop [73] has shown a sub-population of bacteria with higher efflux pump expression has lower expression of mismatch repair genes, facilitating mutations which can lead to permanent genetic changes bestowing resistance. However, the underlying dynamics and influences of efflux heterogeneity in cellular populations remain to be uncovered.

One potential source of heterogeneity in gene expression dynamics and interactions is cell-to-cell differences in available energy supply [119, 131], which is increasingly being characterised experimentally in bacteria [309, 315] (and other species; see Chapter 1). Following theoretical work showing the potentially profound influence of this variability on gene regulatory networks [119, 131], we here investigate the hypothesis that cell-to-cell differences in available energy can contribute to cell heterogeneity in efflux pump expression, intrinsic resistance to antimicrobials, and response to different environmental

stressors.

Because of the experimental challenges involved in revealing general principles of the behaviour of many interacting components, we adopt a modelling approach grounded in and validated by experimental observations. We represent the interacting genetic architecture shaping efflux pumps expression as a GRN. These networks can be modelled using various mathematical approaches [125], including continuous [112, 115, 211, 291] and stochastic [11, 75] frameworks. In Chapter 2, we applied an energy-dependent ordinary differential equation (ODE) framework to a simple GRN, describing the behaviour of many naturally-occurring decision-making circuits. We predicted that differences in cellular energy levels will cause differences in the dynamics and stable outcomes of cellular decision-making. More precisely, without a sufficient energy supply, decision-making circuits lose the capacity to select different phenotypes. Higher intracellular energy levels increase the decision-making capacity, and also stabilise the available phenotypes.

These results from ODE modelling demonstrated the influence of ATP supply on regulatory dynamics. However, while ODE models can in principle be used for GRNs of arbitrary size, the associated parameter space rapidly increases with network size. This expansion makes it harder to investigate general principles, as each parameter's contribution must be investigated for a comprehensive analysis of the system. Because of this challenge, we sought a more scalable way to explore the influence of ATP variability on GRN dynamics. To this end, we adopt an alternative framework. Boolean network models (introduced by Kauffman [127] and Thomas [275]) coarse-grain the complex dynamics within gene regulation by describing the logical interactions contained within a network (allowing genes to be "on" or "off" and allowing simple positive or negative regulatory interactions between genes – see Fig. 3.1(a)). This framework, although a simplification of realistic biological events, has been successfully applied, for example, to the yeast cell-cycle network [150], T helper cell differentiation network [186], and the *Drosophila melanogaster* segment polarity network [6]; although to date few Boolean modelling studies focus on understanding bacterial virulence and resistance mechanisms [24, 61].

While Boolean models have been successfully used in a range of cases, the fact that their edges correspond to ATP-dependent processes has rarely been considered [23, 24, 113, 127]. This ATP dependence is important, because genetically-identical cells in a bacterial population can have substantially different levels of ATP [309, 315]. We therefore developed a simple but highly generalisable modification for including energy dependence in Boolean models of regulatory networks. We construct empirically-derived regulatory networks for efflux pump expression in *E. coli* and *Salmonella*, focusing primarily on behaviour at the transcriptional-level, and validate our models by comparing our results to independent experimental observations. Although these networks are reasonably simple, an ODE-based analysis would still involve a large set of parameters (many of which are not well characterised), so these cases serve to demonstrate our scalable Boolean approach as well as being of intrinsic scientific and translational interest. We then further analyse theoretical outputs, first by considering the case where available energy is not a limiting factor on the dynamics of the system. Subsequently, we consider how the dynamics of these processes change when levels of available energy vary between cells.

3.2 Methods

3.2.1 Literature-based model construction accounting for energy availability

We first sought to construct a Boolean modelling framework describing efflux pump expression dynamics, supported by data and yielding verifiable predictions about biological behaviour. We considered AcrAB-TolC, an efflux pump associated with clinically important drug resistance in *E. coli* [77, 173, 301] and *Salmonella* [27, 71, 225, 237]. TolC is constitutively expressed [277] and predicted to be integrated within at least seven additional efflux transport systems in *E. coli* [310, 316], suggesting the dynamics of *acrAB* are more limiting on producing the complete efflux pump system within each cell. We per-

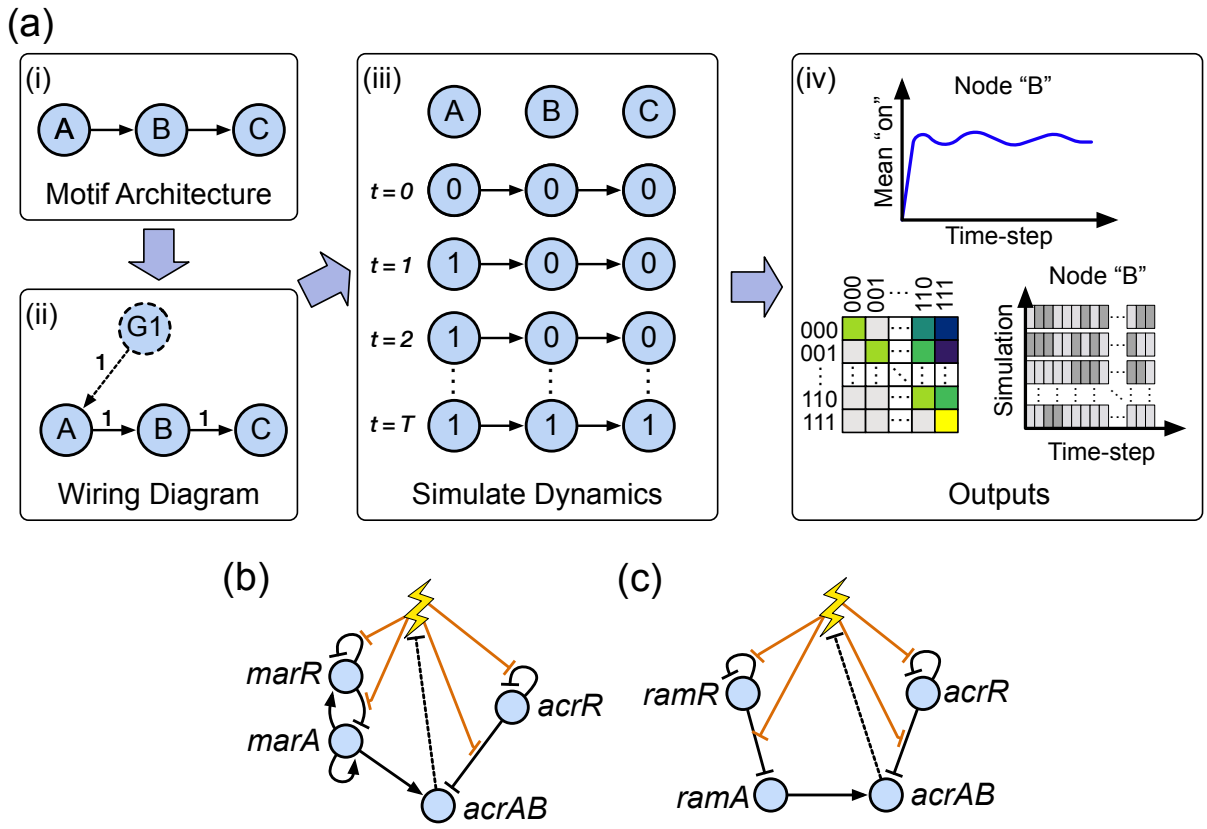


Figure 3.1: **Boolean network modelling of gene regulatory network architectures controlling efflux pump gene expression.** **(a)** Schematic overview of the Boolean network framework. Regulatory architectures (i) are converted into wiring diagram form (ii). These wiring diagrams include "ghost nodes" (G1) if necessary, providing a constant signal to ensure constitutively expressed genes are "on" unless switched "off" by a regulator. The considered network is repeatedly simulated over a set number of timesteps (iii), from some initial configuration, producing individual and averaged pictures of network behaviour through transition matrices and expression dynamics of each node per timestep (iv). In the given transition matrix (iv), '000' can transition to any of '000', '110', '111', and so on, with the probability of each transition given by the colouring in each entry. **(b-c)** Coarse-grained *E. coli* (b) and *Salmonella* (c) AcrAB efflux pump gene regulatory networks, containing a collection of genetic regulators (blue circles) that interact with each other (directed edges) and govern the expression level of efflux genes *acrAB*. Full layouts with parameterizations are given in Fig. 3.2. Antibiotics, or other noxious stressor substances (lightning bolt), can reversibly bind to a gene product (transcription factor) removing its ability to bind to DNA and regulate its target (displayed through orange edges), with consequent downstream effects through the network. Removal of the stressor is integrated through a repressive interaction from *acrAB* to the stress node (dotted edges).

formed a broad literature review of efflux pump gene regulation in *E. coli* and *Salmonella* to define a core GRN of *acrAB* in each species (Fig. 3.1(b)-(c)). This review, which is summarised below, defined the structure of our model, i.e. the nodes, regulatory interactions and edge weights in the corresponding wiring diagrams (Fig. 3.2(a)-(b)).

A key regulator of efflux-related drug resistance in *E. coli* is the multiple antibiotic resistance (*mar*) operon [54, 163]. This genomic region contains the *marRAB* operon, encoding three proteins named MarR, MarA and MarB [54]. MarR has a repressive function, acting on the basal expression of the *mar* operon in stress-free conditions [12, 54, 250]. MarA, the *marA* gene product, is a member of the AraC/XylS family of proteins which induces *acrAB* expression and undergoes positive feedback through auto-activation of the *marRAB* operon [169, 175]; in addition to *acrAB* and *marRAB*, MarA regulates the expression of numerous other genes, either directly or indirectly, in its regulon [25]. The final product from the operon, MarB, is a small periplasmic protein that represses *marRAB* but does so through an unknown mechanism and is therefore omitted from our regulatory network [290].

Salmonella has a second MarA-like protein called RamA, an AraC/XylS transcription activator synthesised from *ramA* [283]. *ramA* is a homologue of the *E. coli marA* gene, having 45% similarity [69]; while not present in *E. coli*, *ramA* is present in a wide range of Enterobacteriaceae. Like MarA in *E. coli*, RamA positively regulates *acrAB* expression in *S. enterica* [195, 236] by binding to the region upstream of *acrAB*. RamR is the transcriptional repressor of *ramR* and *ramA*, binding to the *ramRA* intergenic region [2, 26]. We include the *ramRA* regulatory region in *Salmonella*'s gene regulatory network (Fig. 3.1(c)) and omit the *mar* operon due to previous studies showing inactivation of *marR* and *marA* (as well as *soxR*, and *soxS*) did not affect the susceptibilities of a *S. enterica* serovar Typhimurium strain, whereas inactivating *ramR* resulted in a multidrug resistance phenotype [2]. This observation, along with other supporting studies, suggests the *ramRA* regulatory region primarily controls AcrAB and efflux-mediated multidrug resistance in *Salmonella* [195, 282]. The major roles of the *mar* operon and *ramRA* region in

regulating *acrAB* expression, in *E. coli* and *Salmonella* respectively, supports the choice of these regulators in our coarse-grained GRNs (Fig. 3.1(b)-(c)). Additional regulation of *acrAB* is provided locally in both species by *acrR* [162, 207]. Its product, AcrR, is known to repress both its own synthesis and *acrAB* from a divergently transcribed region [162]. Under normal conditions, *acrR* is transcribed at a basal level, preventing over-expression of *acrAB* [162, 302].

The expression of efflux pumps is inducible by the presence of environmental signals, or chemical compounds (such as antibiotics), which we will generally refer to as 'stressors'. The inducibility of efflux genes enables bacteria to adapt to a wide range of environmental conditions. Noxious substances, such as antibiotics, can bind to GRN products MarR, RamR and AcrR, inhibiting their ability to bind to DNA, and initiating a response cascade that leads to the up-regulation of efflux genes *acrAB* [7, 26, 139, 170, 261, 312]. In our model the expression of *acrAB* corresponds to the capacity of the assembled efflux pump to export molecules.

Having constructed the architecture of each network, we can proceed to query them about biological behaviour. To avoid the aforementioned parameterization issues in continuous models, we developed a general Boolean framework that captures the influence of energy availability. We now describe our framework in more detail.

3.2.2 Wiring diagrams

Wiring diagrams, consisting of nodes (genes) and edges (activatory or inhibitory regulatory interactions), schematically represent the interactions between elements within regulatory networks (Fig. 3.2). Edge weights for our model were chosen through a manual tuning process, seeking to reproduce our "training" observations given the imposed structure of the regulatory network from the literature review. In a Boolean network, edge weights encode logical interactions rather than specific physical rates or quantities, and therefore there is substantial flexibility in the particular values used (see Discussion). Instead of a Boolean modelling approach, an ODE methodology could be applied to the

efflux pump networks. The parameterisation of the ODE model would itself produce parameters that are unknown, however the Boolean framework, due to its coarse-grained approach, would produce fewer parameters to be estimated. Together, the nodes, edges, and edge weights produce a directed graph representation of the regulatory network. A stress, such as antimicrobial exposure, is included through the 'stress node', S, within the wiring diagram. A stress input is modelled as repressing gene-gene regulatory edges, representing a 'repressive' modulation of transcription factors (MarR/RamR and AcrR). The capacity of the assembled efflux pump (AcrAB-TolC) to remove stressors is integrated through a repressive regulation from *acrAB* to S.

In *E. coli*, the *marRA* operon is separated into its individual components to allow each regulation of *marRA* to be captured, but are coupled in all iterative updates; the inability to model a node that both positively and negatively regulates itself is a limitation of the Boolean modelling framework. Experimental data has shown transcription of *marA* and *acrA* increase in *marR* mutant cells [191], and MICs of various antibiotics increase in an *acrR* mutant strain [206] (consistent with the limited repressor function of AcrR [162]). Together, these suggest, under normal conditions, edge weights of interactions from *marR* to *marRA* and *acrR* to *acrAB* should be dominant. Inactivation of the *ramR* gene upstream of *ramA* resulted in increased expression of *ramA* and the AcrAB efflux pump in *Salmonella* [2]. This requires the edge weight from *ramR* to *ramA* to outweigh *ramA* basal expression under normal conditions. Constitutive gene expression is captured using "ghost nodes". These nodes are permanently "on" in the network, supplying a constant positive influx to a node.

Interaction matrix J, used in node updates (Section 3.2.5), is constructed from the data contained within a wiring diagram. A weighted edge from node j to node i, is represented by element $\mathrm{J}_{i,j}$ within J. Each matrix element can take a positive, negative or zero value; edge weight sign corresponds to a negative, positive or absence of regulation respectively. Since edges are directed, $\mathrm{J}_{i,j} \neq \mathrm{J}_{j,i}$, except for $i = j$ or $\mathrm{J}_{i,j} = \mathrm{J}_{j,i} = 0$. Matrix J is implemented in numerical simulations of the Boolean model and has dimension N-

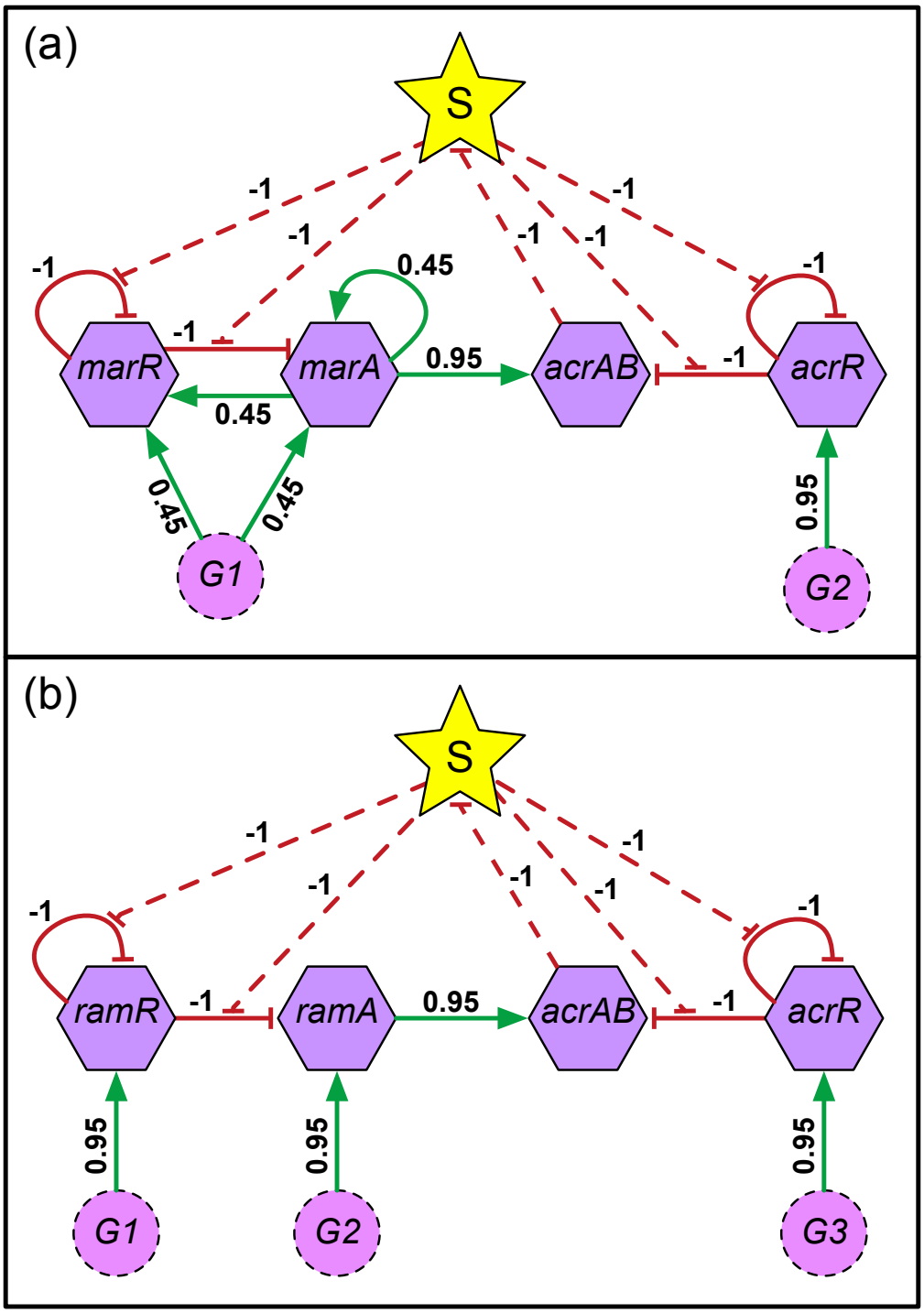


Figure 3.2: **Schematic representation of network wiring diagrams.** Wiring diagrams for (a) *E. coli* and (b) *Salmonella* efflux pump GRNs (Fig. 3.1(b) and (c), respectively). GRN interactions directly modulate network nodes in the wiring diagram depiction, and stress-related regulation (dashed edges to and from node S) modulates network regulatory edges; edge weights within the wiring diagram are estimated from the literature, and enable each element in the network a chance to act. Each regulatory interaction is colour coded to display positive (green) or negative (red) operations.

by-N (N is the total number of nodes in a wiring diagram, including ghost and stress nodes). As an example, the interaction matrix J in Equation (3.2.1) is presented for a simple network containing two nodes A and B, with A activating B and B inhibiting A.

$$\mathrm{J} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{3.2.1}$$

3.2.3 Initial conditions

Two configurations of initial node states were considered in simulations of the *E. coli* and *Salmonella* wiring diagrams. These were named the "unstressed" and "stressed" initial condition, with the initial node states set as follows:

	marR	*marA*	*acrR*	*acrAB*
Unstressed	0	0	1	0
Stressed	1	1	0	1

Table 3.1: ***E. coli* node initial values.** Table defines the stressed and unstressed initial conditions for *E. coli* gene regulatory network components (Fig. 3.1(b)).

	ramR	*ramA*	*acrR*	*acrAB*
Unstressed	1	0	1	0
Stressed	0	1	0	1

Table 3.2: ***Salmonella* node initial values.** Table defines the stressed and unstressed initial conditions for *Salmonella* gene regulatory network components (Fig. 3.1(c)).

3.2.4 Modulation of transition rates allows energy variability to be captured in GRN models

The edges within wiring diagrams (Fig. 3.2) often illustrate combined processes such as transcription, translation and phosphorylation. The processes that are combined within each edge are reduced into a single parameter within our Boolean model, and can be expressed as individual parameters within ODE models [314]. The parameters within our

Boolean model can capture similar behaviour compared to the corresponding parameters within an ODE model. For example, a stronger transcription rate increases the rate with which regulatory interactions can be applied from node X to node Y in the Boolean model, increasing the corresponding edge weight, and similarly it would increase the transcription rate value in the ODE model. However, the information contained within edges in regulatory networks is coarse-grained, omitting a substantial amount of important biological detail. Several of the processes represented by these edges require ATP as an energy source [62, 122], so there exists a core energy dependence in the dynamics modelled within each wiring diagram.

To address this and capture the fact that all the processes involved in gene expression depend on ATP, we constructed a modelling approach modulating the strength of regulation from node X to node Y according to cellular energy levels. This reflects the fact that, for example, cells with low ATP will have a lower rate of transcription elongation per timestep [62]. We interpret this modulation as influencing the rate with which regulatory interactions can be applied from node X to node Y. The rationale for this picture is that if the expression of regulatory components is slowed by reduced energy levels, the corresponding biochemical interactions will occur at a lower rate.

In the absence of energy variability, our asynchronous Boolean modelling framework interprets regulatory interactions as stochastic events occurring with a characteristic mean rate of one per timestep per interaction. To include the influence of energy availability on the system, we allow this rate to be a function of energy level. To capture an order-of-magnitude range across cells, we define three energy levels "high", "intermediate", and "low", respectively involving a characteristic mean rate of 1 (hence, all interactions are always realised), 1/2, and 1/10 per simulation, corresponding to the aforementioned order of magnitude range observed in biology. We thus modulate the probability of any given interaction being realised in a given simulation timestep by the level of available cellular energy.

3.2.5 Node update function

Nodes within the Boolean network are updated iteratively, using an update function. An update function defines the value of each node according to its regulatory nodes and remains fixed for all timesteps [96]. We consider a threshold update function to update the state of node i at timestep $t = \tau + 1$, denoted $\sigma_i(\tau + 1)$. The argument of the update function consists of the regulatory nodes of node i, the states of these nodes at $t = \tau$ and the N-by-N interaction matrix J, where N is the total number of nodes in the wiring diagram. The summation value determines the state of node i at $t = \tau + 1$.

During an active stress time frame, the state of stress node S (Fig. 3.2) is imposed to be "on" at the start of each timestep, regardless of the behaviour from the previous update. This enables the framework to model a bacterial cell being continuously exposed to an antibiotic concentration throughout a set time period. Once the final timestep within the stress period is reached, S is allowed to be switched "off" from regulation within the network.

We consider the update function displayed in Equation (3.2.2). Here, a node requires a non-zero sum of its inputs to change state from $t = \tau$ to $t = \tau + 1$. In the case where the sum is zero, the node state remains unchanged.

$$\sigma_i(\tau+1) = \begin{cases} 1 & \text{if } \sum_{j=1}^{N}(\mathrm{J}_{i,j} \cdot \sigma_j(\tau)) > 0, \\ \sigma_i(\tau) & \text{if } \sum_{j=1}^{N}(\mathrm{J}_{i,j} \cdot \sigma_j(\tau)) = 0, \\ 0 & \text{if } \sum_{j=1}^{N}(\mathrm{J}_{i,j} \cdot \sigma_j(\tau)) < 0. \end{cases} \tag{3.2.2}$$

Alternate threshold update functions are possible. We could, for example, consider the conditions that set a node as "on" if there is an overall positive input, and "off" otherwise. However, this would automatically remove a stress in the system by setting the stress node state to "off" once reaching the end-point of the stress time frame, instead of allowing the efflux node to "expel" the stress through feedback to the stress node S

(Fig. 3.2). By choosing update rule (3.2.2), it allows the efflux pump node to have the capacity to remove stress in the network, through regulatory interactions, modelling the behaviour observed in bacteria.

At each timestep, the binary vector containing all node states is the global, or network, state [125]. The set of all possible global states forms the state space of a network [96]. The state space for *Salmonella*'s GRN is of size 2^M, as we are interested in the dynamics of the M-component GRN network being modelled, and 2^{M-1} for *E. coli*, due to the coupling of *marR* and *marA*.

3.2.6 Node update methods

There are two methods for updating global states, synchronous and asynchronous. In the synchronous mode, all the nodes (including the stress node) are updated at every timestep, with the state of each node at $t = \tau + 1$ depending on the states of its inputs at $t = \tau$. The deterministic nature of synchronous Boolean models, combined with each model containing a finite state space (2^M or 2^{M-1}, where M is the number of elements in the original GRN), guarantees that a synchronously updated time series simulation will eventually end at a steady-state attractor, or a cycle of repeating global states [96]. It also means the network will always reach the same state for a given initial condition and number of timesteps. We do not consider synchronous updating here due to the artefacts that can arise from its imposition, and its inability to model stochastic processes in the case of maximal energy availability. A more complex, but biologically realistic, protocol is the asynchronous mode, which gives a reasonable overview of the random dynamics in the cell and better captures transient behaviour. This method updates nodes of the system according to the last update of their regulator nodes, either from the previous or current iteration [113]. Previous studies have employed numerous different versions of asynchronous updating [47, 61, 101, 297], including methods where all nodes are updated according to a random sequence, or one randomly selected node is updated at a time step [245]. In our framework the asynchronous protocol updates η nodes per timestep, with

η drawn from a Poisson distribution (as the number of nodes updated at one timestep is independent of all other timesteps) with a mean equal to the number of original elements in the regulatory network (M), plus the stress node. Additionally, each node can be updated more than once per timestep and the state of each node at $t = \tau + 1$ depends on the states of its inputs at the most recent update. We used an asynchronous protocol because we believe it better captures the stochasticity in the cell.

3.2.7 Methods to analyse the Boolean network

In summary, our model represents genes as nodes in a network, where each gene can be "on" (expressed) or "off" (not expressed) at a given time. Edges describe activating and repressing interactions, and the state of a network changes over time with asynchronous update rules following these interactions (Sections 3.2.2 and 3.2.6), capturing the stochastic nature of gene regulatory dynamics. Edge weights for our model were chosen through a manual tuning process, seeking to reproduce our 'training' observations given the imposed structure of the regulatory network from the literature review. In a Boolean network, edge weights encode logical interactions rather than specific physical rates or quantities, and therefore there is substantial flexibility in the particular values used (see Discussion). To explicitly capture the fact that the fundamental processes (hence all gene interaction "arrows") involved in gene expression depend on ATP, we modulate the rate with which regulatory interactions from node X to node Y are applied according to cellular energy levels (Section 3.2.4). This reflects the fact that, for example, cells with low ATP will have a lower rate of transcription elongation per timestep [62], slowing the dynamics of the biochemical intermediates involved in these regulatory interactions, and thus slowing the interactions themselves. While fast fluctuations in ATP supply may exist within cells, as gene expression processes are dynamically slower, we assume the network is exposed to a time-averaged ATP level, and thus define three energy levels "high", "intermediate", and "low", respectively involving a characteristic mean rate of 1, 1/2, and 1/10 per simulation.

Taken together, the wiring diagram, initial conditions, update functions and update method form the Boolean network model. Time-dependent numerical solutions of the presented Boolean model were generated in Python 3.8.2 using modules *numpy*, *pandas* and *random*. Numerical solutions were calculated from prescribed initial conditions covering the stressed and unstressed state of each network (Section 3.2.3). Time-series figures, displaying the expression profiles of each node, were constructed from the mean of node expression and standard deviation per timestep, from 10^4 simulations. Expression level coefficient of variation (ELCV) is calculated for each timestep from dividing the standard deviation of the number of simulations a node is "on" at timestep i, by the mean number of simulations a node is "on" at timestep i, and corresponds to population variability in gene expression. A low ELCV is interpreted as a population displaying negligible cell-to-cell variability in gene expression activity, and high ELCV values suggest substantial cell-to-cell variability exists in the population. We will see in Section 3.3.2 that individual simulations exhibit pulses of gene expression. Therefore, an additional metric, the coefficient of variation of pulse lengths (PLCV), is calculated to investigate the variability of efflux pump gene expression pulses (see Section 3.3.5).

Visualisations of "single-cell" expression profiles, and pulse statistics, were constructed from 20 and 5000 randomly selected simulations respectively. Transition matrices were constructed from the simulation data and visualised through a heatmap structure, from 10^7 simulations. Each entry of the transition matrix is a non-negative real number representing the probability of transitioning between two network states from timestep $t = \tau$ to $t = \tau + 1$. Clustering of global states was performed using the same data, with a Euclidean distance matrix from modules *scipy.cluster* and *scipy.cluster.hierarchy*. All scripts are openly accessible through:

`https://github.com/kerr-mathbio/boolean-modelling-efflux-grns.git`.

3.3 Results

In this section we initially validate both of our models by comparing our results to independent experimental observations. We then further analyse theoretical outputs, first by considering the *E. coli* network in the case where available energy is not a limiting factor on the dynamics of the system, and then when levels of available energy vary between cells. The results for *E. coli* are then compared to those from *Salmonella*.

3.3.1 Efflux network models reproduce experimentally observed behaviour in *E. coli* and *Salmonella*

Having constructed our model based on a set of experimental observations, we next asked whether it made predictions that could be tested using independent experimental data. In Table 3.3 we present a table assessing the predictions of our model, for our choice of edge weights, with respect to a multiscale set of experiments at both the population and single-cell level, where agreement is demonstrated across a wide variety of observations (Table 3.3). Estimates of the edge weights were obtained from balancing the data from relevant literature and enabling each component of the networks to function (from performing preliminary simulations with variations in edge weight choices). The overall behaviour of both systems – which we will describe in detail below – predicts a range of single-cell and population-level features observed in independent experiments, including pronounced gene expression heterogeneity at the single-cell level, prominent increases in gene expression when a short or long stress signal is provided, the expression of repressive genes in response to stress, and the details of expression levels of different actors in the GRN in the presence and absence of stress.

Table 3.3: **Theoretical model reproduces realistic behaviour in *E. coli* and *Salmonella*.** Each row in the table describes experimentally observed behaviour (with references) and theoretical results arising from our model that capture analogous behaviour. References supporting experimental observations are labelled according to the dependence of information contained within each reference that is applied to construct the GRNs. The reference markers show: (●) a dependent study where observations are used to flow into GRN construction, making them more difficult to disambiguate from model construction; (●●) a study where the mechanism of the stated observation is unspecified, but other results are used in model construction; (●●●) an independent study where the observation tests the model predictions. In summary, references with more dots have a higher strength of independence, and are a better validation of the model.

Experimental Observation	**Theoretical Observation from Boolean Model**	**References**
An isogenic, uninduced *E. coli* population displays significant cell-to-cell variability in *marA* expression.	In the absence of stress, the simulated behaviour of *marA* displays significant cell-to-cell expression variability at the population and single-cell level. At low, intermediate, and high energy availability, *marA* ELCV (expression level coefficient of variation) is approximately 1.0 in each energy sub-population (Fig. 3.5 (i)). The population-level observations predict a collective population of low, intermediate, and high energy sub-populations of cells would contain high cell-to-cell variability in *marA* expression. A qualitative investigation of *marA* expression dynamics in simulated single-cells further supports the heterogeneity observed at the population-level (Fig. 3.9).	●●●[74]: Fig. 1(c), 2(b) and 4(c). Fig. 1(c) shows cell-to-cell variability in *marA* expression at different time points. Fig. 2(b) plots single-cell readouts of the *marA* promoter (at $t = 0$ min). Fig. 4(c) is a histogram showing frequency (%) of cells with a given *marA* expression level.

An *E. coli* population, under normal conditions, exhibits significant cell-to-cell heterogeneity in the expression of stress response genes *acrAB*.	In stress-free conditions, simulations predict pronounced cell-to-cell variability in *acrAB* expression, stabilising at an ELCV between 1.0-1.35 for low, intermediate, and high energy sub-populations (Fig. 3.5 (i)). Individual simulation plots (Fig. 3.10), representing single-cell expression of *acrAB*, qualitatively predict heterogeneous pulses of *acrAB* expression between cells prior to any stress period. Consistent with this observation, calculating the pulse length coefficient of variation (PLCV) predicts low and intermediate energy sub-populations exhibit significant variability in pulse length (Fig. 3.14 (ii) and (iv)).	[73]: Fig. 2(b) plots single-cell *acrAB* promoter activity. [213]: Fig. 3(b) plots single-cell accumulation of ethidium bromide (EtBr)[1].

[1]This observation applies to a broader range of efflux genes, other than the one we consider, as well as cell membrane protein complexes involved in the uptake of small molecules (e.g. porins). However, *acrAB* are generally considered to be the more dominant genes.

An unstressed *Salmonella* population displays significant inhomogeneity in *ramA* promoter expression.	At the population-level, simulations predict pronounced variability in *ramA* expression (calculated using the ELCV). Increasing the amount of energy availability, increases the variability, with the ELCV ranging from 0.68 to 1 at equilibrium (Fig. 3.8 (i)). Simulated single-cell dynamics (Fig. 3.11) qualitatively predict significant cell-to-cell heterogeneity in *ramA* expression in stress-free conditions. Taken together, simulations predict *Salmonella* cells display pronounced cell-to-cell heterogeneity in *ramA* expression throughout the population in the absence of stress.	●●●[139]: Fig. 2(b) plots individual cell *ramA* promoter activity.

Salmonella enterica serovar Typhimurium, under normal conditions, exhibits heterogeneous cell-to-cell efflux pump expression in a population.	Population-level simulations (Fig. 3.8 (i)) predict high variability in *acrAB* expression within low, intermediate and high energy sub-populations (calculated through the ELCV). The ELCV increases with the amount of energy available, ranging from 0.86 to 1.45, suggesting significant heterogeneity as energy increases (Fig. 3.8 (i)). In agreement, simulated *Salmonella* single-cells qualitatively predict heterogeneous cell-to-cell *acrAB* expression in low, intermediate, and high energy sub-populations (Fig. 3.12). Taken together, this biologically suggests a *Salmonella* population would contain considerable cell-to-cell variability in *acrAB* expression in the absence of stress.	●●●[247][1]: Fig. 5(a)-(b) plots EtBr accumulation in an isogenic culture of *S. enterica*.

[1]Data is not *acrAB* specific, however, AcrAB is the major efflux pump in *Salmonella*, suggesting it is highly probable cell-to-cell differences in EtBr accumulation correlate to variability in *acrAB* gene activity, not the minor players. To support this, Fig. 5(b) shows a strong similarity in EtBr accumulation for wild-type (WT) cells in the presence of an efflux pump inhibitor and $\Delta acrAB$ cells.

The *marRAB* operon can be induced by stress compounds, increasing transcriptional activity of *marRAB* in wild-type *E. coli*. Data shows tetracycline [12, 99, 263], chloramphenicol [99], sodium salicylate [55, 129, 169], some redox-cycling compounds [250] and the uncoupling agent carbonyl cyanide m-chlorophenyl hydrazone [250] all induce *mar* transcript.	The application of stress within simulations triggers a response in *marRA*, increasing the mean expression within low, intermediate and high energy sub-populations (Fig. 3.3 (ii)). Simulated single-cell behaviour is consistent with population-level predictions, qualitatively displaying an increase in transcriptional activity of *marRA* when induced by a stressor (Fig. 3.9).	●●●[55]: Fig. 2 lanes 1 and 2. ●●●[99]: Fig. 5(a) lanes 1-3. ●●●[129]: Fig. 2(a)-(b), column 1. ●●●[263]: Table 1, row 1. ●●[12]: Fig. 4, WT column. ●●[169]: Fig. 6. ●●[250]: Table 3, column 3.

Stress molecules, including fatty acids, bile salts, ethanol and NaCl [162, 163], induce *E. coli* *acrAB* expression, increasing the transcript level.	Population-level simulations predict an induction in *acrAB* expression in the presence of a simulated stressor, for all energy levels (Fig. 3.3 (i)). Simulated single-cell stress responses suggest the increase in transcriptional activity is more pronounced and more dynamic as energy availability and stress period increase (Fig. 3.10).	●● [162]: Fig. 3 lanes 1-2 and 5-6. ●● [163]: Fig. 5.
Transcription of *acrR* increases under ethanol and NaCl stress conditions.	In simulated stress conditions, the behaviour of *acrR*, in both *E. coli* and *Salmonella*, displays an induction of expression at the population level. At low, intermediate, and high energy availability, mean *acrR* expression increases (Fig. 3.3 (ii) and Fig. 3.6 (ii) for *E. coli* and *Salmonella* respectively). The population-level observations suggest a collective population of low, intermediate, and high energy sub-populations of cells would display an increase in *acrR* expression when exposed to a stressor.	●● [162]: Fig. 3(a), compare lanes 3 and 7 to lanes 4 and 8 respectively (this instance is for WT *E. coli* cells).

Noxious substances increase *ramA*, and subsequently *acrAB*, expression in wild-type *Salmonella* cells. Data shows bile [26, 195], indole [139, 193, 194], select biocides [139] and antibiotics (including chloramphenicol, ciprofloxacin, cloxacillin and rifampicin) [139], induce *ramA* and *acrAB* expression.	Population-level behaviour in simulated stress conditions predict an initiation of a stress response at each energy level. This response generally consists of an initial increase in transcriptional activity of *ramR* and *ramA* within each energy sub-population (Fig. 3.6 (ii) and (iv)). The induced *ramA* expression positively modulates efflux pump genes *acrAB*, increasing the mean expression of *acrAB* in each sub-population (Fig. 3.6 (i) and (iii)). In agreement, simulated single-cell dynamics qualitatively predict an increase in *ramA* transcriptional activity (Fig. 3.11), followed by a transition to the active efflux expression state (Fig. 3.12) in the presence of stress, within low, intermediate, and high energy *Salmonella* single-cells.	●●●[139]: Fig. 3(a)-(c) and Fig. S5. ●●●[193]: Fig. 1(c-2) and Table 2. ●●●[194]: Fig. 1(a)-(b). ●●[195]: Figs. 1(a), 2 and 3. ●[26]: Fig. 1(a).

3.3.2 Stress-free efflux network behaviour with a maximal cellular energy budget

To dissect the dynamics of these networks in more detail, we explored the stress-free population-level behaviour of each network. In the absence of stress, the population-wide mean behaviour in expression levels approaches a steady state (pre-stress dynamics in Fig. 3.3). We found at high energy availability, initial conditions had no influence on this population-level equilibrium of the *E. coli* and *Salmonella* networks (compare Figs. 3.3 and 3.4 for *E. coli*, and Figs. 3.6 and 3.7 for *Salmonella* respectively). We therefore proceed by investigating network behaviour starting from a particular initial state – which models an "unstressed" baseline condition – in the rest of the study and focus on the *E. coli* network before comparing the behaviour of the two networks. We note that when allowing for energy variability within the model, the population-level equilibrium of the *E. coli* and *Salmonella* networks was again unaffected by the choice of initial condition (compare Fig. 3.3 to Fig. 3.4 for *E. coli*), supporting the use of the unstressed initial condition within this investigation.

To explore the variability within the equilibrium population without energetic limitations, we calculated the ELCV (expression level coefficient of variation) for each genetic component in the *E. coli* network (Fig. 3.5 (i)). In the absence of stress, ELCV, corresponding to population variability in gene expression, ranged between 1.0-1.35 at equilibrium (see pre-stress dynamics in Fig. 3.5 (i)), suggesting significant cell-to-cell heterogeneity exists in the expression of each gene within the *E. coli* network when there are no energetic limitations. This diversity is due to population-level asynchrony in expression dynamics; in simulated individual bacteria, network components are expressed in heterogeneous pulses (pre-stress dynamics in Figs. 3.9 and 3.10). This diversity, particularly in *acrAB*, suggests that the network structure itself may support a degree of "bet-hedging" in the absence of stress, where some members of a population are primed to respond to stressors at any given time, agreeing with experimentally observed behaviour in *Salmonella enterica* [247] (see Discussion).

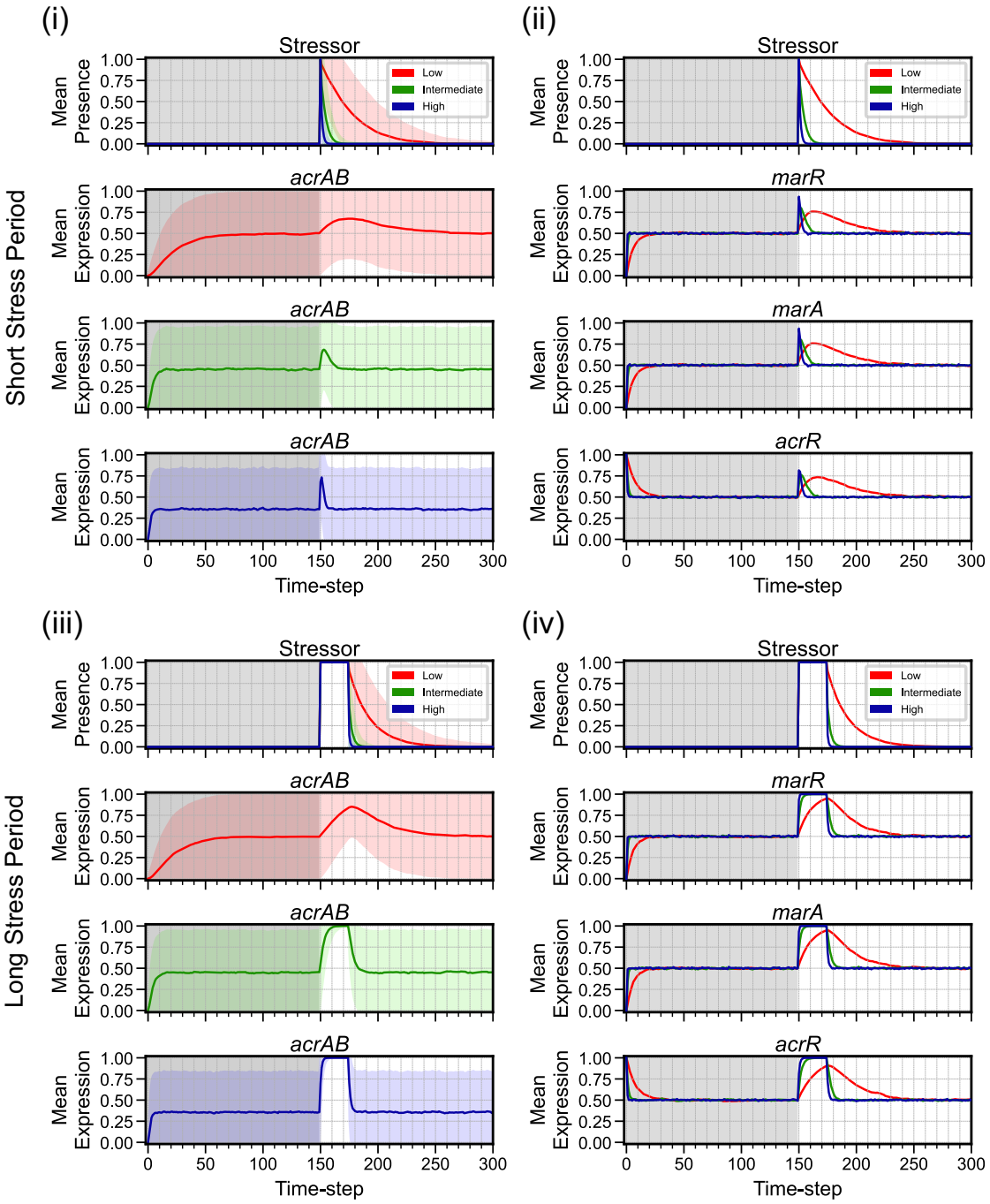


Figure 3.3: **Mean expression dynamics for *E. coli* from the unstressed initial condition in response to a short or long stress period.** Time-series dynamics of mean expression (coloured line for low (red), intermediate (green) and high (blue) energy levels) for ((i), (iii)) efflux genes *acrAB* and ((ii), (iv)) remaining regulatory components in *E. coli*; all panels include stressor behaviour. Behaviour is shown when the network is exposed to a (i)-(ii) short ($t = 150$) or (iii)-(iv) long ($t = 150$-174) stress period, from the unstressed initial condition. The initial time region of each regulatory component prior to stress application (grey shaded region) displays the complete stress-free behaviour. Note: Standard deviation (corresponding coloured shaded area) is included within efflux panels to represent the uncertainty, but the actual expression level cannot exceed the $[0, 1]$ domain.

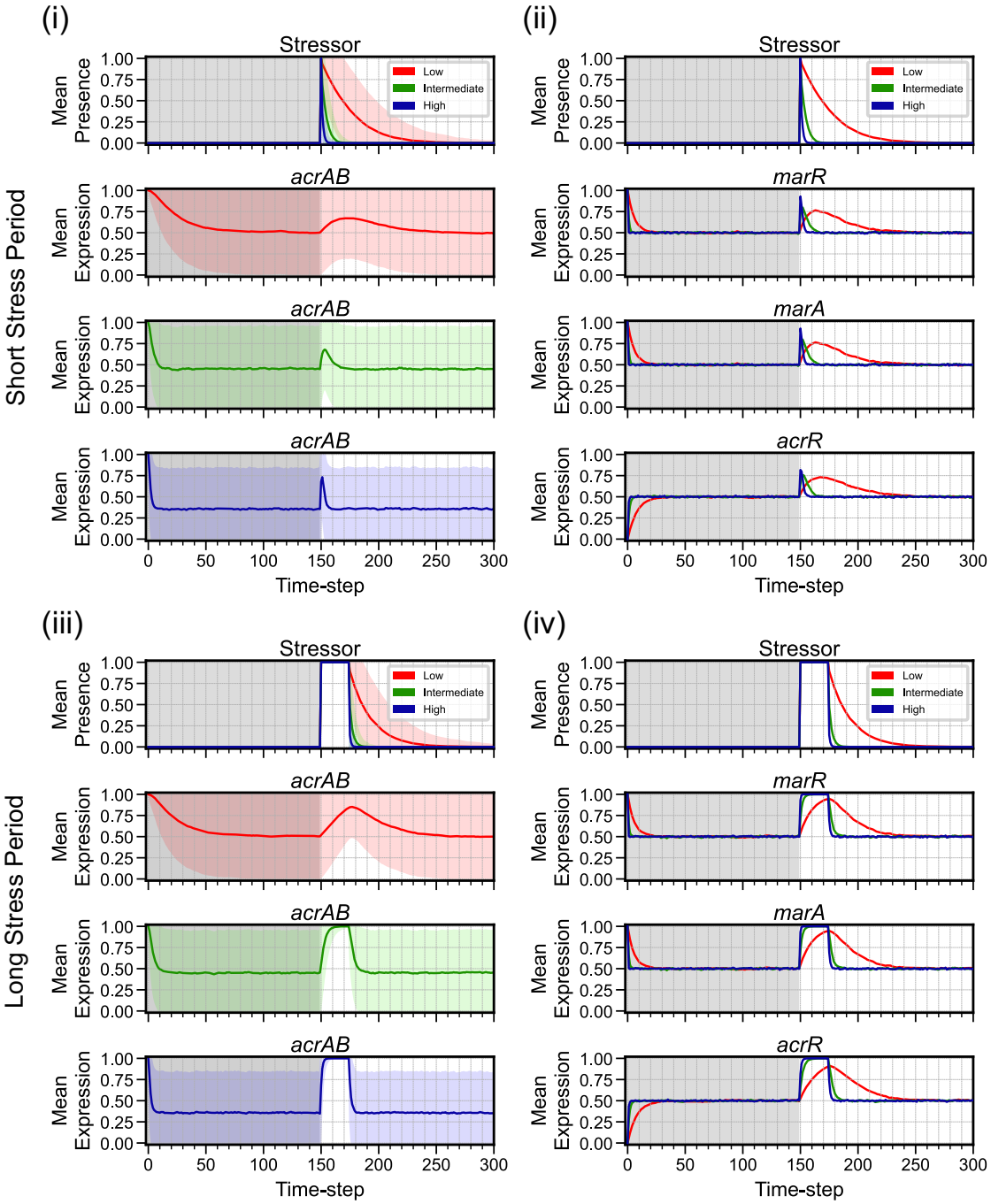


Figure 3.4: ***E. coli* mean expression dynamics from the stressed initial condition.** Time-series dynamics of mean expression (coloured line for low (red), intermediate (green) and high (blue) energy levels) for ((i), (iii)) efflux genes *acrAB* and ((ii), (iv)) remaining regulatory components in *E. coli*, from the stressed initial condition; all panels include stressor behaviour. Behaviour is shown when the network is exposed to a (i)-(ii) short ($t = 150$) or (iii)-(iv) long ($t = 150$-174) stress period. The initial time region of each regulatory component prior to stress application (grey shaded region) displays the complete stress-free behaviour. Note: Standard deviation (corresponding coloured shaded area) is included within efflux panels to represent the uncertainty, but the actual expression level cannot exceed the $[0, 1]$ domain.

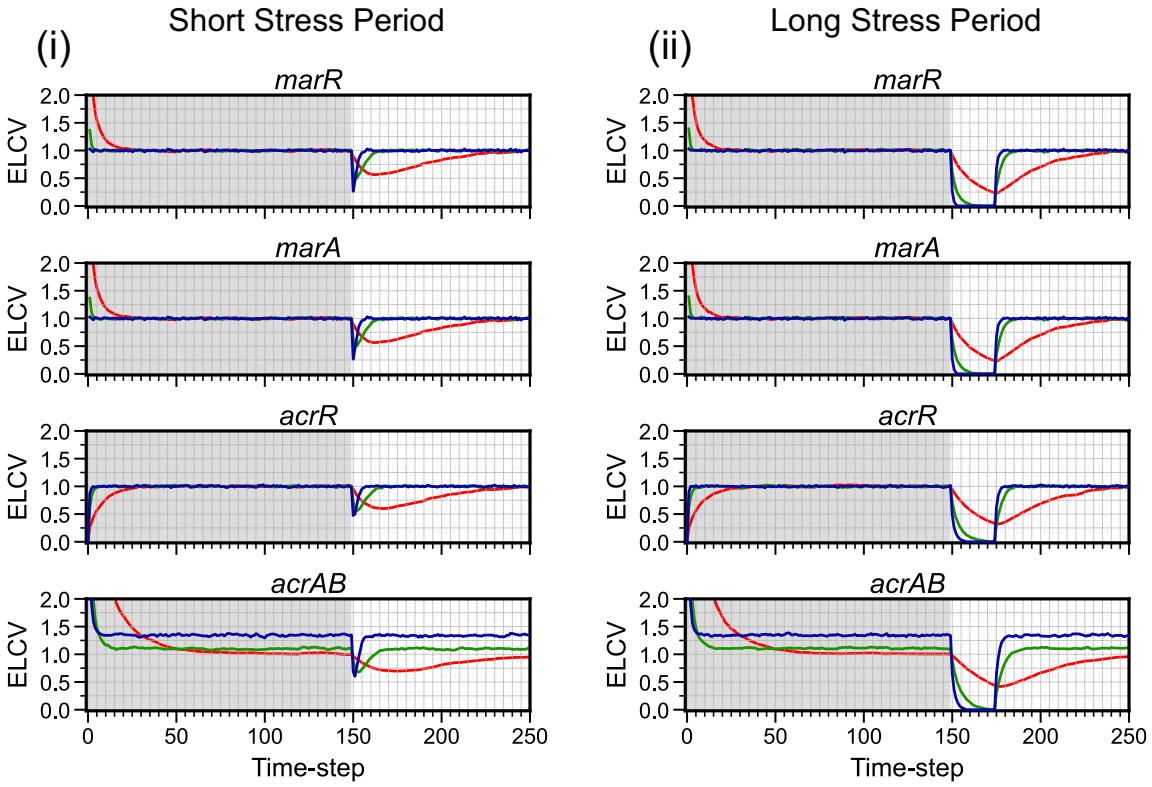


Figure 3.5: ***E. coli* expression level coefficient of variation.** Time-series dynamics of ELCV (expression level coefficient of variation) for each regulatory component in *E. coli*'s GRN, from the unstressed initial condition (coloured line for low (red), intermediate (green) and high (blue) energy levels). Behaviour is shown when the network is exposed to a (i) short ($t = 150$) or (ii) long ($t = 150$-174) stress period. The initial time region of each regulatory component prior to stress application (grey shaded region) displays the complete stress-free behaviour.

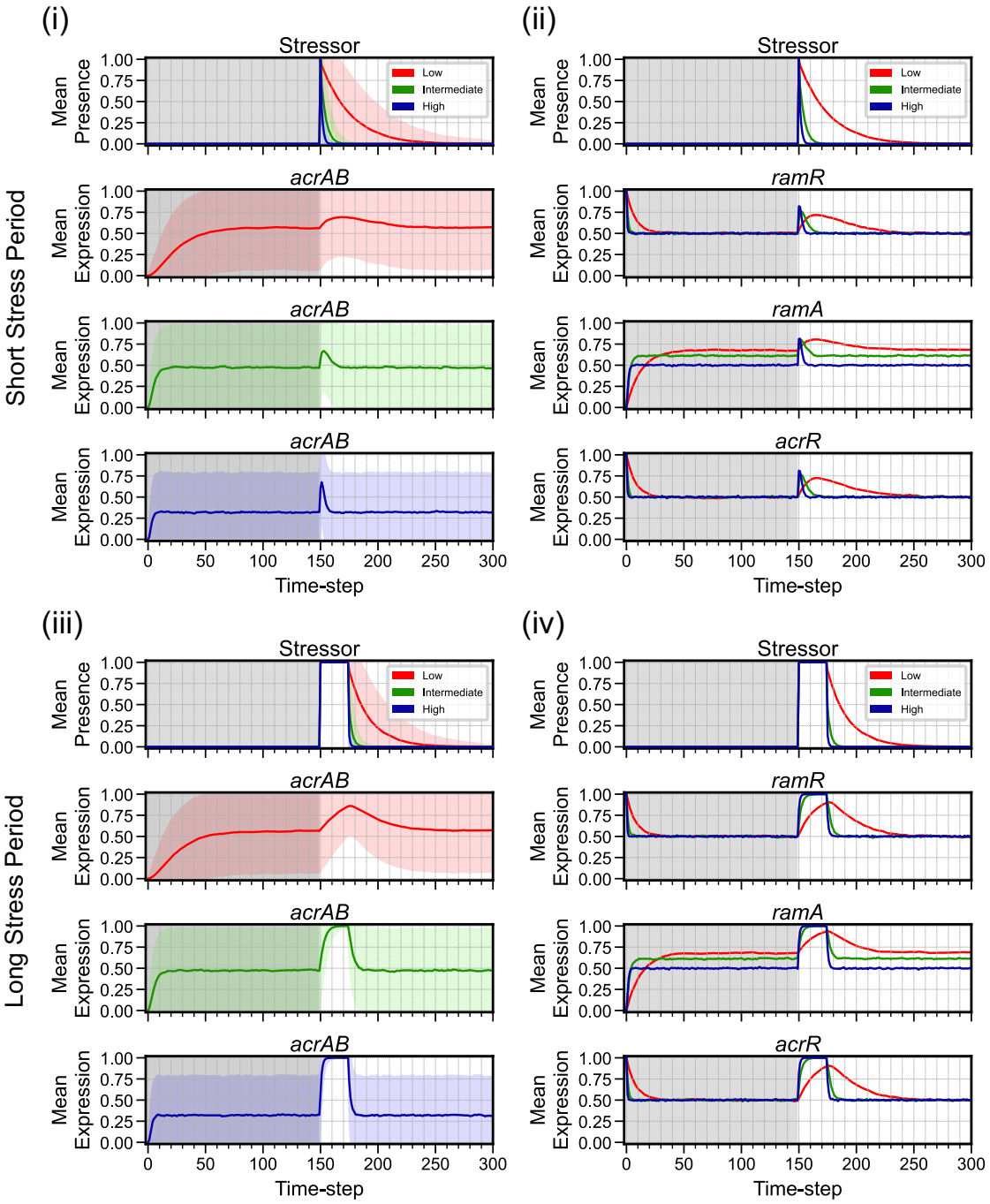


Figure 3.6: **Mean expression dynamics for *Salmonella* from the unstressed initial condition in response to a short or long stressor duration.** Time-series dynamics of mean expression (coloured line for low (red), intermediate (green) and high (blue) energy levels) for ((i), (iii)) efflux genes *acrAB* and ((ii), (iv)) remaining regulatory components in *Salmonella*, from the unstressed initial condition; all panels include stressor behaviour. Behaviour is shown when the network is exposed to a (i)-(ii) short ($t = 150$) or (iii)-(iv) long ($t = 150$-174) stress period. The initial time region of each regulatory component prior to stress application (grey shaded region) displays the complete stress-free behaviour. Note: Standard deviation (corresponding coloured shaded area) is included within efflux panels to represent the uncertainty, but the actual expression level cannot exceed the $[0, 1]$ domain.

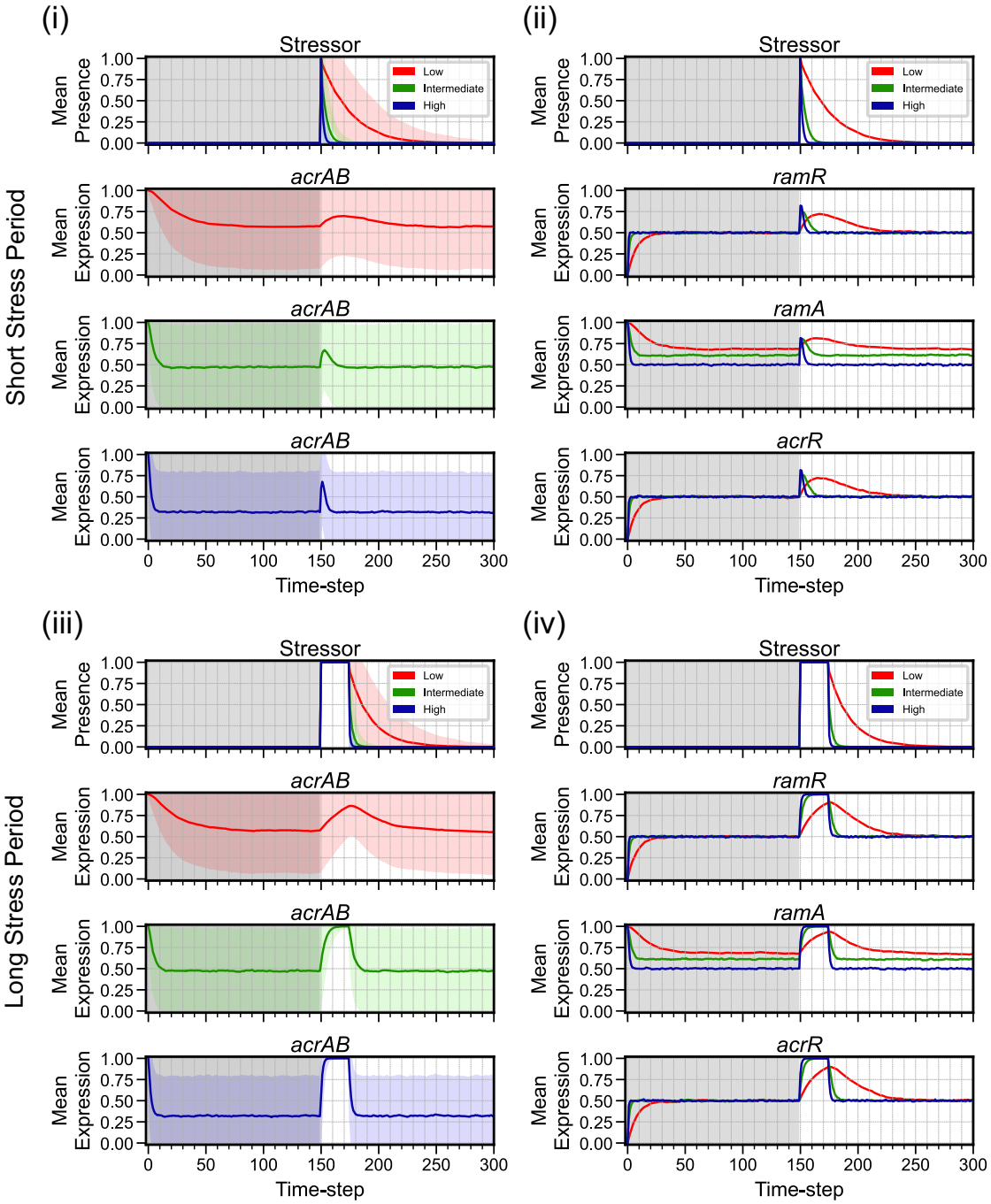


Figure 3.7: ***Salmonella* mean expression dynamics from the stressed initial condition.** Time-series dynamics of mean expression (coloured line for low (red), intermediate (green) and high (blue) energy levels) for ((i), (iii)) efflux genes *acrAB* and ((ii), (iv)) remaining regulatory components in *Salmonella*, from the stressed initial condition; all panels include stressor behaviour. Behaviour is shown when the network is exposed to a (i)-(ii) short ($t = 150$) or (iii)-(iv) long ($t = 150$-174) stress period. The initial time region of each regulatory component prior to stress application (grey shaded region) displays the complete stress-free behaviour. Note: Standard deviation (corresponding coloured shaded area) is included within efflux panels to represent the uncertainty, but the actual expression level cannot exceed the $[0, 1]$ domain.

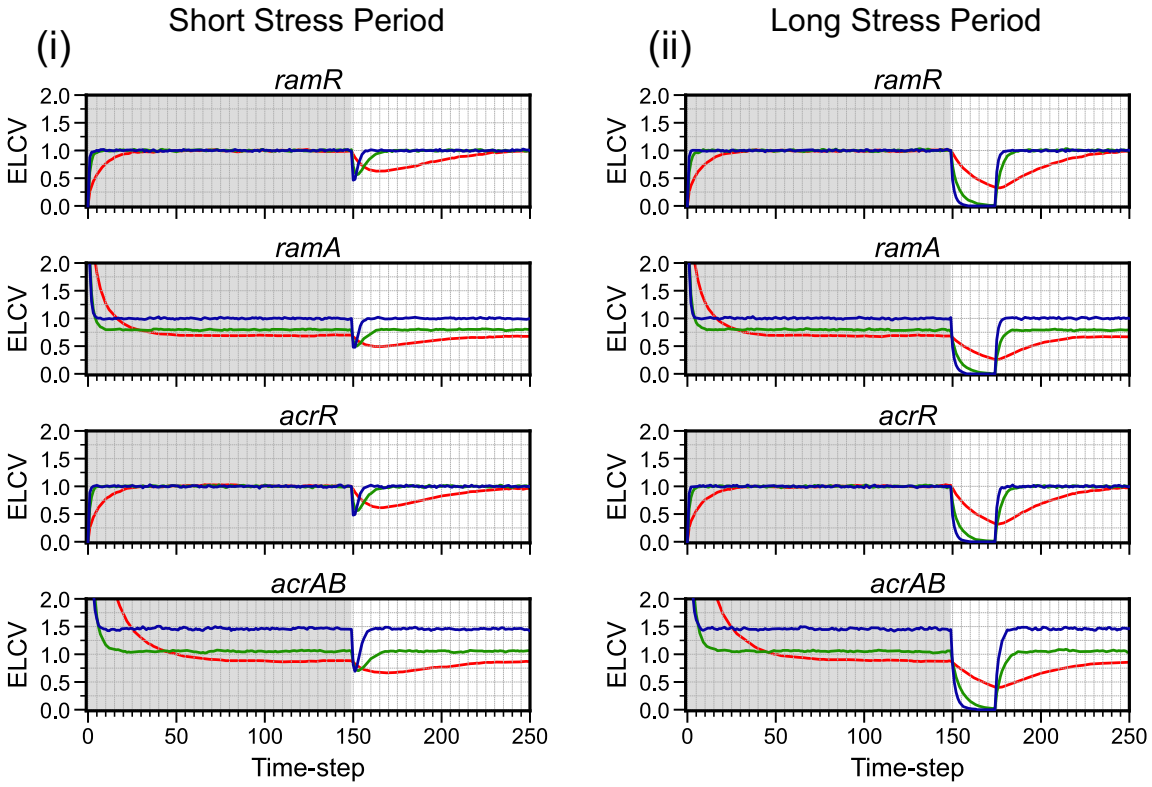


Figure 3.8: ***Salmonella* expression level coefficient of variation.** Time-series dynamics of ELCV (expression level coefficient of variation) for each regulatory component in *Salmonella*'s GRN, from the unstressed initial condition (coloured line for low (red), intermediate (green) and high (blue) energy levels). Behaviour is shown when the network is exposed to a (i) short (t = 150) or (ii) long (t = 150-174) stress period. The initial time region of each regulatory component prior to stress application (grey shaded region) displays the complete stress-free behaviour.

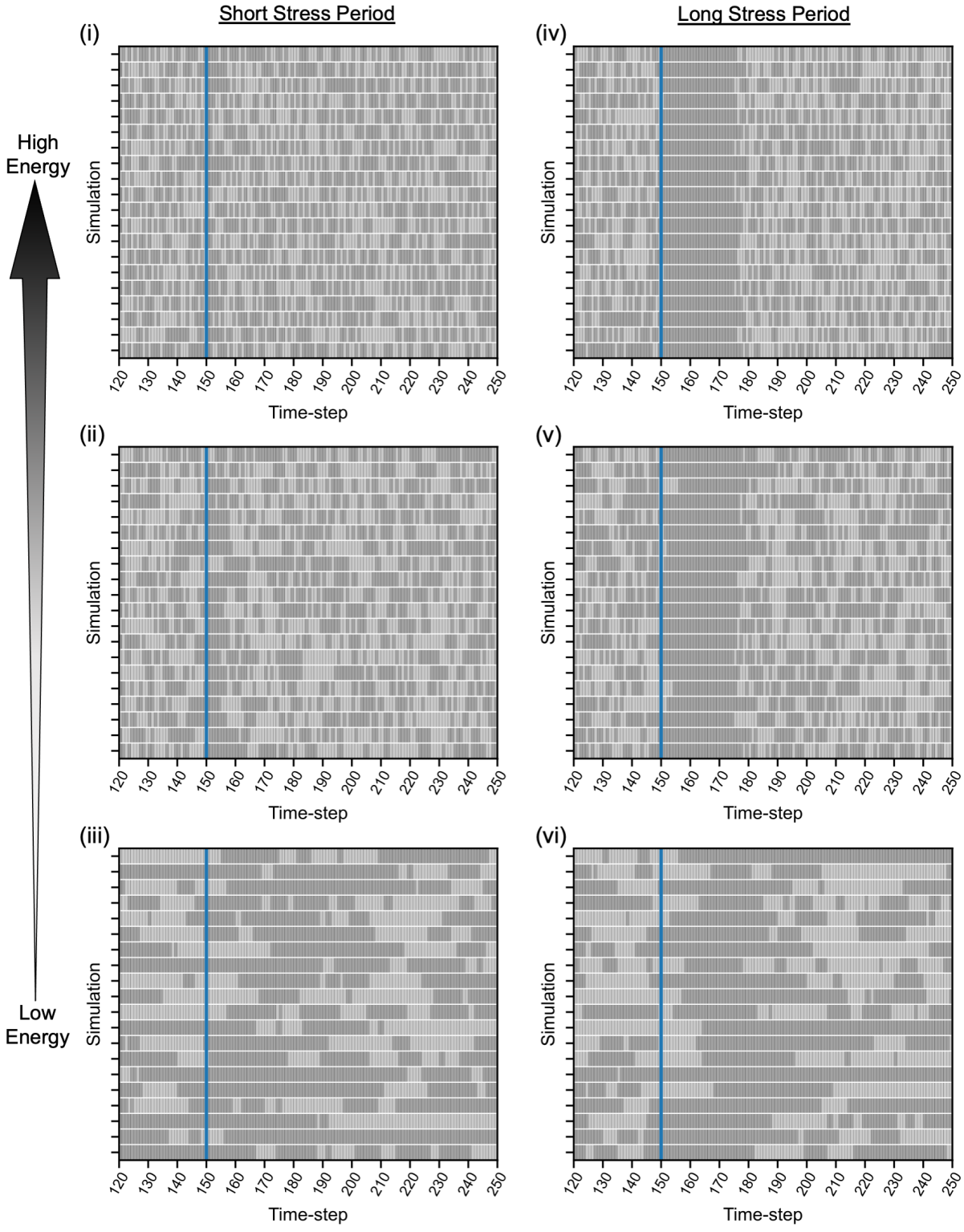


Figure 3.9: ***E. coli* displays significant heterogeneity in *marRA* gene expression in the absence of stress.** Qualitative plots displaying *marRA* "on"/"off" (dark/light grey) dynamics for 20 randomly selected individual *E. coli* simulations, when exposed to a (i)-(iii) short or (iv)-(vi) long stressor period. Panels display *marRA* dynamics between timesteps t = 120-250 at low, intermediate, and high energy availability. The short stressor period is a single timestep pulse activated at $t = 150$ (blue line) and the long stressor period is activated at $t = 150$ (blue line) until $t = 174$.

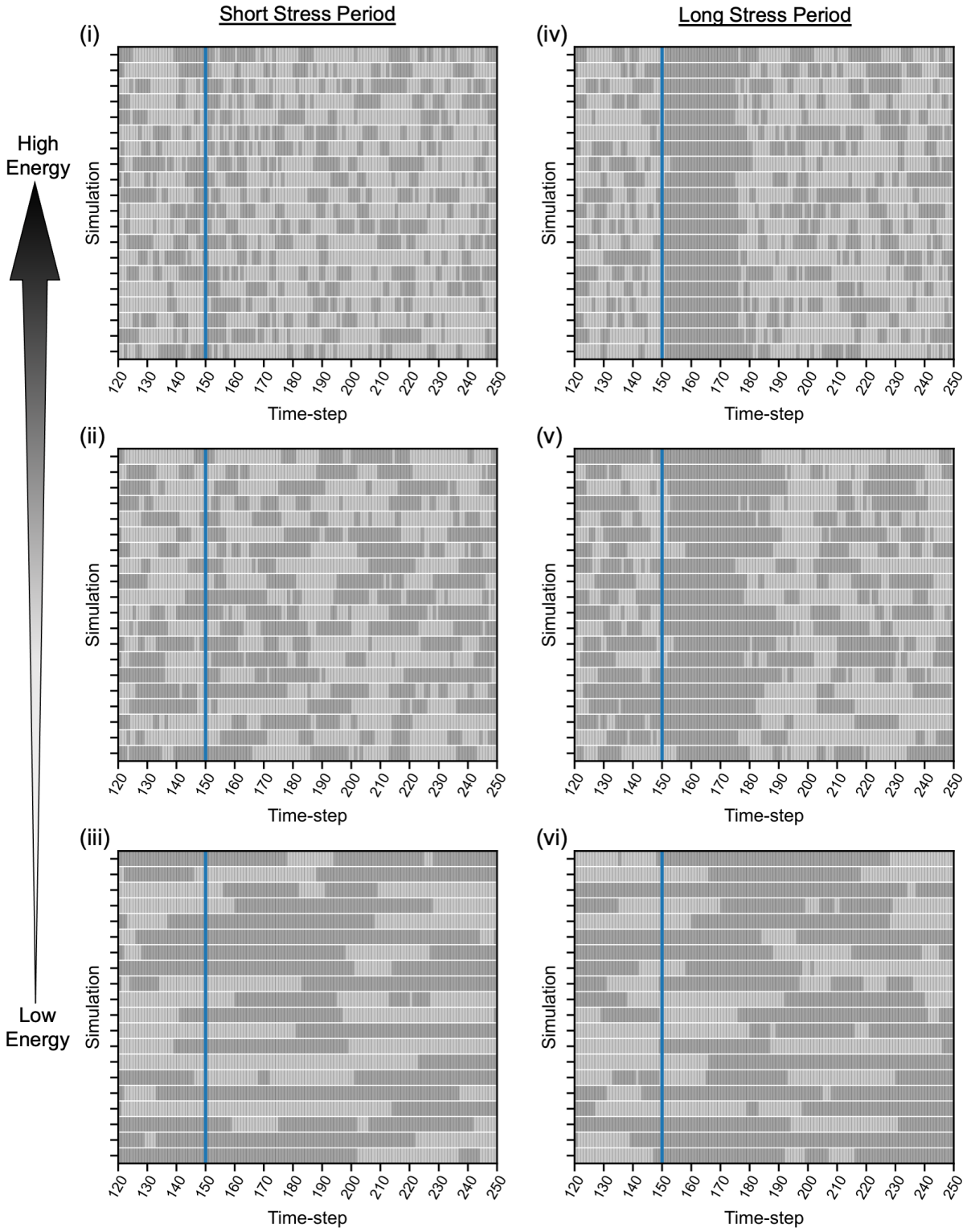


Figure 3.10: **Simulations predict *E. coli* expresses *acrAB* in heterogeneous pulses in the absence of stress.** Qualitative plots displaying *acrAB* "on"/"off" (dark/light grey) dynamics for 20 randomly selected individual *E. coli* simulations, when exposed to a (i)-(iii) short or (iv)-(vi) long stressor period. Panels display *acrAB* dynamics between timesteps $t = 120$-250 at low, intermediate, and high energy availability. The short stressor period is a single timestep pulse activated at $t = 150$ (blue line) and the long stressor period is activated at $t = 150$ (blue line) until $t = 174$.

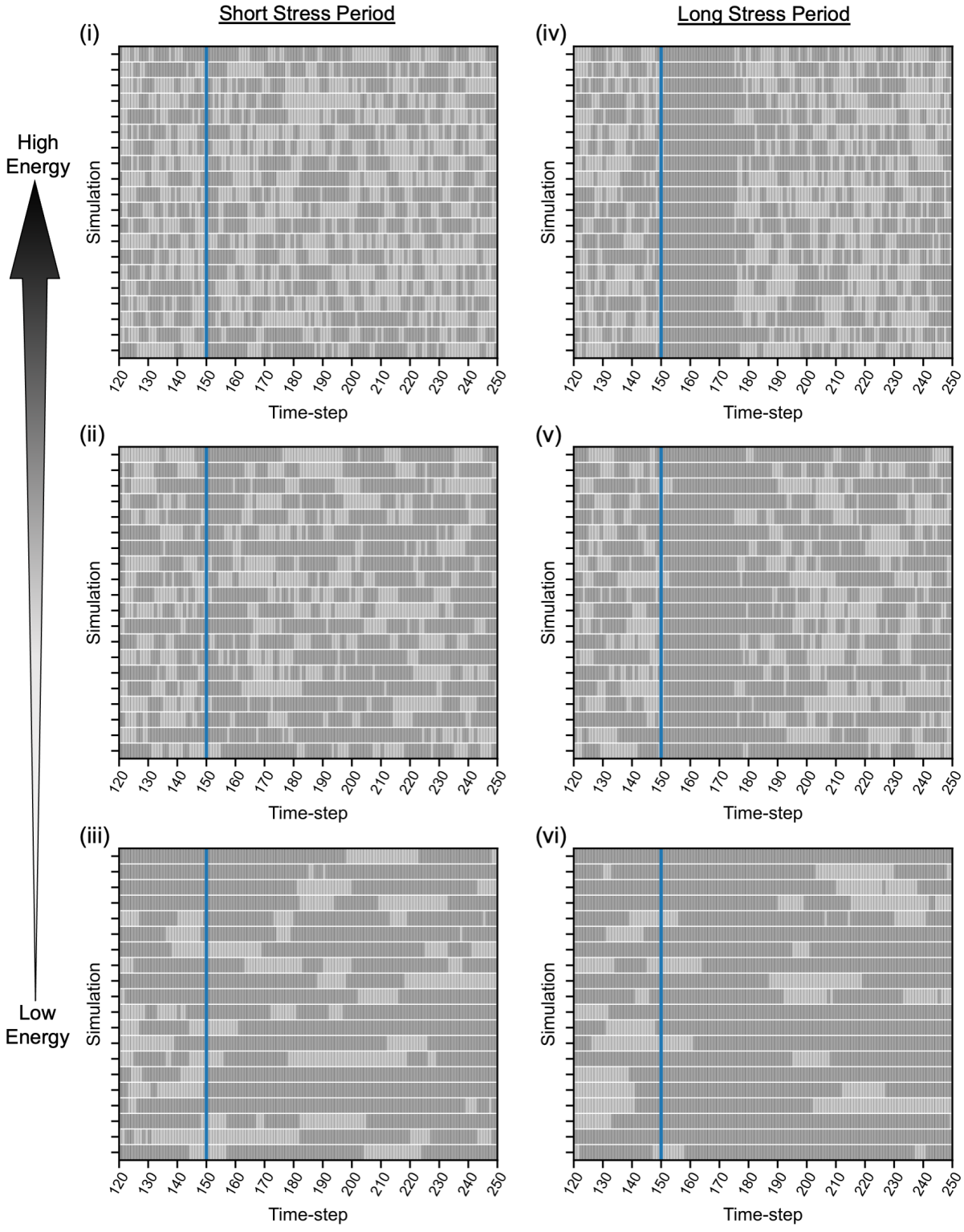


Figure 3.11: ***Salmonella* displays significant heterogeneity in *ramA* gene expression in the absence of stress.** Qualitative plots displaying *ramA* "on"/"off" (dark/light grey) dynamics for 20 randomly selected individual *Salmonella* simulations, when exposed to a (i)-(iii) short or (iv)-(vi) long stressor period. Panels display *ramA* dynamics between timesteps $t = 120$-250 at low, intermediate, and high energy availability. The short stressor period is a single timestep pulse activated at $t = 150$ (blue line) and the long stressor period is activated at $t = 150$ (blue line) until $t = 174$.

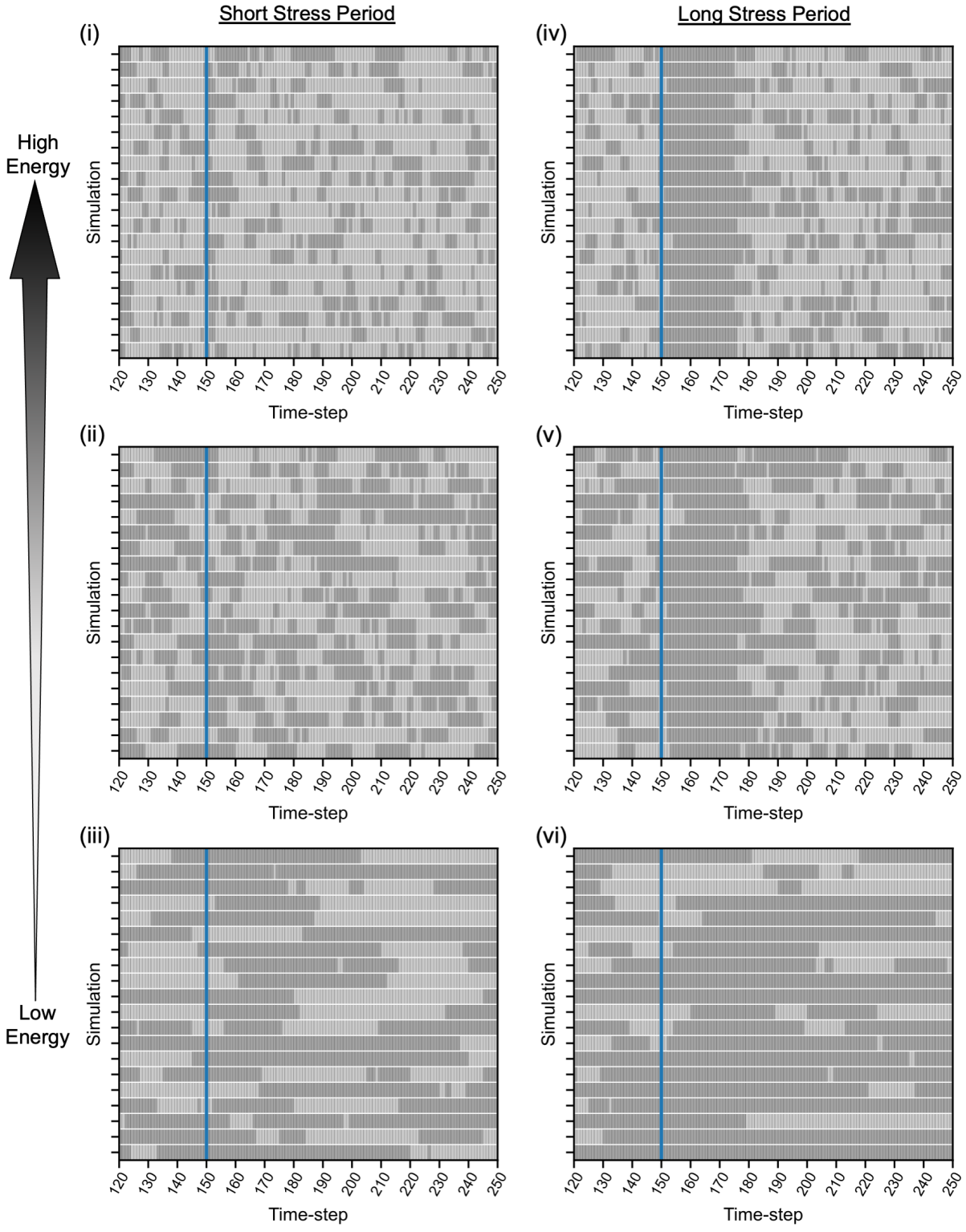


Figure 3.12: **Simulations predict *Salmonella* expresses *acrAB* in heterogeneous pulses in the absence of stress.** Qualitative plots displaying *acrAB* "on"/"off" (dark/light grey) dynamics for 20 randomly selected individual *Salmonella* simulations, when exposed to a (i)-(iii) short or (iv)-(vi) long stressor period. Panels display *acrAB* dynamics between timesteps $t = 120$-250 at low, intermediate, and high energy availability. The short stressor period is a single timestep pulse activated at $t = 150$ (blue line) and the long stressor period is activated at $t = 150$ (blue line) until $t = 174$.

3.3.3 Efflux network behaviour with a maximal cellular energy budget after exposure to a stressor

We next asked how exposure to a stressor affects network behaviour. Following a simulated stress-free period as in the previous section, we simulated exposure to a stressor for either a short (one timestep) or long (25 timesteps) stress period. In response to long stress, the mean expression levels of all genes in the system change, with different rates and magnitude (Fig. 3.3). *acrAB* mean expression rises throughout stress exposure, thus becoming higher for the longer stress duration (Fig. 3.3 (i) and (iii)). After the long stress period, the high energy population displays homogeneous behaviour (active efflux expression), increasing the capacity of each cell to remove the stressor (see Fig. 3.10 (iv) to qualitatively observe this behaviour). The high energy population can then rapidly, and collectively, expel the noxious substance, promptly restoring baseline conditions in mean expression following the steps discussed previously.

When stimulated by a short stress (Fig. 3.3 (i)), high variability is observed in efflux pump activity throughout the entire stress response, corresponding to diverse responses across the bacterial population. Individual simulation dynamics (Fig. 3.10 (i)) suggest this behaviour originates from sub-populations of cells developing unsynchronised responses to stress, and a delayed development of the efflux response. The unsynchronised dynamics prolong the population-level stress response, potentially benefiting a clonal population by providing a hedging mechanism against future stress.

In response to both stress lengths, the expression of repressive genes increases (Figs. 3.3 and 3.6), agreeing with experimental observations [162, 263]. It is predicted that the binding of the stress compound to the repressive transcription factors [7, 261, 312] removes their ability to bind to the DNA regions and down-regulate the expression of their targets (including their own expression), allowing the expression of repressive genes to increase. Once the simulated stressor is removed, repressive transcription factors can resume down-regulation of their GRN target, initially reducing the mean expression level of *marRA* and *acrR*, followed by a reduction in *acrAB* mean expression (Fig. 3.3). This

process switches off the response cascade, allowing *acrAB* mean expression to return to the pre-stress population equilibrium (Fig. 3.3).

We next asked when the population of *E. coli* cells reached an equilibrium in *acrAB* expression after each perturbation – that is, when the statistics of efflux gene expression levels no longer changed with time in the population after the stress period. We found that the number of timesteps for *acrAB* to attain a post-stress population equilibrium, in the maximal energy case, was similar after a short or long stress (9 timesteps *versus* 8 timesteps respectively, Fig. 3.3 (i) and (iii)), and the equilibrium eventually attained was the same for both stress durations. These timings differ when energy is limited, as described in the following sections.

To understand these dynamics in more depth, we next asked how the different patterns of gene expression between which a network can transition vary when a cell is exposed to stress (reflecting the diversity of behaviours supported by the network). We define 'accessible transitions' to be the set of expression patterns that can be reached directly from a given network state. To explore this, we analysed the transition matrices (see Fig. 3.1(a) and Section 3.2.7 for an illustration and description of the process to yield these outputs) describing the dynamics of the *E. coli* regulatory network in the presence and absence of a stressor, visualised through heatmap images (Fig. 3.13).

In the absence of stress, all transitions are accessible (Fig. 3.13 (i)). When a stress is applied, the accessible transitions for each network state are significantly restricted (compare Fig. 3.13 (i) to (iv)). The presence of a stressor removes the negative regulation within *E. coli*'s regulatory architecture (Fig. 3.1(b)), and consequently the ability to actively switch components "off". The stressed network contains a steady-state attractor – a single state that when reached stays constant – corresponding to all components of the network being "on", including the genes encoding repressors in the network. Under stress-free conditions no such single steady-state attractor exists (Fig. 3.13 (i)). The presence of this attractor implies that an *E. coli* cell under stress shifts its regulatory poise to allow the expression of each GRN component, allowing a response to the stressor

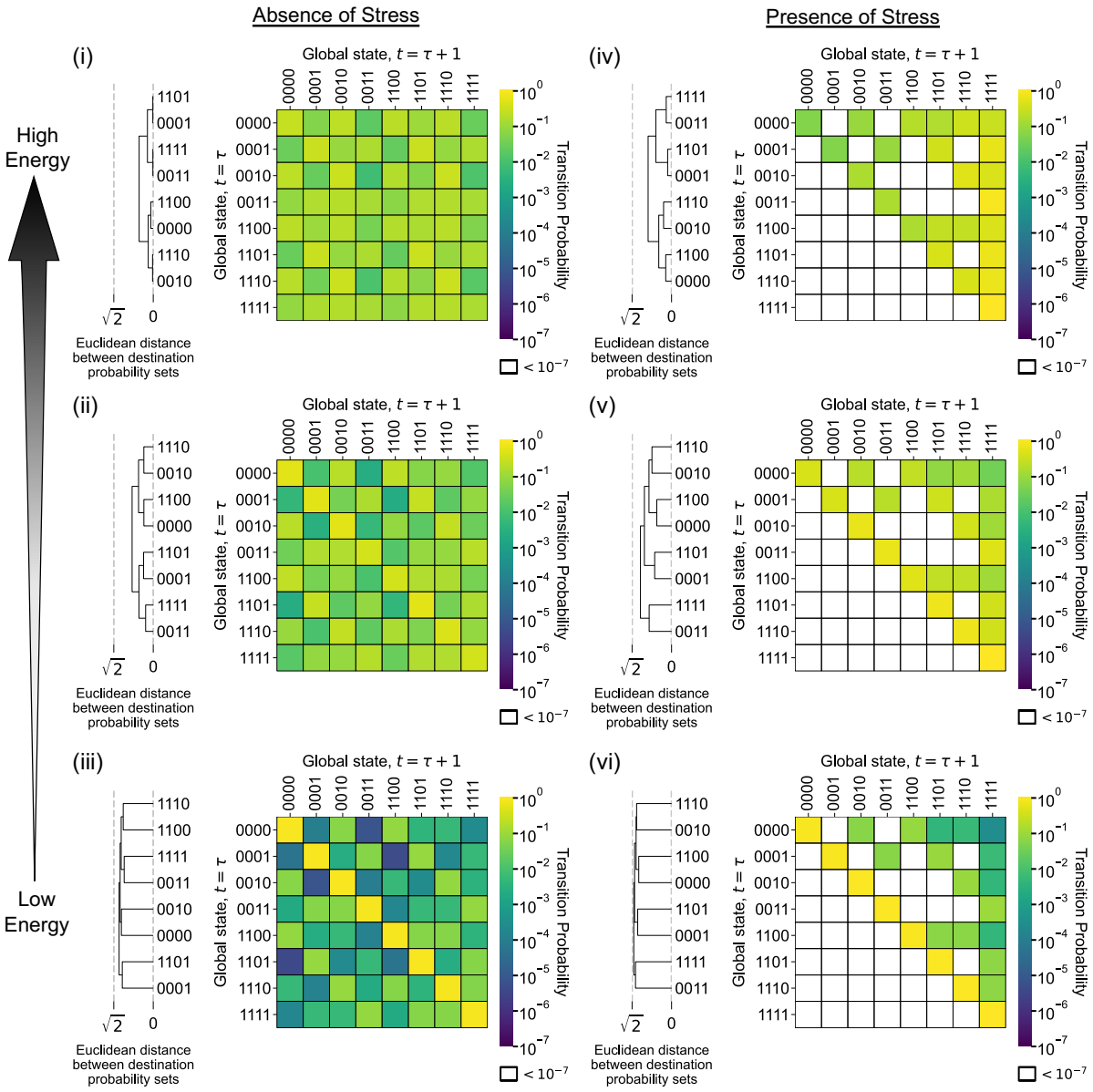


Figure 3.13: ***E. coli* transition dynamics depend on cellular energy and the presence of stress.** Heatmap plots showing *E. coli*'s transition matrix. Each row contains the transition probability distribution from each of possible global states (rows) to each subsequent state (columns), in the absence (i)-(iii) and presence (iv)-(vi) of a stressor, with increasing energy availability; here, due to the coupling of *marR* and *marA* (as they form the *marRA* operon), we do not artificially implement unattainable network states, for example, '1011'. Each matrix element is coloured to indicate the probability p (using a logarithmic scale) that global state A transitions to global state B from timestep $t = \tau$ to $t = \tau + 1$. Hierarchical clustering of global states is calculated using the destination possibilities ($t = \tau + 1$) for each $t = \tau$ global state, with clustering employing a Euclidean distance metric to determine the distances between each element. Global state binary vectors are displayed in the order [*marR*, *marA*, *acrR*, *acrAB*].

and a faster dynamic response when the stress is removed (as repressive transcription factors can immediately resume down-regulation of their targets.)

3.3.4 Efflux network behaviour is modulated by available cellular energy

We next sought to understand how energy availability changes network behaviour in *E. coli* cells, in an environment with and without a stressor. We first explored how the amount of energy available to a cell affects the accessible network transitions. Both with and without stress, the accessible transitions for each network state are the same across all energy levels, but the probabilities of these transitions vary dramatically with energy (Fig. 3.13). At lower energies, it is more likely that the system remains in the same state from one timestep to the next; increasing energy distributes the probability of each accessible transition more evenly amongst the total accessible transitions.

When a stressor is present, accessible transitions are restricted to subsets of the stress-free case at each energy level (Fig. 3.13 (i)-(iii) *versus* (iv)-(vi)). As energy increases, the likelihood of transitioning between supported network behaviours expands, transitioning more directly toward the steady-state attractor, '1111'. Therefore, increasing energy availability supports more rapid shifting of behaviour in the *E. coli* efflux network (Fig. 3.13), particularly towards a high efflux state.

Next, we asked whether the equilibrium mean expression of efflux pump genes changed at different energy levels, both in the presence and absence of stress. Counterintuitively, we found that increasing energy availability decreased the mean expression equilibrium of *acrAB* in simulated stress-free conditions (Fig. 3.3). This behaviour is a possible consequence of the less dynamic behaviour at low energy (Fig. 3.13 (i)). In the case where a simulation transitions to a state with active efflux expression, the network may become 'stuck' in this state, contributing to the increase in *acrAB* mean activation. Although low energy cells would be expected to conserve energy by dampening gene expression, an increased sub-population of active efflux cells at each time-step could provide a beneficial

bet-hedging mechanism.

In the presence of a stressor, the population response is more rapid and of higher magnitude as energy availability increases (Fig. 3.3), suggesting high energy cells develop a more superior response to stress. Simulated single-cell dynamics predict cell-to-cell variability in the time taken to transition to an active efflux expression state during stress exposure in each energy state (Fig. 3.10). The delayed response is more pronounced as energy decreases (Fig. 3.10 (iv) *versus* (vi)), indicating low energy cells may be disadvantaged in responding to stress, and potentially why the low energy population supports a higher mean expression equilibrium of *acrAB* in simulated stress-free conditions, compared to the high energy population. Following either stress period, the network returns to the pre-stress equilibrium more rapidly at higher energy levels (Fig. 3.3).

3.3.5 Energy levels influence heterogeneity in individual expression level dynamics

To investigate how energy level influences the detailed dynamics of single-cell pulses of efflux pump genes *acrAB*, we calculated the mean pulse length in *acrAB* expression at different energy levels (Fig. 3.14(i) and (iii) displays mean pulse length in *acrAB* expression for a short and long stress respectively). Intuitively, from time-scaling effects, mean pulse length increased as energy decreased (Fig. 3.14 (i)), agreeing with qualitative observations in the simulations (Fig. 3.10 (i)-(iii)). The introduction of a short stress results in only a very small transient increase in mean pulse length, which is more pronounced at higher energy (the increase is observed at each energy level in Fig. 3.14(i)). Pulse lengths during the long stress period are more dependent on energy level, with mean pulse length in high-energy cells increasing more rapidly from the baseline value compared to low-energy cells (Fig. 3.14(iii)).

We next addressed our central hypothesis, whether energy availability could be a cause of the observed cell-to-cell variability in efflux pump expression. To explore this, we calculated the dynamics of the ELCV for *E. coli acrAB* at each energy level (Fig.

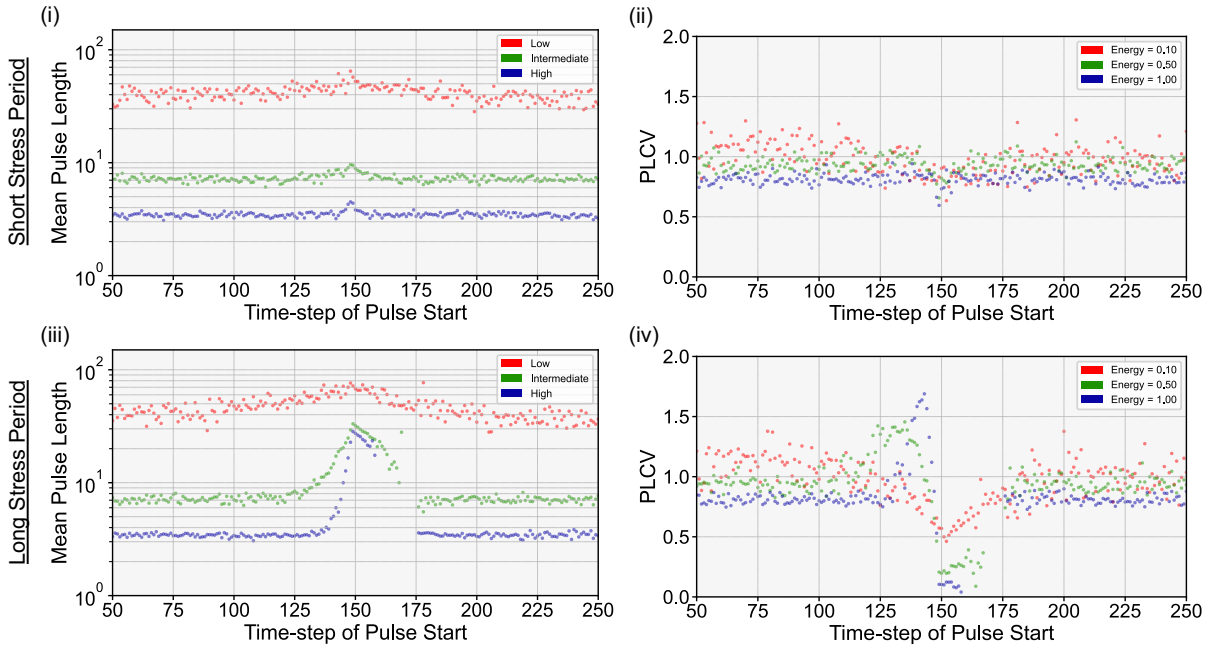


Figure 3.14: **Energy availability and stress modulate *E. coli acrAB* pulse variability.** Scatter plots displaying the (logarithmic) mean and PLCV (coefficient of variation of *acrAB* pulse lengths) starting at each timestep in *E. coli* (coloured dots for low (red), intermediate (green) and high (blue) energy levels). Statistics are displayed between timesteps t = 50-250 to observe pulse behaviour before, during and after a (i)-(ii) short or (iii)-(iv) long stress period. The short stressor period is a single timestep pulse activated at $t = 150$ and the long stressor period is activated at $t = 150$ until $t = 174$.

3.5). As seen previously, substantial cell-to-cell variability exists in the system in the absence of a stressor, reflecting asynchrony and intrinsic noise in the system (see pre-stress dynamics in Fig. 3.5 (i)). This variability is initially higher for low-energy cells, before reaching a lower value at steady state compared to high-energy cells (ELCV 1.01 at low energy, 1.34 at high energy in Fig. 3.5 (i)). Exposure to stress reduces this variability, as the population synchronises and responds in a similar way to the imposed stressor (observe the reduction in ELCV in Fig. 3.5 (i) and (ii)). Exposure to a longer stress period imposed a more substantial decrease in efflux variability than a shorter stress at all energy levels (compare Fig. 3.5 (i) and (ii)) and, as observed for the maximal energy case (Section 3.3.3), the ELCV converged to zero at intermediate to high energy (Fig. 3.5 (ii)). After the stress exposure, the original level of variability is recovered, allowing the same potential 'hedging' against further stress in the future (Fig. 3.5 (i) and (ii)) [21, 118, 148, 286].

Next, we quantitatively investigated how energy level influences the variability of efflux pump gene expression pulses, by calculating the PLCV (coefficient of variation of pulse lengths) starting from each timestep in *acrAB* (Fig. 3.14 (ii) and (iv)). A pulse is defined as a gene being "on" for one or more consecutive time-steps until it is switched "off". The PLCV is calculated for each timestep from dividing the standard deviation of the length of pulses starting from timestep i, by the mean length of pulses starting from timestep i, and corresponds to population variability in pulse lengths; a low PLCV value is interpreted as negligible variability in gene expression pulses between cells, and high values as significant cell-to-cell variability in pulse lengths. In the absence of stress, heterogeneity in pulse length is higher at lower energy (Fig. 3.14 (ii) and (iv)), whilst high energy levels display relatively low PLCV < 1, corresponding to lower variability when dynamics are faster and individual events less rare. As with expression level dynamics, when stimulated by a short stress period, PLCV decreases, more dramatically at lower energy levels (Fig. 3.14 (ii)), as the population becomes more synchronised by the imposition of an external stress. For longer stress periods (Fig. 3.14 (iv)), differences between different energy levels are more dramatic. Pulse lengths remain highly variable for the low energy population, but PLCV decreases to zero at higher energy. Thus, higher energy availability allows more concerted and synchronous population-level behaviour in dealing with the stressor. Following stress exposure, PLCV behaviour returns to its pre-stress level for each energy level.

3.3.6 *E. coli* and *Salmonella* efflux networks share functional similarity

We next compared the behaviour of the *Salmonella* network to identify similarities and differences between bacterial species. We first explored the capacity of the *Salmonella* network to transition between expression states. As with *E. coli*, the *Salmonella* network architecture supports the same accessible transitions (Fig. 3.15). Clustering by the set of accessible transitions for each global state at a given energy level shows analogous features in both bacterial species, with transitions between states limited by energy (Fig. 3.15).

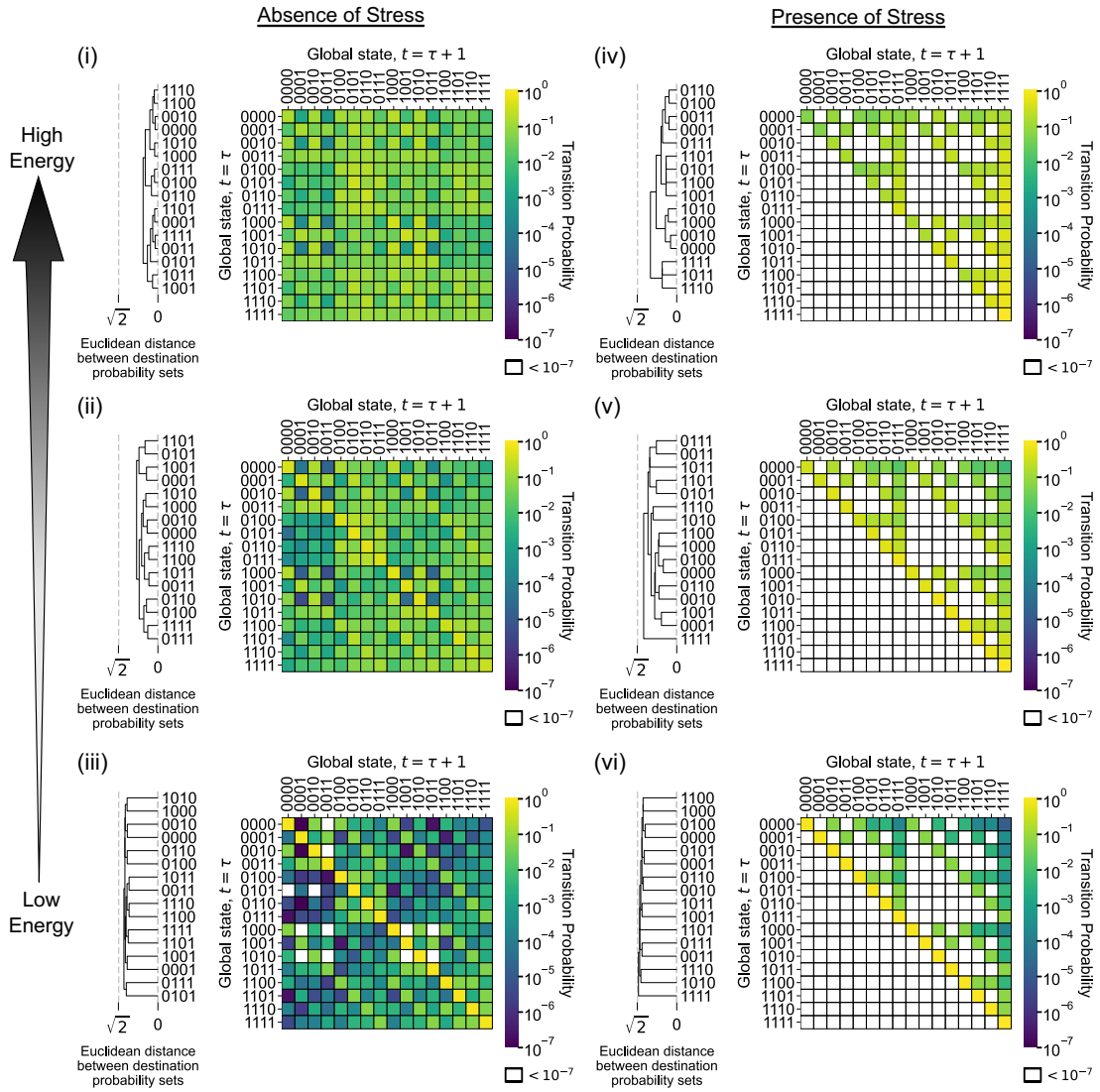


Figure 3.15: ***Salmonella* displays analogous coarse-grained qualitative behaviour compared to *E. coli*.** Heatmap plots showing *Salmonella*'s transition matrix. Each row contains the transition probability distribution for each of 2^M possible global states, in the absence (i)-(iii) and presence (iv)-(vi) of a stressor, with increasing energy availability; M here is the number of nodes in the initial GRN. Each matrix element is coloured to indicate the probability p (using a logarithmic scale) that global state A transitions to global state B from timestep $t = \tau$ to $t = \tau + 1$. Hierarchical clustering of global states is calculated using the destination possibilities ($t = \tau + 1$) for each $t = \tau$ global state, with clustering employing a Euclidean distance metric to determine the distances between each element. Global state binary vectors are displayed in the order [*ramR*, *ramA*, *acrR*, *acrAB*].

Next, we asked how *Salmonella*'s *acrAB* efflux pump genes respond to a stressor, compared to the homologous efflux pump in *E. coli*, by analysing the population-level activity of *acrAB* in both species in response to a short and long stress period (Fig. 3.3

and Fig. 3.6 for *E. coli* and *Salmonella* respectively). Simulation results predict both species respond more dynamically to stress at higher energy, regardless of stress duration and, at each energy level, rapidly return to baseline conditions after the stress period (compare Fig. 3.3 and Fig. 3.6).

Salmonella displays several other analogous behaviours compared to *E. coli*. At maximal energy availability, the number of timesteps for *Salmonella* to return to the mean expression baseline after a short and long stress period is similar (Fig. 3.6), as observed for *E. coli* (see Section 3.3.2). Other analogous behaviours include a zero *acrAB* ELCV at the maximum stress response to the long stress period (Fig. 3.8 (ii)), and the observation of heterogeneous pulses in network components (Figs. 3.11-3.12).

Simulations predict the stress-free *acrAB* ELCV is lower in *Salmonella* at low and intermediate energy, but greater at high energy, compared to *E. coli* (compare Fig. 3.5 (i) and Fig. 3.8 (i)). This behaviour is a downstream consequence of the mean and variability in global activator expression in each species, which is modulated by the energy level, and the decoupled expression of *Salmonella*'s activator *ramA* and repressor *ramR*, compared with *E. coli*'s coupled *marRA* operon (Fig. 3.1(b)-(c)).

Together, this suggests the gene regulatory networks for both bacterial species in this investigation display behavioural similarities in the presence and absence of a stressor, but despite structural similarities in their architectures (Fig. 3.1(b)-(c)) there are differences in the dynamics of efflux pump expression (Fig. 3.3 and Fig. 3.6).

3.4 Discussion

We have applied a Boolean modelling framework to coarse-grained *E. coli* and *Salmonella* efflux pump GRNs, to explore how efflux pump activity depends on the energy available to fuel the expression of efflux proteins and their regulators. We find quantitative support for our hypothesis that differences in available energy can contribute to cell heterogeneity in efflux pump expression, and dissect this heterogeneity in different mechanisms (see

	Low Energy	Intermediate Energy	High Energy
Speed of network dynamics.	●	●●	●●●
Diversity of supported network behaviour in the absence of stress	●●●	●●●	●●●
Diversity of supported network behaviour in the presence of stress	●	●	●
Cell-to-cell variability in population-level *acrAB* expression in the absence of stress.	●●●	●●●	●●●
Cell-to-cell variability in population-level *acrAB* expression in the presence of stress.	●●	●	●
Cell-to-cell variability in pulses of *acrAB* expression in the absence of stress.	●●●	●●	●
Cell-to-cell variability in pulses of *acrAB* expression in the presence of stress.	●●●	●	●

Table 3.4: **Summary of energy availability effects on regulatory network behaviour.** The table displays a summary of our main results from applying the modified Boolean network framework to *E. coli* and *Salmonella acrAB* GRNs. The strength of the effect of low, intermediate, and high energy availability on each key result is presented through a discretised scale: weak (●), medium (●●) and strong (●●●).

Table 3.4 for a summary).

In the absence of differences in available energy, our model suggests (supported by experiments) that substantial intrinsic variability exists in efflux gene expression. The imposition of a population-wide stressor reduces this intrinsic variability by partially synchronising cell behaviour; stochastic effects after stress exposure recover the original variability. This intrinsic noise can potentially be acting to hedge the population against future stress, while maintaining cells that invest less resource in defence mechanisms and/or protect the lineage from diverse environmental conditions. When extrinsic differences in ATP availability between cells are additionally imposed, the population variability in efflux response is amplified in the simulations. Cells with lower energy levels experience more variable expression of efflux components (amplifying intrinsic variability), slower responses to stressors, and hence slower removal of stressors from the environment. In future, experimental tests of this hypothesis could involve, for example, jointly tracking bacterial ATP levels (using the assay in Section 2.4) and efflux gene expression in a culture prior to, during and post antibiotic stress using fluorescent reporters [30, 73, 213, 309].

In our model, stress-free bacteria naturally express heterogeneous pulses of efflux genes (Fig. 3.10 and Fig. 3.12). This behaviour may be biologically valuable for several reasons. The products of *acrAB* have an extremely long half-life in *E. coli* [44], suggesting that the genes may only need to be transcribed in short pulses (which is predicted to be a less energetically demanding process compared to translation [161]) for the protein to be in an abundant supply. Dealing with external stressors is one role of the AcrAB-TolC efflux pump, but it also, for example, removes endogenous waste products [35], suggesting the pulse behaviour could aid physiological functions by expressing *acrAB* mRNA to be used in these processes. The pulse behaviour may also confer a natural resistance phenotype to a fraction of cells in stress-free environmental conditions at all times, as reported experimentally in *Salmonella enterica* [247]. Therefore, the predicted pulsing behaviour may contribute to several essential intracellular functions.

Recent mathematical models have studied the *mar* network at the protein level, suggesting the motif is monostable, and supports a single deterministic steady-state in the absence of stress [87, 240]. A theoretical study has also shown, at the protein level, the development of stochastic pulses in *marRA* products [87]. The cell-to-cell heterogeneity in *mar* expression that we predict agrees with ODE models of the *mar* operon – which account for the inherent stochasticity of biological systems – and experimental demonstrations of heterogeneous cell-to-cell steady-state *mar* expression within a population, in the presence and absence of stress [74, 87, 241] (Figs. 3.3-3.9). With the development of advanced experimental assays, future studies may resolve the behaviour at the transcript level to test our theoretical predictions.

Even for these relatively simple networks, our Boolean modelling approach has demonstrated general principles of energy influence without requiring characterization of a large parameter space. Developed with a view to scalability, this approach can be applied to many naturally-occurring decision-making phenomena. Our work in Chapter 2 demonstrated the profound influence that energy variability can have on simple decision-making circuits. This GRN approach allows the expansion of this philosophy to a much

wider range of architectures. In the broader regulatory context of these core networks (Fig. 3.1(b)-(c)), this could include, for example, the inclusion of *soxRS*, *rob* and *acrZ* within our GRNs, which are known to be integrated into regulating *acrAB* expression [13, 104, 171, 178], providing additional layers of control; the current understanding is that *soxRS*, *rob*, and *acrZ* play a minor role compared to the *mar*, *ram* and *acrR* regulators [2, 29, 104, 195, 224, 302]. As cell-to-cell variability in ATP is increasingly elucidated [63, 270, 313] we believe that our approach will help reveal the theoretical underpinnings of this important influence on regulatory dynamics.

It is important to note some limitations of our modelling approach. First, the coarse-grained nature of a Boolean model limits the precise quantitative agreement of individual observations with either more detailed models or biological reality. Instead we focus on the relative magnitudes and directions of behaviour – for example, does X increase or decrease Y, and does Z do so more. Second, our models are deliberately simple. Model refinements, including allowing heterogeneity in the timescales [47] and ATP dependencies of different genes, could be invoked to allow a more detailed connection with biology, and doing so as more data becomes available is an important target for future work. Third, our models are not guaranteed to be unique. The edge weights in our models were manually chosen for compatibility with our training observations. As they encode logical interactions rather than physical quantities, there are ranges of values around our choices that will give networks that behave identically. The existence of non-overlapping sets of parameter values that would also describe and predict observations has not been addressed here. Parameter inference approaches including approximate Bayesian computation [117, 278] could be used in future work to explore parameter space more fully.

Our hypothesised relationship between energy availability and efflux-pump activity, variability, and capacity to respond to antimicrobials has, as far as we are aware, not been previously considered. However, our hypothesis is not restricted to this AMR mechanism or motif. A connected phenomenon that may be mechanistically related is the formation of bacterial persisters in the presence of antibiotics [21, 31, 67]. Here, phenotypic variants

display a transient tolerance to external stress [135, 149], causing a leading factor in the recalcitrance of chronic infections [80, 136, 182, 317]. Increasing evidence has reported a mechanism linking persister cells and intracellular energy, or more specifically ATP [56, 153, 184, 251, 315]. One study has gone further, suggesting there may be a general low energy mechanism of persister formation in bacteria [166]. This relationship bears substantial similarities to our model where we predict high variability and noise in gene expression pulses in low energy cells, compared to intermediate and high energy cells. High expression variability contributes to cell-to-cell variation within a clonal population, implying a different set of phenotypes may be expressed in the low energy sub-population. As energy, or more precisely ATP, increases, cell-to-cell variability decreases, implying a uniform phenotype displayed across the population. We would predict, in agreement with growing evidence, intracellular ATP is a factor in the mechanisms of persister generation, and future work will explore this using our approach when the governing GRNs can be identified.

CHAPTER 4

IDENTIFYING EVOLUTIONARY PATHWAYS OF ANTIBIOTIC RESISTANCE IN *P. AERUGINOSA* USING BAYESIAN MODELLING

Identifying the antimicrobial resistance profile of an isolate taken from a patient's infection is key to managing their treatment and recovery. Based on the current resistance profile of a sample, it is difficult to predict whether the bacteria causing the infection will acquire resistance to other antimicrobials (reducing the available treatments for a patient). In this chapter, we apply the HyperTraPS algorithm [121], a statistical framework combining stochastic modelling with Bayesian inference, to phenotypic susceptibility data of clinical *P. aeruginosa* isolates. This framework is used to infer the dynamic probabilistic pathways of antibiotic resistance acquisition in the *P. aeruginosa* isolates. In general, this could allow more optimal treatment strategies based on current real-time data.

4.1 Introduction

P. aeruginosa is a bacterial pathogen associated with a broad spectrum of infections in humans and has been ranked as a critical priority for the research and development of new and effective antibiotic treatments by the WHO. This is due to its intrinsic resistance to several classes of antibiotics, and the ability to acquire resistance to last-line therapeutic options [267]. Strategies to prevent *P. aeruginosa* infection and reduce the likelihood

of antimicrobial resistance should be prioritised when treating patients vulnerable to, or infected with, this microorganism.

The current platform of genome sequencing and phenotypic susceptibility testing across microbiology has the ability to monitor the epidemiology of infectious diseases, and enable earlier and more accurate diagnoses [48, 65, 76]. Computational tools can be applied to genome sequencing and phenotypic susceptibility data to complement current diagnostic methodologies, providing predictive tools that suggest appropriate antibiotics based on current isolate data. This could enable advanced, real-time treatment for patients. Advancing diagnostic frameworks, that involve both computational and experimental developments, could be a powerful tool in combatting antimicrobial resistant infections.

In relation to this, the recently developed HyperTraPS (hypercubic transition path sampling) algorithm aims to quantitatively resolve evolutionary questions in different organisms, regarding the acquisition of numerous discrete, coupled features [95, 121, 307]. HyperTraPS, a statistical framework combining stochastic modelling with Bayesian inference, takes a set of related observations displaying a set of features, and infers the potential ordering in which individual features were most likely acquired. Here, we apply the HyperTraPS algorithm to phenotypic susceptibility data of clinical *P. aeruginosa* isolates to infer the dynamic probabilistic pathways of antibiotic resistance acquisition in these cells. We identify general trends of early and late acquisition of resistance to each antibiotic, which can be used to predict future behaviours, such as the order in which resistance may arise to individual antibiotics within a *P. aeruginosa* isolate from a patient's sample. This data could aid in suggesting the best paths of treatment based on current real-time data from a patient's sample, allowing more optimal treatment strategies that are quick and cost-effective to get results.

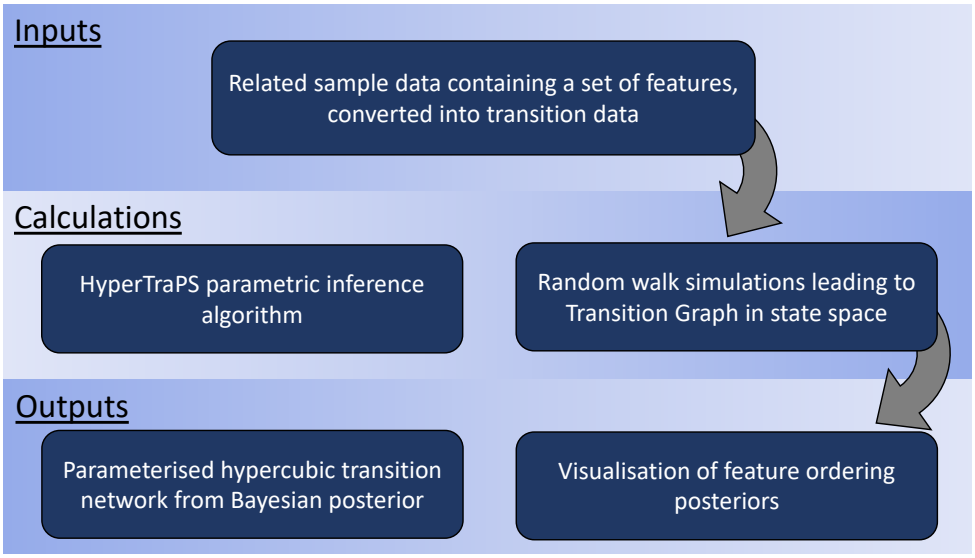


Figure 4.1: **Coarse-grained illustration of the HyperTraPS pipeline.** Schematic illustration of the HyperTraPS framework from inputs to outputs. The input for HyperTraPS is a set of transition data, produced from related sample data containing a set of L features. These inputs feed into the HyperTraPS parametric inference algorithm and random walk simulations, returning a parameterised hypercubic transition network from the Bayesian posterior. The inferred feature ordering posteriors can then be visualised.

4.2 Methods

4.2.1 Methodology overview

HyperTraPS describes the features, or traits, of each state of a system by a binary string of length L (total features), where the ith position corresponds to the absence (0) or presence (1) of the ith trait. A system can be in any of the 2^L possible states, and our interest is in the transitions between each of these states. Therefore, we consider a set of related data (phylogenetically coupled), D, which can be represented as a set of transitions $D = \{s_i, t_i\}$, where s_i, t_i are the ith start and target states, and there are n_D observations in total (see Fig. 4.1 for illustration of the HyperTraPS pipeline). The different states, corresponding to all possible combinations of the presence (1) and absence (0) of L features, can be represented as a hypercubic space, connecting all possible patterns of feature presence and absence, with neighbouring nodes differing by a single feature.

On the hypercubic space we have a transition network π, describing the probability of a transition between two states. The set of transition probabilities in π leading away from any node are, by definition, constrained to the sum of 1. We allow evolutionary transitions between nodes on the hypercubic space, so that traits are acquired one by one, and each transition between nodes has a probability which is determined from the data, D. Evolution is modelled as a walk on the hypercubic state space starting at the state of all 0's, and potentially ending at the state of all 1's, with individual transitions along the walk occurring randomly, weighted by transition probability.

We will use a Bayesian inference framework for inferring parameters (transition probabilities) for the set of edge weights π on the hypercubic transition graph that explain data D. Initially, we assign an uninformative uniform prior over the outgoing edge weights from any given node, interpreted as all possible options leaving a node as being equally likely *a priori*.

Next, we apply the core algorithm of HyperTraPS to efficiently estimate the likelihood, $\mathcal{L}(\pi|D)$. We initially note that we work in the picture of a hidden Markov model [231]. Trajectories emit 'signals' at random with a characteristic rate, allowing observation of a sequence of states visited in the trajectory. Of interest is the probability that randomly emitted signal(s) from a randomly sampled evolutionary trajectory matches an observation [121]. The product of this probability over all n_D independent observations constitutes the likelihood $\mathcal{L}(\pi|D)$ of a given transition network π given data D.

A set of random walkers are simulated along the hypercubic space and we write $P_{emission}$ to denote the probability of emitting signals at the required times along a walk to allow the observation of a transition between the initial and final state. The probability of observing a transition matching an observation is generally proportional to the product of the transition probability between start and target states, and the probability of reaching the initial state ($P(t_i|s_i,\pi) = P_{reach}(s_i|\pi)$). In general we have

$$P_{observe}(s_i \to t_i) = P_{emission}(s_i,t_i)P_{reach}(s_i|\pi)P(t_i|s_i,\pi). \tag{4.2.1}$$

As signals are emitted uniformly with a characteristic rate, and independently of π, $P_{emission}$ is a constant multiplicative factor, leaving $P_{observe}(s_i \to t_i) = P(t_i|s_i, \pi)P_{reach}(s_i|\pi)$. It is shown by Johnston and Williams [121] that the term $P_{reach}(s_i|\pi)$ is unity both for convergent (distantly related organisms independently evolve similar traits) and non-convergent evolution (taxonomically related relationships). In convergent evolution, the initial state is always the origin state O where all trajectories start, and $P_{reach}(O|\pi) = 1$. In non-convergent evolution, if s_i is the end state of one observation, then in the overall likelihood product we can take $P_{reach}(s_i|\pi)$ to be 1. This leaves $P_{observe}(s_i \to t_i) = P(t_i|s_i, \pi)$ in both evolutionary pictures. The likelihood can then be decomposed into the form in Equation 4.2.2, which is the product of the probabilities of observed transitions to t_i, from s_i, for a given parameterization of π, and returns a single estimated likelihood for a given parameterization and set of observations. For computational convenience, log-likelihood $l = \log \mathcal{L}$ is used.

$$\mathcal{L}(\pi|D) = \prod_{i=1}^{n_D} P_{observe}(s_i \to t_i). \tag{4.2.2}$$

To compute the probabilities associated with all trajectories that lead from s_i to t_i, we need to sample the set of paths that start at the first state and finish at the second, preferentially sampling the paths that are most likely. This calculation becomes computationally expensive for systems of dimensions approximately $L \gtrsim 10$ [95]. Instead, this problem is tackled through path sampling, by performing biased random walks restricted to pathways that end in t_i, introduced by Johnston and Williams [121]. In brief, a trajectory is initialised, and the probability of making a move to a state consisting of features displayed in state t_i (called t-compatible) is noted. This step is repeatedly performed until t_i is reached, with the product of probabilities computed, and the amount of bias required to keep in t-compatible states recorded. After performing numerous trajectories, the average probability product is computed, and the aggregated amount of bias controlled for. This sampling method is an efficient framework to accurately estimate the likelihood, and

allows systems with more features to be considered than previously has been the case.

After estimating the likelihood, it is then embedded within a Bayesian inference framework to infer the set of edge weights π on the hypercubic transition graph that explain the data D. We are therefore concerned with obtaining posterior distributions on π, using Bayes' theorem (Equation 4.2.3),

$$P(\pi|D) = \frac{P(D|\pi)}{\int P(D|\pi)P(\pi)d\pi}P(\pi). \tag{4.2.3}$$

As we choose a uniform prior, $P(\pi)$, we only need to consider the calculation of $\mathcal{L}(\pi|D)$ to derive samples from the posterior probability distribution.

To sample parameterizations that are associated with high likelihoods, the HyperTraPS framework embeds its likelihood estimation algorithm within a Markov chain Monte Carlo (MCMC) sampler [300], sequentially drawing new parameter estimates π' (called the trial transition matrix), after applying a Gaussian-distributed perturbation to each element of the current transition matrix π. The standard deviation of the perturbation is dependent on the kernel value used in simulation inputs. If the trial transition matrix π' is outside the prior, it is automatically rejected and a different trial matrix is considered. Otherwise, the likelihood is calculated for trial matrix π' and accepted based on comparisons to the likelihood of the current transition matrix π, using the Metropolis–Hastings algorithm [41]. This inference process yields inferred posterior distributions on the hypercubic edge weights π (displayed within "Calculations" in Fig. 4.1).

Although recently developed, this framework has been successfully applied to diverse real-world problems, including the prediction of clinical risk of children with severe malaria and the evolution of multi-drug resistance in Tuberculosis, displaying its capabilities [95, 120].

Antibiotic Class	Antibiotics
Fluoroquinolones	Gatifloxacin, Levofloxacin, Ciprofloxacin.
Nitrofurans	Nitrofurantoin.
Aminoglycosides	Amikacin, Tobramycin, Gentamicin.
Antifolates	Trimethoprim.
Cephalosporins	Ceftazidime, Cefotetan, Cefazolin, Cefepime, Cefuroxime, Cefotaxime, Ceftriaxone.
Carbapenems	Meropenem, Imipenem.
Penicillins	Amoxicillin, Ampicillin, Piperacillin/Tazobactam, Ticarcillin/Clavulanic acid.

Table 4.1: **Antibiotics grouped by antibiotic class.** Each row in the table states an antibiotic class and the group of antibiotics from the 21 tested compounds that are within it; based on information from [133]. Note: 1) Cephalosporins, carbapenems and penicillins are all part of the β-lactam family, and antibiotics within the same class can be from different generations of antibiotic development; 2) Tazobactam and clavulanic acid are both β-lactamase inhibitors.

4.2.2 Application of Bayesian inference framework to *P. aeruginosa* data

We applied HyperTraPS and Bayesian inference methodologies to a dataset consisting of antibiotic resistance profiles for *P. aeruginosa* isolates causing nosocomial infections in patients located at public hospitals of northeastern Mexico [93]. The data consists of 92 isolates tested against 21 antibiotic treatments, from several classes, including cephalosporins, aminoglycosides and carbapenems (see Table 4.1 and Fig. 4.2). To produce this data, González-Olvera *et al.* [93] obtained antibiotic resistance profiles for each isolate using the automated MicroScan autoSCAN-4 System. For each antibiotic a minimum inhibitory concentration (MIC) was determined, relating to the resistance status of each isolate (sensitive, resistant, or indeterminate). In addition, the presence of β-lactamases in response to each antibiotic was evaluated. Out of the 92 isolates, 26 (28.6%) were found to be multidrug resistant (MDR), and 21 (22.83%) were classified as extremely drug resistant (XDR); categorization of *P. aeruginosa* isolates as MDR and XDR is as defined in [164].

An initial problem within the dataset was that isolates were categorized as producing β-lactamases in response to an antibiotic (some of which are from the β-lactam family)

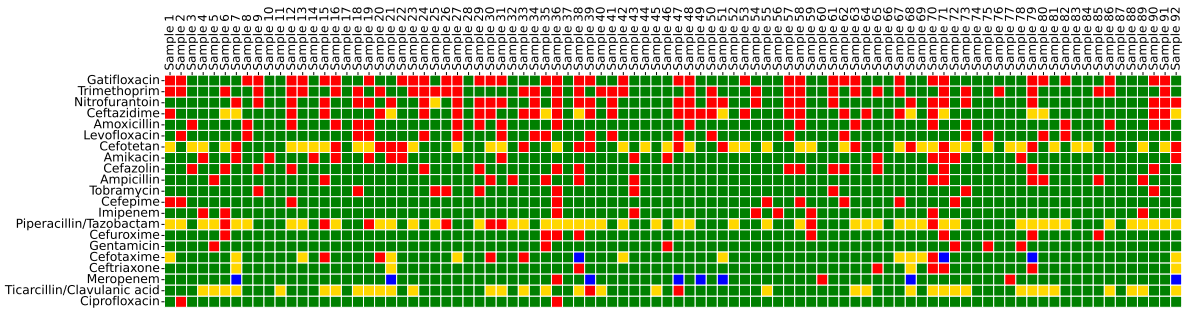


Figure 4.2: **Antibiotic resistance profiles of 92 *P. aeruginosa* isolates.** Replica dataset displaying antibiotic resistance profiles of 92 clinical isolates isolated from nosocomial infections in various anatomical sites, such as urine, respiratory and auditive tract, wounds (from different parts of the body), in addition to bloodstream samples. The squares below each isolate represent the resistance pattern of each isolate to different antibiotics, indicating sensitive (green), induction of beta-lactamase (yellow), resistance (red) and indeterminate data (dark blue). Replicated from the original version in [93] which includes phylogenetic relationships of all 92 *P. aeruginosa* isolates.

without being defined as displaying a resistant phenotype, even though β-lactamases are known to be a sophisticated, clinically-relevant resistance mechanism [233]. To investigate any underlying relationships within the data that may be missed by categorising an isolate as susceptible and β-lactamases producing, we have taken two interpretations of the data and will perform the analysis separately on both datasets. In the first dataset, we define traits as a resistance or induction of β-lactamase phenotype (red and yellow squares in Fig. 4.2), and the second defines traits as a resistant phenotype only (red squares in Fig. 4.2); the two datasets are named the "Test Dataset" and "Resistance Dataset" respectively, conveying that one dataset is testing an alternative resistance phenotype definition compared to that by González-Olvera *et al.* [93].

We describe the antibiotic resistance profile of each isolate by a binary string of length $L = 21$, where the ith position corresponds to the absence (0) or presence (1) of the ith trait, which here is being susceptible (0) or resistant (1) to antibiotic i. The full set of all binary strings of length $L = 21$ can be represented as a hypercubic graph, with each node representing a possible state.

It is typically perceived that antimicrobial-resistant bacteria might be less fit, or the AMR machinery imposes a fitness cost on the host, leading to the loss of the resistance

mechanism or resistant bacterial population, but it has been shown this is often not the case [20, 60, 216]. In the case where resistant bacteria might be less fit, compensatory mutations can restore fitness to wild-type levels, enabling resistant clinical isolates to proliferate and spread [167]. We therefore proceed by assuming the acquisition of resistance in the dataset is not lost once gained. This assumption removes the possibility of continuous losses and gains of a trait, equating to eliminating the prospect of walks containing loops on the hypercube, which would otherwise be computationally exhaustive. Proceeding with this assumption enables the application of HyperTraPS to the *P. aeruginosa* data.

With this assumption, we work backwards through the phylogeny calculated by González-Olvera *et al.* [93] and infer, using maximum parsimony, the antibiotic resistance profile of every unobserved parent state. Briefly, this includes comparing the resistance profiles of the descendants from each internal node within the phylogenetic tree; we performed this computationally using Python 3.9.4. From these estimates, we can reconstruct the transitions from parent nodes to descendent nodes on the phylogeny (see [93]) using the phylogenetic relationship to establish the resistance profile of a parent node which can evolve to any of its connecting descendants (using an OR function in Python 3.9.4 computations). These transitions then form the observations used by the HyperTraPS platform, which is embedded within a Bayesian inference framework to sample from the posterior distribution for the parameters within our model.

For each of the datasets in our results, 2000 HyperTraPS trajectories were used to estimate likelihoods. 4×10^5 MCMC iterations are performed with a burn-in period of 2×10^4 iterations, and posterior samples taken every 10^3 iterations. To determine an appropriate value of the perturbation kernel for Gaussian-distributed perturbations on transition matrix π, and that there was good chain mixing and convergence, we compared the results from several different simulation protocols, involving different random number seeds and step sizes (by modulating the kernel value to yield an acceptance rate between 20%-40%), in our preliminary investigation. We found for both datasets, a perturbation

with standard deviation of 0.05 on each element in the current transition matrix π when obtaining new parameters, produced good mixing (reasonable for preliminary investigations) and a good acceptance rate of transitions. The agreement between the posteriors in different random number seeds cases then confirmed the convergence of the algorithm. Information from the inferred posterior distributions are then visually summarised by recording the probability with which a given feature is gained/lost at a given ordering on an evolutionary pathway.

4.3 Posterior orderings for the Test and Resistance Datasets

In Figs. 4.3-4.4 we show the inferred orderings of antibiotic resistance acquisition for the Test and Resistance Dataset respectively. Traits for each isolate are defined as displaying β-lactamase activity or a resistance phenotype in the Test Dataset, and displaying a resistance phenotype in the Resistance Dataset, when tested against each antibiotic; we will call these the "resistance" traits of each dataset (see Fig. 4.1 for illustration of the process to yield these outputs). Inferred orderings of the Test and Resistance Datasets are simulated with different random numbers, with each random number being an integer initialising the random number generator within the MCMC algorithm, and different values are run to check for convergence (outputs display confluent behaviour). The convergent and contrasting behaviour will be discussed.

The inferred ordering distributions of the Test Dataset display some confluent behaviour when simulated with different random number seeds (Fig. 4.3), however, there is a lack of convergence in the outputs. The lack of convergence means that there is strong uncertainty in the produced results for the majority of antibiotics within the data. To improve this convergence, simulations of the Test Dataset need to be extended, but this preliminary stage of the project is time-limited. The features that are conserved across simulations are for a small subset of the antibiotics, including cefotetan, gentamicin, and

meropenem. It is predicted that resistance to cefotetan is acquired earlier than other drugs tested, and resistance to gentamicin and meropenem is acquired later (Fig. 4.3). Gentamicin also displays minor bimodality in the inferred ordering, suggesting resistance to this drug might be acquired at different stages, but with a skew towards later acquisition.

Next, we investigated the inferred ordering distributions of individual antibiotics for the Resistance Dataset across different random number seeds (Fig. 4.4). In this dataset, resistance traits are defined as an isolate phenotypically displaying resistance to any of the 21 antibiotics in susceptibility testing. It is initially observed the general structure and variability in ordering distributions is more confluent between different random number seed simulations, suggesting better convergence for the Resistance Dataset and giving more confidence in the posterior orderings. There is, however, a degree of difference within the posterior distributions for piperacillin/tazobactam and meropenem still present.

The Resistance Dataset displays general trends of early, intermediate, and late acquisition of resistance to each antibiotic. Gatifloxacin, trimethoprim and nitrofurantoin resistance is predicted to be acquired earlier than other drugs tested, and resistance to drugs such as cefuroxime, and meropenem is acquired later (Fig. 4.4). Between the drugs with heavily skewed early or late distributions, resistance to some antibiotics is predicted to be acquired at intermediate orderings (for example, imipenem), or the ordering distribution is evenly distributed over the majority of possible positions (for example, amikacin).

There is significant multimodality in the ordering distributions of individual antibiotics (Fig. 4.4), suggesting resistance to these drugs might be acquired at different stages and through independent evolutionary pathways. Notably, levofloxacin and ampicillin display bimodality in their ordering distributions, predicting several different pathways in which these traits may be acquired early or late, but not at intermediate orderings (Fig. 4.4).

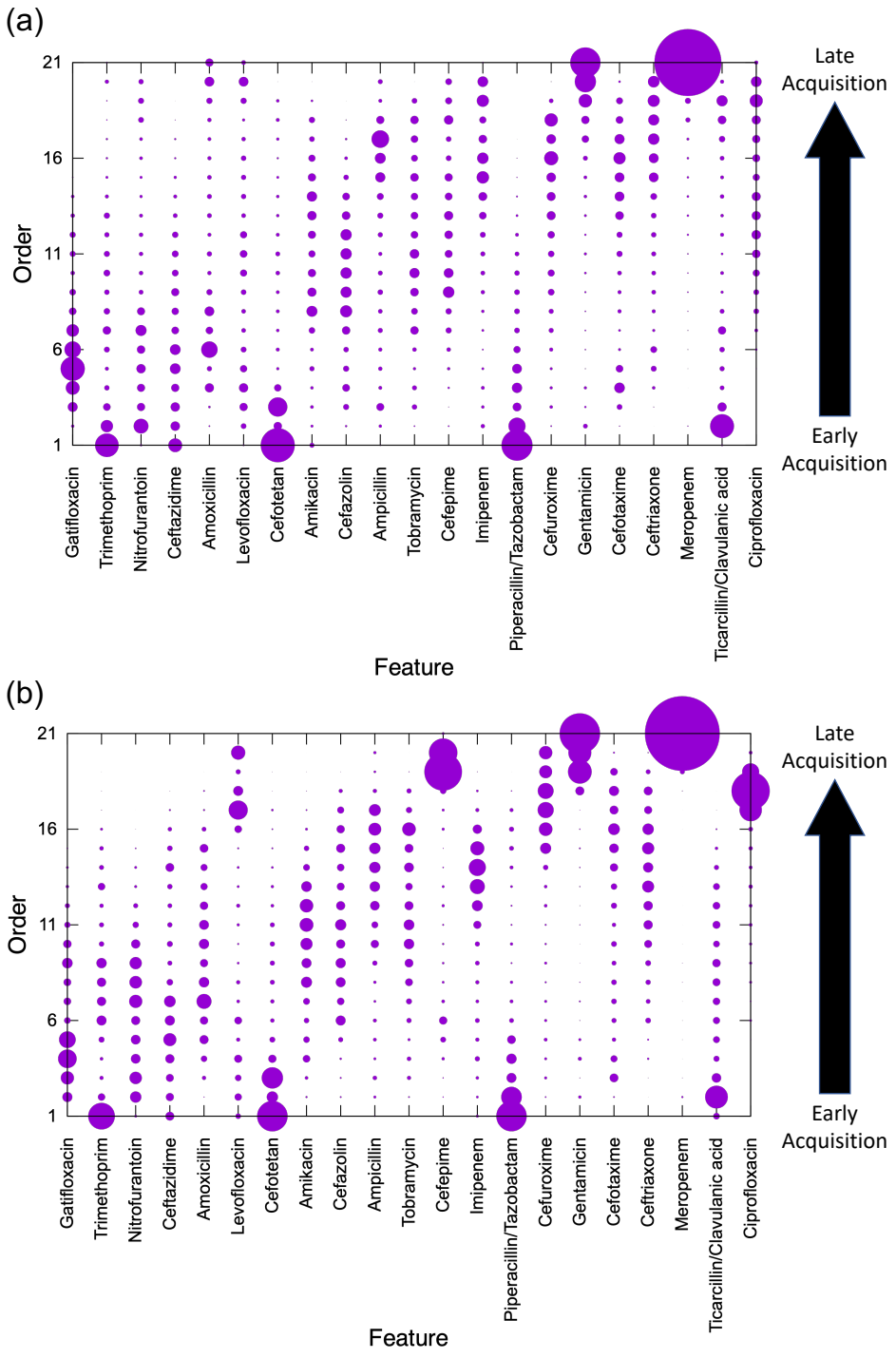


Figure 4.3: **Inferred pathways of antibiotic resistance acquisition in *P. aeruginosa* isolates.** Posterior orderings of drug-resistance features for the Test Dataset with random number seeds (a) 2 and (b) 3. Columns display the inferred ordering for resistance of 21 antibiotics, and the radius of a disk is proportional to the posterior probability that resistance to a given drug (horizontal axis) is acquired at a given position (vertical axis) in an evolutionary path starting with no drug resistance. Distributions are a relative statement about the ordering of different events and do not enforce that evolution must happen from no resistance to a full resistance profile.

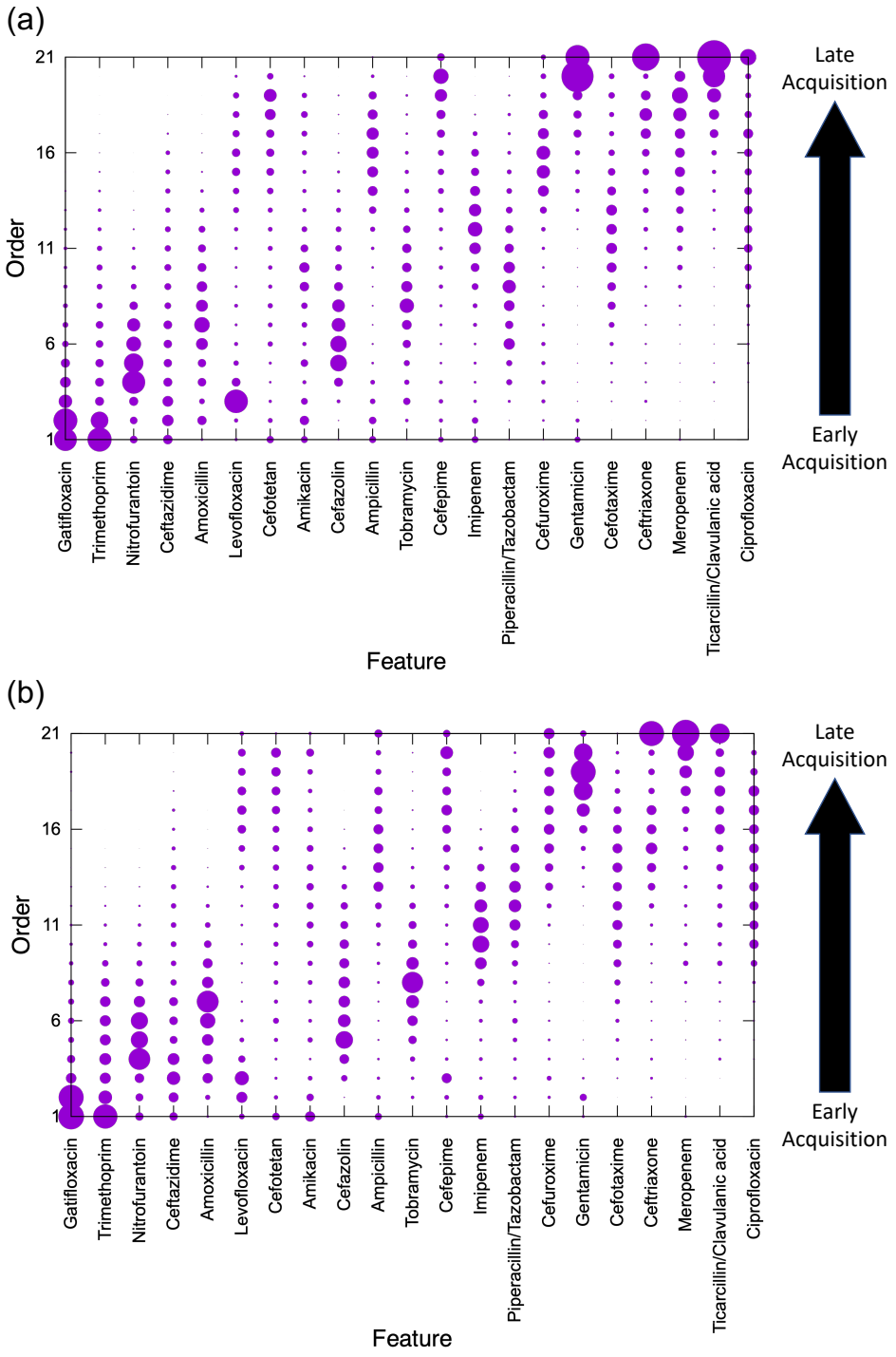


Figure 4.4: **Inferred acquisition order for drug resistance in *P. aeruginosa* clinical isolates.** Posterior orderings of drug-resistance features for the Resistance Dataset with random number seeds (a) 2 and (b) 3. Columns display the inferred ordering for resistance of 21 antibiotics, and the probability of drug-resistance acquisition in each position (purple circles increasing in radius for higher probabilities). Distributions are a relative statement about the ordering of different events and do not enforce that evolution must happen from no resistance to a full resistance profile.

4.4 Discussion

We have applied a highly generalizable statistical framework, combining stochastic modelling with Bayesian inference, to infer the structure of evolutionary pathways to explore the acquisition of antibiotic resistance across clinical *P. aeruginosa* isolates. We compared ordering distributions for two versions of the *P. aeruginosa* dataset, identifying structure and variability in antibiotic resistance acquisition (Figs. 4.3-4.4).

The inferred orderings from simulations with different random number seeds for the Resistance Dataset display a substantial level of convergence, providing a strong degree of confidence in the predicted behaviour, compared with the Test Dataset, which displays lack of confluence in the posterior distributions. The results from the Test Dataset will not contribute to the final discussion due to the uncertainty in the outputs. By considering the Resistance Dataset, the behaviour is being reported from the initial categorization of isolate susceptibility in the original study [93], where β-lactamase activity was not defined as an isolate displaying a resistance phenotype.

Our dynamic pathway inference results predict new insights into the structure of the trajectories by which antibiotic resistance is acquired. The Resistance Dataset provides predictions of drugs with early resistance acquisition, and multimodality in the ordering distributions of individual antibiotics, for example, levofloxacin (Figs. 4.4). Resistance to imipenem (Fig. 4.4) is predicted to be acquired at intermediate orderings, suggesting use of these drugs initially, or as last-line choices, may be more effective for optimal anti-*P. aeruginosa* therapy.

In the Resistance Dataset, antibiotics with predicted early resistance acquisition come from several antibiotic families (Table 4.1) and have a selection of primary targets [133]. An example of this is the fluoroquinolone family, where there is a possibility of acquiring levofloxacin and gatifloxacin resistance early, but ciprofloxacin resistance is inferred to be acquired at intermediate or late orderings. More generally, it can also be observed that ordering distributions of individual antibiotics from the same antibiotic family or primary mechanism of action do not exhibit consistent ordering behaviour, suggesting,

from this phenotypic data, the choice of prescribed antibiotic needs to be considered on an individual basis, and not by family or drug target.

Multimodality is the signature of independent evolutionary paths in a system, where a feature can be acquired at different points, for example, early or late but not at intermediate orderings. It is observed, in the original susceptibility data, resistance to levofloxacin is distributed evenly throughout the isolates (Fig. 4.2). Contained within the antibiotic susceptibility data are isolates resistant to amoxicillin and susceptible to levofloxacin, isolates susceptible to amoxicillin and resistant to levofloxacin, and isolates resistant to both antibiotics. This observation implies there are multiple independent trajectories in which these traits can be acquired, relating to independent evolutionary paths through the set of acquisitions, and is supported by the predicted bimodality of levofloxacin in posterior orderings (Fig. 4.4).

Resistance to the aminoglycosides, amikacin and tobramycin, is predicted to be acquired throughout the majority of orderings, with tobramycin resistance peaking at intermediate orderings. This suggests, tobramycin would be a reasonable choice as a later therapeutic option, but it may be better to limit the use of amikacin as resistance can be acquired at any stage. In a clinical-setting, it is suggested that broad spectrum β-lactams are the first therapeutic option for the treatment of *P. aeruginosa* infections, with colistin and aminoglycosides kept as later therapeutic options [266]. Our prediction, that resistance to gentamicin (an aminoglycoside, see Table 4.1) is heavily skewed towards late acquisition (Fig. 4.4), conforms with the clinical use of this drug as a later therapeutic option. The predicted ordering for gentamicin suggests that this antibiotic is a good choice to treat *P. aeruginosa* infections early on, indicating that results from our simulations need to be considered in a clinical context to provide the most appropriate proposals for drug use.

Together, the posterior of the Resistance Dataset provides a predictive framework which can be utilized, for example, to predict the next likely antibiotic resistance acquisition given an isolate from a real-world patient, and suggest an optimal treatment strategy

based on the isolate's current state. Whilst the posterior for the Resistance Dataset provides reasonable convergence, it is clear that simulations of both datasets need to be extended. This would provide longer chains to improve convergence and confidence in the posteriors.

CHAPTER 5

CONCLUSIONS

The accelerating spread of AMR is a huge global challenge. The methods used to treat infections for almost 100 years are failing, and the current development of antimicrobial treatments is not able to keep pace with the escalation of resistance. Numerous pathways conferring resistance contain an underlying regulatory network. Traditional research of AMR considers specific elements within a network thought to be involved with a resistance mechanism, rather than the system as a whole. Mathematical analysis is able to observe how a complete system makes a decision, rather than the effects of a small subset of network elements. By considering a resistance system as a whole, theoretical frameworks can complement traditional experimental methods, providing a combined strategy that may progress the current approach to tackling antimicrobial resistance.

In Chapter 2, we introduced a simple, ODE model of a general gene regulatory network consisting of two genes, representative of the core of several known decision-making motifs [81, 108, 109, 205, 318], to explore the effects of cell-to-cell ATP variability on cellular decision-making. From the analysis of our model, it is predicted that differences in cellular energy levels will cause differences in the dynamics and stable outcomes of cellular decision-making (Sections 2.3.1 and 2.3.6). Without a sufficient energy supply, decision-making circuits lose the capacity to select different phenotypes (Section 2.3.1), agreeing with experimentally observed behaviour within both eukaryotic and prokaryotic cells [78, 248, 251]. As energy availability increases, the decision-making landscape

expands, allowing, and stabilising, separate and intermediate states. Investigating the influence of noise in gene expression [215, 217, 274] on the behaviour and stability of this system will be valuable to shed further light on how decisions are made by a cell. Whilst this model gave us basic insights into how energy availability modulates the decision-making landscape, scaling this framework to larger GRNs develops a large parameter space with numerous unknown parameter values, resulting in the analysis of parameter sensitivity becoming computationally exhaustive.

The framework presented in Chapter 2 could be tailored to model different AMR mechanisms. The relationship between energy level and a cellular decision reducing antibiotic efficacy in prokaryotes (persister cells) has been previously predicted [166]. Applying our framework to GRNs predicted to be involved with the persister decision may improve our current understanding of the role of energy availability in persister formation. Developing our understanding of the role of energy availability in persister formation may help develop novel or enhanced treatments for recalcitrant infections caused by persister cells. This is of increasing importance, as recent theoretical models predict persisters facilitate the evolution of resistance within a bacterial population [145], identifying the larger potential role of persister cells in antimicrobial resistance. The broader application of this general framework to networks controlling AMR mechanisms, including efflux pumps and β-lactamase production, may identify novel methods for the treatment of AMR infections.

To investigate how heterogeneity in energy availability may affect cellular decision-making within a clonal population, we developed a preliminary experimental assay to image, segment and output quantitative single-cell values of ATP levels in an *E. coli* population (Section 2.4). The degree of heterogeneity we observed in an *E. coli* population was in qualitative agreement with the study from where the assay was first given [309], suggesting significant cell-to-cell variability in energy availability. This degree of heterogeneity in a clonal population has been experimentally observed in additional studies in eukaryotic and prokaryotic cells [63, 270, 313, 315]. By jointly tracking ATP concentrations and a pre-specified cellular decision in single-cells (or, for example, monitoring

continuous-time efflux gene expression), the role of energy availability on cellular-decision making may be further elucidated.

In Chapter 3, we applied a more coarse-grained modelling framework to investigate how efflux pump activity (in the presence and absence of stress) depends on the energy available to fuel the expression of efflux proteins and their regulators. Developed with a view for scalability to networks consisting of more genetic actors, a Boolean network methodology was proposed. We show that the inclusion of energy dependence within a Boolean network framework – something that has not been observed to be done previously – predicted that differences in available energy can contribute to pronounced cell heterogeneity in efflux pump gene expression. We found both the level of available energy and the stress duration influence the population-level variability in efflux pump expression (Section 3.3.4). Cells, in the absence of stress, are predicted to express heterogeneous pulses of efflux genes, and at low energy availability, we predict high variability and noise in gene expression pulses (Section 3.3.5). Pulse behaviour has been predicted at the protein level in a previous study of the *mar* operon [87], a regulator of *acrAB* in *E. coli*, and future studies may resolve the behaviour at the transcript level to test our theoretical predictions.

It is important to note some limitations of our modelling approach in Chapter 3. First, the coarse-grained nature of a Boolean model limits the precise quantitative agreement of individual observations with either more detailed models or biological reality. Instead we focus on the relative magnitudes and directions of behaviour – for example, does X increase or decrease Y, and does Z do so more. Second, the Boolean modelling framework is deliberately simple, and model refinements, including allowing cell death in the presence of stress (which could be an antibiotic exposure), would allow a more accurate representation of cell dynamics, and thus produce more detailed predictions on population behaviour, in stressed and stress-free environments. This development of the framework could improve its use in AMR applications.

There is accumulating evidence suggesting efflux pumps have broader functional roles

in antibiotic survival unrelated to the transport of toxic compounds. First, a study by El Meouche and Dunlop [73] reported bacteria with increased expression of the AcrAB-TolC system have lower expression of mismatch repair genes, facilitating mutations which can lead to permanent genetic changes bestowing resistance. Second, a recent publication reports multi-drug efflux-associated genes, *acrA*, *acrB*, and *tolC* showed significantly higher expression in persisters, and that *tolC* is important in promoting persister formation [229]; further research has suggested a link between the AcrAB-TolC efflux pump and persister development [40, 296]. The broader influences of heterogeneity in efflux expression in cellular populations remains to be uncovered, and further investigation into the predicted link between efflux pump activity and persisters may determine a general role of energy availability in antimicrobial resistance-related mechanisms.

In Chapter 4, we applied a Bayesian inference framework to predict pathways leading to AMR through the potential ordering of resistance acquisition to several antibiotics, in clinical *Pseudomonas aeruginosa* isolates. This ordering offers a different insight into the evolution of AMR bacteria. We have shown, from this alternative method to investigate the global AMR issue, general trends of early and late resistance acquisition in the ordering distributions of antibiotics (Section 4.3). In the early and late groups of antibiotics, no common characteristics, such as antibiotic family and target, were observed. It was also predicted that resistance to several antibiotics is acquired through multiple independent evolutionary paths. The potential possibilities of this, and similar applications, are extensive and could provide a basis for personalised and tailored treatments for patients, improving the outcomes against AMR infections. Here, we have only displayed one of the many capabilities that HyperTraPS may provide, and applications to further real-world genomic or phenotypic data will show the possible clinical application of this tool. In connection with the work performed here, there is the prospect of feeding the HyperTraPS framework outputs into mathematical models of GRNs governing resistance mechanisms. For example, following the development of a (energy-dependent) mathematical model of a GRN governing a resistance mechanism, HyperTraPS can be applied to phylogenetically-

related mutation data considering the genes in the modelled GRN [137, 234] and help predict the evolutionary paths of resistance acquisition through mutations in genes contained within the GRN. This information can then drive the relevant modifications and simulations of the mathematical model, providing predictions of the effect of each mutation on model outputs (such as the level of efflux pump expression). The collaborative strategy of investigating AMR mechanisms, could accelerate the development of novel targets or treatment strategies to combat AMR bacterial infections.

We believe that this work provides useful frameworks to investigate bacterial AMR. In Chapters 2-3, we have developed two generalisable methodologies that enable an investigation into the influence of intrinsic cell-to-cell ATP variability on cellular decision-making and AMR-related network behaviour. With our theoretical results supporting our hypotheses throughout, this should provide motivation for further investigation, both theoretically and experimentally, into the effect of cell-to-cell ATP heterogeneity on cellular behaviour. In Chapter 4, we have applied the HyperTraPS framework to predict evolutionary pathways of AMR acquisition, which we hope can progress forward and be used together with the energy-dependent mathematical model methodologies to improve the current understanding of AMR mechanisms, and inspire a broader investigation into employing ATP, or ATP-synthase, as an antimicrobial drug target.

LIST OF REFERENCES

[1] S. N. Abdi, R. Ghotaslou, K. Ganbarov, A. Mobed, A. Tanomand, M. Yousefi, M. Asgharzadeh, and H. S. Kafil. *Acinetobacter baumannii* efflux pumps and antibiotic resistance. **Infection and Drug Resistance**, 13:423, 2020.

[2] Y. M. Abouzeed, S. Baucheron, and A. Cloeckaert. *ramr* mutations involved in efflux-mediated multidrug resistance in *Salmonella enterica* serovar Typhimurium. **Antimicrobial Agents and Chemotherapy**, 52(7):2428–2434, 2008.

[3] E. P. Abraham and E. Chain. An enzyme from bacteria able to destroy penicillin. **Nature**, 146(3713):837–837, 1940.

[4] M. Acar, J. T. Mettetal, and A. Van Oudenaarden. Stochastic switching as a survival strategy in fluctuating environments. **Nature Genetics**, 40(4):471, 2008.

[5] G. K. Ackers, A. D. Johnson, and M. A. Shea. Quantitative model for gene regulation by λ phage repressor. **Proceedings of the National Academy of Sciences**, 79(4):1129–1133, 1982.

[6] R. Albert and H. G. Othmer. The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in *Drosophila melanogaster*. **Journal of Theoretical Biology**, 223(1):1–18, 2003.

[7] M. N. Alekshun and S. B. Levy. Alteration of the repressor activity of *marR*, the negative regulator of the *Escherichia coli marRAB* locus, by multiple chemicals in vitro. **Journal of Bacteriology**, 181(15):4669–4672, 1999.

[8] S. M. Amato, M. A. Orman, and M. P. Brynildsen. Metabolic control of persister formation in *Escherichia coli*. **Molecular Cell**, 50(4):475–487, 2013.

[9] A. Andersson, U. Ronner, and P. E. Granum. What problems does the food industry have with the spore-forming pathogens *Bacillus cereus* and *Clostridium perfringens*? **International Journal of Food Microbiology**, 28(2):145–155, 1995.

[10] J. Anes, M. P. McCusker, S. Fanning, and M. Martins. The ins and outs of RND efflux pumps in *Escherichia coli*. **Frontiers in Microbiology**, 6:587, 2015.

[11] P. Ao. Metabolic network modelling: Including stochastic effects. **Computers and Chemical Engineering**, 29(11-12):2297–2303, 2005.

[12] R. Ariza, S. Cohen, N. Bachhawat, S. Levy, and B. Demple. Repressor mutations in the *marRAB* operon that activate oxidative stress genes and multiple antibiotic resistance in *Escherichia coli*. **Journal of Bacteriology**, 176(1):143–148, 1994.

[13] R. R. Ariza, Z. Li, N. Ringstad, and B. Demple. Activation of multiple antibiotic resistance and binding of stress-inducible promoters by *Escherichia coli* Rob protein. **Journal of Bacteriology**, 177(7):1655–1661, 1995.

[14] L. Armstrong, K. Tilgner, G. Saretzki, S. P. Atkinson, M. Stojkovic, R. Moreno, S. Przyborski, and M. Lako. Human induced pluripotent stem cell lines show stress defense mechanisms and mitochondrial regulation similar to those of human embryonic stem cells. **Stem Cells**, 28(4):661–673, 2010.

[15] D. Ashiru-Oredope, M. Sharland, E. Charani, C. McNulty, and J. Cooke. Improving the quality of antibiotic prescribing in the NHS by developing a new antimicrobial stewardship programme: Start smart—then focus. **Journal of Antimicrobial Chemotherapy**, 67(1):51–63, 2012.

[16] S. Atsumi and J. W. Little. A synthetic phage λ regulatory circuit. **Proceedings of the National Academy of Sciences**, 103(50):19045–19050, 2006.

[17] E. Aun, V. Kisand, M. Laht, K. Telling, P. Kalmus, Ü. Väli, A. Brauer, M. Remm, and T. Tenson. Molecular characterization of *Enterococcus* isolates from different sources in Estonia reveals potential transmission of resistance genes among different reservoirs. **Frontiers in Microbiology**, 12:585, 2021.

[18] E. Aurell and K. Sneppen. Epigenetics as a first exit problem. **Physical Review Letters**, 88(4):048101, 2002.

[19] D. Austin, N. White, and R. Anderson. The dynamics of drug action on the within-host population growth of infectious agents: melding pharmacokinetics with pathogen population dynamics. **Journal of Theoretical Biology**, 194(3):313–339, 1998.

[20] S. Baker, P. T. Duy, T. V. T. Nga, T. T. N. Dung, V. V. Phat, T. T. Chau, A. K. Turner, J. Farrar, and M. F. Boni. Fitness benefits in fluoroquinolone-resistant *Salmonella* typhi in the absence of antimicrobial pressure. **eLife**, 2:e01229, 2013.

[21] N. Q. Balaban, J. Merrin, R. Chait, L. Kowalik, and S. Leibler. Bacterial persistence as a phenotypic switch. **Science**, 305(5690):1622–1625, 2004.

[22] G. Balázsi, A. van Oudenaarden, and J. J. Collins. Cellular decision making and biological noise: from microbes to mammals. **Cell**, 144(6):910–925, 2011.

[23] V. Balbas-Martinez, L. Ruiz-Cerda, I. Irurzun-Arana, I. Gonzalez-Garcia, A. Vermeulen, J. D. Gomez-Mantilla, and I. F. Troconiz. A systems pharmacology model for inflammatory bowel disease. **PloS One**, 13(3), 2018.

[24] M. Banzhaf, O. Resendis-Antonio, and M. L. Zepeda-Mendoza. Uncovering the dynamic mechanisms of the *Pseudomonas aeruginosa* quorum sensing and virulence networks using boolean modelling. **IEEE Transactions on NanoBioscience**, 2020.

[25] T. M. Barbosa and S. B. Levy. Differential expression of over 60 chromosomal genes in *Escherichia coli* by constitutive expression of MarA. **Journal of Bacteriology**, 182(12):3467–3474, 2000.

[26] S. Baucheron, K. Nishino, I. Monchaux, S. Canepa, M.-C. Maurel, F. Coste, A. Roussel, A. Cloeckaert, and E. Giraud. Bile-mediated activation of the *acrAB* and *tolc* multidrug efflux genes occurs mainly through transcriptional derepression of *ramA* in *Salmonella enterica* serovar Typhimurium. **Journal of Antimicrobial Chemotherapy**, 69(9):2400–2406, 2014.

[27] S. Baucheron, S. Tyler, D. Boyd, M. R. Mulvey, E. Chaslus-Dancla, and A. Cloeckaert. AcrAB-TolC directs efflux-mediated multidrug resistance in *Salmonella enterica* serovar Typhimurium DT104. **Antimicrobial Agents and Chemotherapy**, 48(10):3729–3735, 2004.

[28] J. Bendig, P. Kyle, P. Giangrande, D. Samson, and B. Azadian. Two neutropenic patients with multiple resistant *Pseudomonas aeruginosa* septicaemia treated with ciprofloxacin. **Journal of the Royal Society of Medicine**, 80(5):316–317, 1987.

[29] M. H. Bennik, P. J. Pomposiello, D. F. Thorne, and B. Demple. Defining a *rob* regulon in *Escherichia coli* by using transposon mutagenesis. **Journal of Bacteriology**, 182(13):3794–3801, 2000.

[30] T. Bergmiller, A. M. Andersson, K. Tomasek, E. Balleza, D. J. Kiviet, R. Hauschild, G. Tkačik, and C. C. Guet. Biased partitioning of the multidrug efflux pump AcrAB-TolC underlies long-lived phenotypic heterogeneity. **Science**, 356(6335):311–315, 2017.

[31] J. Bigger. Treatment of *Staphylococcal* infections with penicillin by intermittent sterilisation. **Lancet**, pages 497–500, 1944.

[32] J. M. Blair and L. J. Piddock. Structure, function and inhibition of RND efflux pumps in Gram-negative bacteria: an update. **Current Opinion in Microbiology**, 12(5):512–519, 2009.

[33] J. M. Blair, M. A. Webber, A. J. Baylay, D. O. Ogbolu, and L. J. Piddock. Molecular mechanisms of antibiotic resistance. **Nature Reviews Microbiology**, 13(1):42–51, 2015.

[34] W. J. Blake, G. Balázsi, M. A. Kohanski, F. J. Isaacs, K. F. Murphy, Y. Kuang, C. R. Cantor, D. R. Walt, and J. J. Collins. Phenotypic consequences of promoter-mediated transcriptional noise. **Molecular Cell**, 24(6):853–865, 2006.

[35] P. Blanco, S. Hernando-Amado, J. A. Reales-Calderon, F. Corona, F. Lira, M. Alcalde-Rico, A. Bernardini, M. B. Sanchez, and J. L. Martinez. Bacterial multidrug efflux pumps: much more than antibiotic resistance determinants. **Microorganisms**, 4(1):14, 2016.

[36] P. Bokes, J. R. King, A. T. Wood, and M. Loose. Transcriptional bursting diversifies the behaviour of a toggle switch: hybrid simulation of stochastic gene expression. **Bulletin of Mathematical Biology**, 75(2):351–371, 2013.

[37] M. Bootsma, M. van der Horst, T. Guryeva, B. Ter Kuile, and O. Diekmann. Modeling non-inherited antibiotic resistance. **Bulletin of Mathematical Biology**, 74(8):1691–1705, 2012.

[38] D. Botman, J. H. van Heerden, and B. Teusink. An improved ATP FRET sensor for yeast shows heterogeneity during nutrient transitions. **bioRxiv**, 2019.

[39] C. G. Bowsher and P. S. Swain. Environmental sensing, information transfer, and cellular decision-making. **Current Opinion in Biotechnology**, 28:149–155, 2014.

[40] B. A. Byrd, B. Zenick, M. C. Rocha-Granados, H. E. Englander, P. J. Hare, T. J. LaGree, A. M. DeMarco, and W. W. Mok. The AcrAB-TolC efflux pump impacts persistence and resistance development in stationary-phase *Escherichia coli* following delafloxacin treatment. **Antimicrobial Agents and Chemotherapy**, 65(8):e00281–21, 2021.

[41] E. Carson and C. Cobelli. **Modelling Methodology for Physiology and Medicine**. Newnes, 2013.

[42] G. Carvalho, C. Guilhen, D. Balestrino, C. Forestier, and J.-D. Mathias. Relating switching rates between normal and persister cells to substrate and antibiotic concentrations: a mathematical modelling approach supported by experiments. **Microbial Biotechnology**, 10(6):1616–1627, 2017.

[43] N. Casali, V. Nikolayevskyy, Y. Balabanova, S. R. Harris, O. Ignatyeva, I. Kontsevaya, J. Corander, J. Bryant, J. Parkhill, S. Nejentsev, *et al.* Evolution and transmission of drug-resistant tuberculosis in a Russian population. **Nature Genetics**, 46(3):279–286, 2014.

[44] Q. Chai, S. R. Webb, Z. Wang, R. E. Dutch, and Y. Wei. Study of the degradation of a multidrug transporter using a non-radioactive pulse chase method. **Analytical and Bioanalytical Chemistry**, 408(27):7745–7751, 2016.

[45] H. H. Chang, M. Hemberg, M. Barahona, D. E. Ingber, and S. Huang. Transcriptome-wide noise controls lineage choice in mammalian progenitor cells. **Nature**, 453(7194):544, 2008.

[46] D. A. Charlebois, G. Balázsi, and M. Kærn. Coherent feedforward transcriptional regulatory motifs enhance drug resistance. **Physical Review E**, 89(5):052708, 2014.

[47] M. Chaves, E. D. Sontag, and R. Albert. Methods of robustness analysis for boolean models of gene control networks. **IEE Proceedings-Systems Biology**, 153(4):154–167, 2006.

[48] J. Chen, X. Wang, S. Wang, C. Chen, W. Zhang, and Y. Zhang. Ultra-rapid drug susceptibility testing for *Klebsiella pneumoniae* clinical isolates in 60min by SYBR green I/propidium iodide viability assay. **Frontiers in Microbiology**, 12:2137, 2021.

[49] J. L. Cherry and F. R. Adler. How to make a biological switch. **Journal of Theoretical Biology**, 203(2):117–133, 2000.

[50] Y. Chitkara and T. Feierabend. Endogenous and exogenous infection with *Pseudomonas aeruginosa* in a burns unit. **International Surgery**, 66(3):237–240, 1981.

[51] S. Chung, P. P. Dzeja, R. S. Faustino, C. Perez-Terzic, A. Behfar, and A. Terzic. Mitochondrial oxidative metabolism is required for the cardiac differentiation of stem cells. **Nature Reviews Cardiology**, 4(S1):S60, 2007.

[52] G. Cingolani and T. M. Duncan. Structure of the ATP synthase catalytic complex (F_1) from *Escherichia coli* in an autoinhibited conformation. **Nature Structural and Molecular Biology**, 18(6):701, 2011.

[53] N. R. Cohen, M. A. Lobritz, and J. J. Collins. Microbial persistence and the road to drug resistance. **Cell Host and Microbe**, 13(6):632–642, 2013.

[54] S. P. Cohen, H. Hächler, and S. Levy. Genetic and functional analysis of the multiple antibiotic resistance (*mar*) locus in *Escherichia coli*. **Journal of Bacteriology**, 175(5):1484–1492, 1993.

[55] S. P. Cohen, S. B. Levy, J. Foulds, and J. L. Rosner. Salicylate induction of antibiotic resistance in *Escherichia coli*: activation of the *mar* operon and a *mar*-independent pathway. **Journal of Bacteriology**, 175(24):7856–7862, 1993.

[56] B. P. Conlon, S. E. Rowe, A. B. Gandt, A. S. Nuxoll, N. P. Donegan, E. A. Zalis, G. Clair, J. N. Adkins, A. L. Cheung, and K. Lewis. Persister formation in *Staphylococcus aureus* is associated with ATP depletion. **Nature Microbiology**, 1(5):1–7, 2016.

[57] P. P. Cook and M. Gooch. Long-term effects of an antimicrobial stewardship programme at a tertiary-care teaching hospital. **International Journal of Antimicrobial Agents**, 45(3):262–267, 2015.

[58] F. F. Correia, A. D'Onofrio, T. Rejtar, L. Li, B. L. Karger, K. Makarova, E. V. Koonin, and K. Lewis. Kinase activity of overexpressed *hipA* is required for growth arrest and multidrug tolerance in *Escherichia coli*. **Journal of Bacteriology**, 188(24):8360–8367, 2006.

[59] S. Cortijo, Z. Aydin, S. Ahnert, and J. C. Locke. Widespread inter-individual gene expression variability in *Arabidopsis thaliana*. **Molecular Systems Biology**, 15(1):e8591, 2019.

[60] J. L. Cottell, M. A. Webber, and L. J. Piddock. Persistence of transferable extended-spectrum-β-lactamase resistance in the absence of antibiotic pressure. **Antimicrobial Agents and Chemotherapy**, 56(9):4703–4706, 2012.

[61] S. E. Dallidis and I. G. Karafyllidis. Boolean network model of the *Pseudomonas aeruginosa* quorum sensing circuits. **IEEE Transactions on Nanobioscience**, 13(3):343–349, 2014.

[62] R. P. das Neves, N. S. Jones, L. Andreu, R. Gupta, T. Enver, and F. J. Iborra. Connecting variability in global transcription rate to mitochondrial variability. **PLoS Biology**, 8(12):e1000560, 2010.

[63] V. De Col, P. Fuchs, T. Nietzel, M. Elsaesser, C. P. Voon, A. Candeo, I. Seeliger, M. D. Fricker, C. Grefen, I. M. Møller, *et al.* ATP sensing in living plant cells reveals tissue gradients and stress dynamics of energy physiology. **eLife**, 6:e26770, 2017.

[64] G. Demarre, V. Prudent, H. Schenk, E. Rousseau, M.-A. Bringer, N. Barnich, G. T. Van Nhieu, S. Rimsky, S. De Monte, and O. Espéli. The Crohn's disease-associated *Escherichia coli* strain LF82 relies on SOS and stringent responses to survive, multiply and tolerate antibiotics within macrophages. **PLoS Pathogens**, 15(11), 2019.

[65] X. Didelot, R. Bowden, D. J. Wilson, T. E. Peto, and D. W. Crook. Transforming clinical microbiology with bacterial genome sequencing. **Nature Reviews Genetics**, 13(9):601–612, 2012.

[66] T. Dörr, M. Vulić, and K. Lewis. Ciprofloxacin causes persister formation by inducing the TisB toxin in *Escherichia coli*. **PLoS Biology**, 8(2), 2010.

[67] S. P. M. Drescher, S. W. Gallo, P. M. A. Ferreira, C. A. S. Ferreira, and S. D. de Oliveira. *Salmonella enterica* persister cells form unstable small colony variants after in vitro exposure to ciprofloxacin. **Scientific Reports**, 9(1):7232, 2019.

[68] I. Duncan and B. Jeans. Lincomycin in hospital practice. **Canadian Medical Association Journal**, 93(13):685, 1965.

[69] V. Duval and I. M. Lister. MarA, SoxS and Rob of *Escherichia coli* – global regulators of multidrug resistance, virulence and stress response. **International Journal of Biotechnology for Wellness Industries**, 2(3):101, 2013.

[70] H. Easwaran, H.-C. Tsai, and S. B. Baylin. Cancer epigenetics: tumor heterogeneity, plasticity of stem-like states, and drug resistance. **Molecular Cell**, 54(5):716–727, 2014.

[71] D. J. Eaves, V. Ricci, and L. J. Piddock. Expression of *acrB*, *acrF*, *acrD*, *marA*, and *soxS* in *Salmonella enterica* serovar Typhimurium: role in multiple antibiotic resistance. **Antimicrobial Agents and Chemotherapy**, 48(4):1145–1150, 2004.

[72] European Centre for Disease Prevention and Control. Antimicrobial resistance surveillance in Europe 2015. **Annual Report of the European Antimicrobial Resistance Surveillance Network (EARS-Net)**, 2017.

[73] I. El Meouche and M. J. Dunlop. Heterogeneity in efflux pump expression predisposes antibiotic-resistant cells to mutation. **Science**, 362(6415):686–690, 2018.

[74] I. El Meouche, Y. Siu, and M. J. Dunlop. Stochastic expression of a multiple antibiotic resistance activator confers transient resistance in single cells. **Scientific Reports**, 6(1):1–9, 2016.

[75] H. El Samad, M. Khammash, L. Petzold, and D. Gillespie. Stochastic modelling of gene regulatory networks. **International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal**, 15(15):691–711, 2005.

[76] E. D. Esplin, L. Oei, and M. P. Snyder. Personalized sequencing and the future of medicine: discovery, diagnosis and defeat of disease. **Pharmacogenomics**, 15(14):1771–1790, 2014.

[77] M. J. Everett, Y. F. Jin, V. Ricci, and L. Piddock. Contributions of individual mechanisms to fluoroquinolone resistance in 36 *Escherichia coli* strains isolated from humans and animals. **Antimicrobial Aagents and Chemotherapy**, 40(10):2380–2386, 1996.

[78] T. Ezashi, P. Das, and R. M. Roberts. Low O_2 tensions and the prevention of differentiation of hES cells. **Proceedings of the National Academy of Sciences**, 102(13):4783–4788, 2005.

[79] M. E. Falagas and D. E. Karageorgopoulos. Pandrug resistance (PDR), extensive drug resistance (XDR), and multidrug resistance (MDR) among Gram-negative bacilli: need for international harmonization in terminology. **Clinical Infectious Diseases**, 46(7):1121–1122, 2008.

[80] M. Fauvart, V. N. De Groote, and J. Michiels. Role of persister cells in chronic infections: clinical relevance and perspectives on anti-persister therapies. **Journal of Medical Microbiology**, 60(6):699–709, 2011.

[81] H. Feng and J. Wang. A new mechanism of stem cell differentiation through slow binding/unbinding of regulators to genes. **Scientific Reports**, 2:550, 2012.

[82] L. Fernández and R. E. Hancock. Adaptive and mutational resistance: role of porins and efflux pumps in drug resistance. **Clinical Microbiology Reviews**, 25(4):661–681, 2012.

[83] H. G. Floss and T.-W. Yu. Rifamycin mode of action, resistance, and biosynthesis. **Chemical Reviews**, 105(2):621–632, 2005.

[84] K. J. Forsberg, A. Reyes, B. Wang, E. M. Selleck, M. O. Sommer, and G. Dantas. The shared antibiotic resistome of soil bacteria and human pathogens. **Science**, 337(6098):1107–1111, 2012.

[85] D. Fraser and M. Kaern. A chance at survival: gene expression noise and phenotypic diversification strategies. **Molecular Microbiology**, 71(6):1333–1340, 2009.

[86] J. A. Gama, R. Zilhão, and F. Dionisio. Plasmid interactions can improve plasmid persistence in bacterial populations. **Frontiers in Microbiology**, 11:2033, 2020.

[87] J. Garcia-Bernardo and M. J. Dunlop. Tunable stochastic pulsing in the *Escherichia coli* multiple antibiotic resistance network from interlinked positive and negative feedback loops. **PLoS Computational Biology**, 9(9):e1003229, 2013.

[88] J. Garcia-Bernardo and M. J. Dunlop. Noise and low-level dynamics can coordinate multicomponent bet hedging mechanisms. **Biophysical Journal**, 108(1):184–193, 2015.

[89] L. Gelens, L. Hill, A. Vandervelde, J. Danckaert, and R. Loris. A general model for toxin-antitoxin module dynamics can explain persister cell formation in *Escherichia coli*. **PLoS Computational Biology**, 9(8), 2013.

[90] A. M. George and S. B. Levy. Amplifiable resistance to tetracycline, chloramphenicol, and other antibiotics in *Escherichia coli*: involvement of a non-plasmid-determined efflux of tetracycline. **Journal of Bacteriology**, 155(2):531–540, 1983.

[91] E. Germain, D. Castro-Roa, N. Zenkin, and K. Gerdes. Molecular mechanism of bacterial persistence by *hipA*. **Molecular Cell**, 52(2):248–254, 2013.

[92] M. J. Gill, N. P. Brenwald, and R. Wise. Identification of an efflux pump gene, *pmrA*, associated with fluoroquinolone resistance in *Streptococcus pneumoniae*. **Antimicrobial Agents and Chemotherapy**, 43(1):187–189, 1999.

[93] E. M. González-Olvera, R. Pérez-Morales, A. G. Zamora, G. Castro-Escarpulli, I. Palma-Martínez, and J. J. Alba-Romero. Antibiotic resistance, virulence factors and genotyping of *Pseudomonas aeruginosa* in public hospitals of northeastern Mexico. **The Journal of Infection in Developing Countries**, 13(05):374–383, 2019.

[94] M. L. Grayson, S. E. Cosgrove, S. Crowe, W. Hope, J. S. McCarthy, J. Mills, J. W. Mouton, and D. L. Paterson. **Kucers' The Use of Antibiotics: A Clinical Review of Antibacterial, Antifungal, Antiparasitic, and Antiviral Drugs**. CRC Press, 2017.

[95] S. F. Greenbury, M. Barahona, and I. G. Johnston. Hypertraps: Inferring probabilistic patterns of trait acquisition in evolutionary and disease progression pathways. **Cell Systems**, 10(1):39–51, 2020.

[96] F. Greil. Boolean networks as modeling framework. **Frontiers in Plant Science**, 3:178, 2012.

[97] T. Gumbo, A. Louie, M. R. Deziel, L. M. Parsons, M. Salfinger, and G. L. Drusano. Selection of a moxifloxacin dose that suppresses drug resistance in *Mycobacterium tuberculosis*, by use of an in vitro pharmacodynamic infection model and mathematical modeling. **Journal of Infectious Diseases**, 190(9):1642–1651, 2004.

[98] H. Guo, T. Suzuki, and J. L. Rubinstein. Structure of a bacterial ATP synthase. **eLife**, 8:e43128, 2019.

[99] H. Hächler, S. P. Cohen, and S. B. Levy. *marA*, a regulated locus which controls expression of chromosomal multiple antibiotic resistance in *Escherichia coli*. **Journal of Bacteriology**, 173(17):5532–5538, 1991.

[100] H. Hanberger, D. Arman, H. Gill, V. Jindrák, S. Kalenic, A. Kurcz, M. Licker, P. Naaber, E. A. Scicluna, V. Vanis, *et al.* Surveillance of microbial resistance in European intensive care units: a first report from the care-ICU programme for improved infection control. **Intensive Care Medicine**, 35(1):91–100, 2009.

[101] I. Harvey and T. Bossomaier. Time out of joint: attractors in asynchronous random boolean networks. **Proceedings of the Fourth European Conference on Artificial Life**. MIT Press, 1997.

[102] P. Higgins, A. Fluit, D. Milatovic, J. Verhoef, and F.-J. Schmitz. Mutations in GyrA, ParC, MexR and NfxB in clinical isolates of *Pseudomonas aeruginosa*. **International Journal of Antimicrobial Agents**, 21(5):409–413, 2003.

[103] T. Hishinuma, T. Tada, K. Kuwahara-Arai, N. Yamamoto, M. Shimojima, and T. Kirikae. Spread of GES-5 carbapenemase-producing *Pseudomonas aeruginosa* clinical isolates in Japan due to clonal expansion of ST235. **PLoS One**, 13(11):e0207134, 2018.

[104] E. C. Hobbs, X. Yin, B. J. Paul, J. L. Astarita, and G. Storz. Conserved small protein associates with the multidrug efflux pump AcrB and differentially affects antibiotic resistance. **Proceedings of the National Academy of Sciences**, 109(41):16696–16701, 2012.

[105] G. Hochhaus and H. Derendorf. Dose optimization based on pharmacokinetic-pharmacodynamic modeling. **Handbook of Pharmacokinetic/Pharmacodynamic Correlation**. CRC Press Inc. Boca Raton, 1995.

[106] S. Y. Holbrook and S. Garneau-Tsodikova. Evaluation of aminoglycoside and carbapenem resistance in a collection of drug-resistant *Pseudomonas aeruginosa* clinical isolates. **Microbial Drug Resistance**, 24(7):1020–1030, 2018.

[107] P.-R. Hsueh, L.-J. Teng, P.-C. Yang, Y.-C. Chen, S.-W. Ho, and K.-T. Luh. Persistence of a multidrug-resistant *Pseudomonas aeruginosa* clone in an intensive care burn unit. **Journal of Clinical Microbiology**, 36(5):1347–1351, 1998.

[108] S. Huang. Non-genetic heterogeneity of cells in development: more than just noise. **Development**, 136(23):3853–3862, 2009.

[109] S. Huang, Y.-P. Guo, G. May, and T. Enver. Bifurcation dynamics in lineage-commitment in bipotent progenitor cells. **Developmental Biology**, 305(2):695–713, 2007.

[110] B. S. Hughes, A. J. Cullum, and A. F. Bennett. An experimental evolutionary study on adaptation to temporally fluctuating pH in *Escherichia coli*. **Physiological and Biochemical Zoology**, 80(4):406–421, 2007.

[111] H. Imamura, K. P. H. Nhat, H. Togawa, K. Saito, R. Iino, Y. Kato-Yamada, T. Nagai, and H. Noji. Visualization of ATP levels inside single living cells with fluorescence resonance energy transfer-based genetically encoded indicators. **Proceedings of the National Academy of Sciences**, 106(37):15651–15656, 2009.

[112] N. T. Ingolia. Topology and robustness in the *Drosophila* segment polarity network. **PLoS Biology**, 2(6):e123, 2004.

[113] I. Irurzun-Arana, J. M. Pastor, I. F. Trocóniz, and J. D. Gómez-Mantilla. Advanced boolean modeling of biological networks applied to systems pharmacology. **Bioinformatics**, 33(7):1040–1048, 2017.

[114] S. Iwata, J. W. Lee, K. Okada, J. K. Lee, M. Iwata, B. Rasmussen, T. A. Link, S. Ramaswamy, and B. K. Jap. Complete structure of the 11-subunit bovine mitochondrial cytochrome bc_1 complex. **Science**, 281(5373):64–71, 1998.

[115] S. Jabbari, J. T. Heap, and J. R. King. Mathematical modelling of the sporulation-initiation network in *Bacillus subtilis* revealing the dual role of the putative quorum-sensing signal molecule PhrA. **Bulletin of Mathematical Biology**, 73(1):181–211, 2011.

[116] P. R. Jensen and O. Michelsen. Carbon and energy metabolism of ATP mutants of *Escherichia coli*. **Journal of Bacteriology**, 174(23):7635–7641, 1992.

[117] I. G. Johnston. Efficient parametric inference for stochastic biological systems with measured variability. **Statistical Applications in Genetics and Molecular Biology**, 13(3):379–390, 2014.

[118] I. G. Johnston and G. W. Bassel. Identification of a bet-hedging network motif generating noise in hormone concentrations and germination propensity in *Arabidopsis*. **Journal of The Royal Society Interface**, 15(141):20180042, 2018.

[119] I. G. Johnston, B. Gaal, R. P. das Neves, T. Enver, F. J. Iborra, and N. S. Jones. Mitochondrial variability as a source of extrinsic cellular noise. **PLoS Computational Biology**, 8(3):e1002416, 2012.

[120] I. G. Johnston, T. Hoffmann, S. F. Greenbury, O. Cominetti, M. Jallow, D. Kwiatkowski, M. Barahona, N. S. Jones, and C. Casals-Pascual. Precision identification of high-risk phenotypes and progression pathways in severe malaria without requiring longitudinal data. **NPJ Digital Medicine**, 2(1):1–9, 2019.

[121] I. G. Johnston and B. P. Williams. Evolutionary inference across eukaryotes identifies specific pressures favoring mitochondrial gene retention. **Cell Systems**, 2(2):101–111, 2016.

[122] M. Kafri, E. Metzl-Raz, G. Jona, and N. Barkai. The cost of protein production. **Cell Reports**, 14(1):22–31, 2016.

[123] T. Kalmar, C. Lim, P. Hayward, S. Munoz-Descalzo, J. Nichols, J. Garcia-Ojalvo, and A. M. Arias. Regulated fluctuations in Nanog expression mediate cell fate decisions in embryonic stem cells. **PLoS Biology**, 7(7):e1000149, 2009.

[124] U. Karaoz, T. Murali, S. Letovsky, Y. Zheng, C. Ding, C. R. Cantor, and S. Kasif. Whole-genome annotation by using evidence integration in functional-linkage networks. **Proceedings of the National Academy of Sciences**, 101(9):2888–2893, 2004.

[125] G. Karlebach and R. Shamir. Modelling and analysis of gene regulatory networks. **Nature Reviews Molecular Cell Biology**, 9(10):770–780, 2008.

[126] P. Katajisto, J. Döhla, C. L. Chaffer, N. Pentinmikko, N. Marjanovic, S. Iqbal, R. Zoncu, W. Chen, R. A. Weinberg, and D. M. Sabatini. Asymmetric apportioning of aged mitochondria between daughter cells is required for stemness. **Science**, 348(6232):340–343, 2015.

[127] S. A. Kauffman. Metabolic stability and epigenesis in randomly constructed genetic nets. **Journal of Theoretical Biology**, 22(3):437–467, 1969.

[128] I. Keren, D. Shah, A. Spoering, N. Kaldalu, and K. Lewis. Specialized persister cells and the mechanism of multidrug tolerance in *Escherichia coli*. **Journal of Bacteriology**, 186(24):8172–8180, 2004.

[129] W. Kern, M. Oethinger, A. Jellen-Ritter, and S. Levy. Non-target gene mutations in the development of fluoroquinolone resistance in *Escherichia coli*. **Antimicrobial Agents and Chemotherapy**, 44(4):814–820, 2000.

[130] R. Kerr, S. Jabbari, J. M. Blair, and I. G. Johnston. Dynamic boolean modelling reveals the influence of energy supply on bacterial efflux pump expression. **Journal of the Royal Society Interface**, 19(186):20210771, 2021.

[131] R. Kerr, S. Jabbari, and I. G. Johnston. Intracellular energy variability modulates cellular decision-making capacity. **Scientific Reports**, 9(1):1–12, 2019.

[132] M. Kielhofner, R. L. Atmar, R. J. Hamill, and D. M. Musher. Life-threatening *Pseudomonas aeruginosa* infections in patients with human immunodeficiency virus infection. **Clinical Infectious Diseases**, 14(2):403–411, 1992.

[133] S. Kim, J. Chen, T. Cheng, A. Gindulyte, J. He, S. He, Q. Li, B. A. Shoemaker, P. A. Thiessen, B. Yu, *et al.* PubChem in 2021: new data content and improved web interfaces. **Nucleic Acids Research**, 49(D1):D1388–D1395, 2021.

[134] S. B. Korch, T. A. Henderson, and T. M. Hill. Characterization of the *hipA7* allele of *Escherichia coli* and evidence that high persistence is governed by (p)ppGpp synthesis. **Molecular Microbiology**, 50(4):1199–1213, 2003.

[135] E. Kussell, R. Kishony, N. Q. Balaban, and S. Leibler. Bacterial persistence: a model of survival in changing environments. **Genetics**, 169(4):1807–1814, 2005.

[136] M. D. LaFleur, Q. Qi, and K. Lewis. Patients with long-term oral carriage harbor high-persister mutants of *Candida albicans*. **Antimicrobial Agents and Chemotherapy**, 54(1):39–44, 2010.

[137] M. Lam, R. R. Wick, S. C. Watts, L. T. Cerdeira, K. L. Wyres, and K. E. Holt. A genomic surveillance framework and genotyping tool for *Klebsiella pneumoniae* and its related species complex. **Nature Communications**, 12(1):1–16, 2021.

[138] J.-P. Langlois, G. Millette, I. Guay, A. Dube-Duquette, S. Chamberland, P.-É. Jacques, S. Rodrigue, K. Bouarab, E. Marsault, and F. Malouin. Bactericidal activity of the bacterial ATP synthase inhibitor tomatidine and the combination of tomatidine and aminoglycoside against persistent and virulent forms of *Staphylococcus aureus*. **Frontiers in Microbiology**, 11:805, 2020.

[139] A. Lawler, V. Ricci, S. Busby, and L. Piddock. Genetic inactivation of *acrAB* or inhibition of efflux induces expression of *ramA*. **Journal of Antimicrobial Chemotherapy**, 68(7):1551–1557, 2013.

[140] F. Le Goffic, M.-L. Capmau, D. Bonnet, C. Cerceau, C. Soussy, A. Dublanchet, and J. Duval. Plasmid-mediated pristinamycin resistance. **The Journal of Antibiotics**, 30(8):665–669, 1977.

[141] B. S. Lee, N. P. Kalia, X. E. F. Jin, E. J. Hasenoehrl, M. Berney, and K. Pethe. Inhibitors of energy metabolism interfere with antibiotic-induced death in mycobacteria. **Journal of Biological Chemistry**, 294(6):1936–1943, 2019.

[142] I. Lee, S. V. Date, A. T. Adai, and E. M. Marcotte. A probabilistic functional network of yeast genes. **Science**, 306(5701):1555–1558, 2004.

[143] N. A. Lerminiaux and A. D. Cameron. Horizontal transfer of antibiotic resistance genes in clinical environments. **Canadian Journal of Microbiology**, 65(1):34–44, 2019.

[144] B. R. Levin and D. E. Rozen. Non-inherited antibiotic resistance. **Nature Reviews Microbiology**, 4(7):556–562, 2006.

[145] I. Levin-Reisman, I. Ronin, O. Gefen, I. Braniss, N. Shoresh, and N. Q. Balaban. Antibiotic tolerance facilitates the evolution of resistance. **Science**, 355(6327):826–830, 2017.

[146] S. B. Levy. Active efflux mechanisms for antimicrobial resistance. **Antimicrobial Agents and Chemotherapy**, 36(4):695, 1992.

[147] S. F. Levy. Cellular heterogeneity: benefits besides bet-hedging. **Current Biology**, 26(9):R355–R357, 2016.

[148] S. F. Levy, N. Ziv, and M. L. Siegal. Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. **PLoS Biology**, 10(5):e1001325, 2012.

[149] K. Lewis. Persister cells. **Annual Review of Microbiology**, 64:357–372, 2010.

[150] F. Li, T. Long, Y. Lu, Q. Ouyang, and C. Tang. The yeast cell-cycle network is robustly designed. **Proceedings of the National Academy of Sciences**, 101(14):4781–4786, 2004.

[151] X.-Z. Li, D. M. Livermore, and H. Nikaido. Role of efflux pump(s) in intrinsic resistance of *Pseudomonas aeruginosa*: resistance to tetracycline, chloramphenicol,

and norfloxacin. **Antimicrobial Agents and Chemotherapy**, 38(8):1732–1741, 1994.

[152] A. Lipshtat, A. Loinger, N. Q. Balaban, and O. Biham. Genetic toggle switch without cooperative binding. **Physical Review Letters**, 96(18):188101, 2006.

[153] X. Liu, C. Wang, B. Yan, L. Lyu, H. E. Takiff, and Q. Gao. The potassium transporter KdpA affects persister formation by regulating ATP levels in *Mycobacterium marinum*. **Emerging Microbes and Infections**, 9(1):129–139, 2020.

[154] M. A. Lobas, R. Tao, J. Nagai, M. T. Kronschläger, P. M. Borden, J. S. Marvin, L. L. Looger, and B. S. Khakh. A genetically encoded single-wavelength sensor for imaging cytosolic and cell surface ATP. **Nature Communications**, 10(1):711, 2019.

[155] A. Loinger, A. Lipshtat, N. Q. Balaban, and O. Biham. Stochastic simulations of genetic switch systems. **Physical Review E**, 75(2):021904, 2007.

[156] O. Lomovskaya, M. S. Warren, A. Lee, J. Galazzo, R. Fronko, M. Lee, J. Blais, D. Cho, S. Chamberland, T. Renau, *et al.* Identification and characterization of inhibitors of multidrug resistance efflux pumps in *Pseudomonas aeruginosa*: novel agents for combination therapy. **Antimicrobial Agents and Chemotherapy**, 45(1):105–116, 2001.

[157] C. Lou, Z. Li, and Q. Ouyang. A molecular model for persister in *E. coli*. **Journal of Theoretical Biology**, 255(2):205–209, 2008.

[158] M. Lovmar, K. Nilsson, E. Lukk, V. Vimberg, T. Tenson, and M. Ehrenberg. Erythromycin resistance by L4/L22 mutations and resistance masking by drug efflux pump deficiency. **The EMBO Journal**, 28(6):736–744, 2009.

[159] M. Lovmar, K. Nilsson, V. Vimberg, T. Tenson, M. Nervall, and M. Ehrenberg. The molecular mechanism of peptide-mediated erythromycin resistance. **Journal of Biological Chemistry**, 281(10):6742–6750, 2006.

[160] L. Lv, M. Wan, C. Wang, X. Gao, Q. Yang, S. R. Partridge, Y. Wang, Z. Zong, Y. Doi, J. Shen, *et al.* Emergence of a plasmid-encoded resistance-nodulation-division efflux pump conferring resistance to multiple drugs, including tigecycline, in *Klebsiella pneumoniae*. **mBio**, 11(2), 2020.

[161] M. Lynch and G. K. Marinov. The bioenergetic costs of a gene. **Proceedings of the National Academy of Sciences**, 112(51):15690–15695, 2015.

[162] D. Ma, M. Alberti, C. Lynch, H. Nikaido, and J. E. Hearst. The local repressor AcrR plays a modulating role in the regulation of *acrAB* genes of *Escherichia coli* by global stress signals. **Molecular Microbiology**, 19(1):101–112, 1996.

[163] D. Ma, D. N. Cook, M. Alberti, N. G. Pon, H. Nikaido, and J. E. Hearst. Genes *acrA* and *acrB* encode a stress-induced efflux system of *Escherichia coli*. **Molecular Microbiology**, 16(1):45–55, 1995.

[164] A.-P. Magiorakos, A. Srinivasan, R. B. Carey, Y. Carmeli, M. E. Falagas, C. G. Giske, S. Harbarth, J. F. Hindler, G. Kahlmeter, B. Olsson-Liljequist, *et al.* Multidrug-resistant, extensively drug-resistant and pandrug-resistant bacteria: an international expert proposal for interim standard definitions for acquired resistance. **Clinical Microbiology and Infection**, 18(3):268–281, 2012.

[165] R. S. Mahla. Stem cells applications in regenerative medicine and disease therapeutics. **International Journal of Cell Biology**, 2016, 2016.

[166] S. Manuse, Y. Shan, S. J. Canas-Duarte, S. Bakshi, W.-S. Sun, H. Mori, J. Paulsson, and K. Lewis. Bacterial persisters are a stochastically formed subpopulation of low-energy cells. **PLoS Biology**, 19(4):1–18, 04 2021.

[167] L. L. Marcusson, N. Frimodt-Møller, and D. Hughes. Interplay in the selection of fluoroquinolone resistance and bacterial fitness. **PLoS Pathogens**, 5(8):e1000541, 2009.

[168] P. N. Markham, E. Westhaus, K. Klyachko, M. E. Johnson, and A. A. Neyfakh. Multiple novel inhibitors of the Nora multidrug transporter of *Staphylococcus aureus*. **Antimicrobial Agents and Chemotherapy**, 43(10):2404–2408, 1999.

[169] R. G. Martin, K.-W. Jair, R. Wolf, and J. L. Rosner. Autoactivation of the *marRAB* multiple antibiotic resistance operon by the MarA transcriptional activator in *Escherichia coli*. **Journal of Bacteriology**, 178(8):2216–2223, 1996.

[170] R. G. Martin and J. L. Rosner. Binding of purified multiple antibiotic-resistance repressor protein (MarR) to *mar* operator sequences. **Proceedings of the National Academy of Sciences**, 92(12):5456–5460, 1995.

[171] R. G. Martin and J. L. Rosner. Genomics of the *mara/soxs/rob* regulon of *Escherichia coli*: identification of directly activated promoters by application of molecular genetics and informatics to microarray data. **Molecular Microbiology**, 44(6):1611–1624, 2002.

[172] N. Massjouni, C. G. Rivera, and T. Murali. Virgo: computational prediction of gene functions. **Nucleic Acids Research**, 34(suppl$_2$):W340–W344, 2006.

[173] A. Mazzariol, Y. Tokue, T. M. Kanegawa, G. Cornaglia, and H. Nikaido. High-level fluoroquinolone-resistant clinical isolates of *Escherichia coli* overproduce multidrug efflux protein AcrA. **Antimicrobial Agents and Chemotherapy**, 44(12):3441–3443, 2000.

[174] A. Mazzariol, J. Zuliani, G. Cornaglia, G. M. Rossolini, and R. Fontana. AcrAB efflux system: expression and contribution to fluoroquinolone resistance in *Klebsiella* spp. **Antimicrobial Agents and Chemotherapy**, 46(12):3984–3986, 2002.

[175] L. M. McMurry, A. George, and S. Levy. Active efflux of chloramphenicol in susceptible *Escherichia coli* strains and in multiple-antibiotic-resistant (*mar*) mutants. **Antimicrobial Agents and Chemotherapy**, 38(3):542–546, 1994.

[176] H. R. Meredith, A. J. Lopatkin, D. J. Anderson, and L. You. Bacterial temporal dynamics enable optimal design of antibiotic treatment. **PLoS Computational Biology**, 11(4), 2015.

[177] H. Mi, A. Muruganujan, X. Huang, D. Ebert, C. Mills, X. Guo, and P. D. Thomas. Protocol update for large-scale genome and gene function analysis with the panther classification system (v. 14.0). **Nature Protocols**, 14(3):703–721, 2019.

[178] P. F. Miller, L. F. Gambino, M. C. Sulavik, and S. J. Gracheck. Genetic relationship between *soxRS* and *mar* loci in promoting multiple antibiotic resistance in *Escherichia coli*. **Antimicrobial Agents and Chemotherapy**, 38(8):1773–1779, 1994.

[179] S.-H. Min, A. R. French, K. J. Trull, K. Tat, S. A. Varney, and M. Tantama. Ratiometric BRET measurements of ATP with a genetically-encoded luminescent sensor. **Sensors**, 19(16):3502, 2019.

[180] A. Mitchell, G. H. Romano, B. Groisman, A. Yona, E. Dekel, M. Kupiec, O. Dahan, and Y. Pilpel. Adaptive prediction of environmental changes by microorganisms. **Nature**, 460(7252):220, 2009.

[181] J. Mitchell, I. G. Johnston, and G. W. Bassel. Variability in seeds: biological, ecological, and agricultural implications. **Journal of Experimental Botany**, 68(4):809–817, 2016.

[182] L. R. Mulcahy, J. L. Burns, S. Lory, and K. Lewis. Emergence of *Pseudomonas aeruginosa* strains producing high levels of persister cells in patients with cystic fibrosis. **Journal of Bacteriology**, 192(23):6191–6199, 2010.

[183] H. Muramatsu, T. Horii, A. Takeshita, H. Hashimoto, and M. Maekawa. Characterization of fluoroquinolone and carbapenem susceptibilities in clinical isolates of levofloxacin-resistant *Pseudomonas aeruginosa*. **Chemotherapy**, 51(2-3):70–75, 2005.

[184] D. L. Nabb, S. Song, K. E. Kluthe, T. A. Daubert, B. E. Luedtke, and A. Nuxoll. Polymicrobial interactions induce multidrug tolerance in *Staphylococcus aureus* through energy depletion. **Frontiers in Microbiology**, 10:2803, 2019.

[185] M. Nakano, H. Imamura, T. Nagai, and H. Noji. Ca^{2+} regulation of mitochondrial ATP synthesis visualized at the single cell level. **ACS Chemical Biology**, 6(7):709–715, 2011.

[186] A. Naldi, J. Carneiro, C. Chaouiya, and D. Thieffry. Diversity and plasticity of Th cell types predicted from regulatory network modelling. **PLoS Computational Biology**, 6(9), 2010.

[187] M. L. Nelson and S. B. Levy. Reversal of tetracycline resistance mediated by different bacterial tetracycline resistance determinants by an inhibitor of the Tet(B) antiport protein. **Antimicrobial Agents and Chemotherapy**, 43(7):1719–1724, 1999.

[188] J. R. Newman, S. Ghaemmaghami, J. Ihmels, D. K. Breslow, M. Noble, J. L. DeRisi, and J. S. Weissman. Single-cell proteomic analysis of *S. cerevisiae* reveals the architecture of biological noise. **Nature**, 441(7095):840, 2006.

[189] A. A. Neyfakh, C. Borsch, and G. Kaatz. Fluoroquinolone resistance protein NorA of *Staphylococcus aureus* is a multidrug efflux transporter. **Antimicrobial Agents and Chemotherapy**, 37(1):128–129, 1993.

[190] W. W. Nichols. Modeling the kinetics of the permeation of antibacterial agents into growing bacteria and its interplay with efflux. **Antimicrobial Agents and Chemotherapy**, 61(10):e02576–16, 2017.

[191] H. Nicoloff, V. Perreten, L. M. McMurry, and S. B. Levy. Role for tandem duplication and Lon protease in AcrAB-TolC-dependent multiple antibiotic resistance (Mar) in an *Escherichia coli* mutant without mutations in *marRAB* or *acrAB*. **Journal of Bacteriology**, 188(12):4413–4423, 2006.

[192] K. M. Nielsen and J. P. Townsend. Monitoring and modeling horizontal gene transfer. **Nature Biotechnology**, 22(9):1110–1114, 2004.

[193] E. Nikaido, E. Giraud, S. Baucheron, S. Yamasaki, A. Wiedemann, K. Okamoto, T. Takagi, A. Yamaguchi, A. Cloeckaert, and K. Nishino. Effects of indole on drug resistance and virulence of *Salmonella enterica* serovar Typhimurium revealed by genome-wide analyses. **Gut Pathogens**, 4(1):1–13, 2012.

[194] E. Nikaido, I. Shirosaka, A. Yamaguchi, and K. Nishino. Regulation of the AcrAB multidrug efflux pump in *Salmonella enterica* serovar Typhimurium in response to indole and paraquat. **Microbiology**, 157(3):648–655, 2011.

[195] E. Nikaido, A. Yamaguchi, and K. Nishino. AcrAB multidrug efflux pump regulation in *Salmonella enterica* serovar Typhimurium by RamA in response to environmental signals. **Journal of Biological Chemistry**, 283(35):24245–24253, 2008.

[196] H. Nikaido. Outer membrane barrier as a mechanism of antimicrobial resistance. **Antimicrobial Agents and Chemotherapy**, 33(11):1831, 1989.

[197] H. Nikaido. Prevention of drug access to bacterial targets: permeability barriers and active efflux. **Science**, 264(5157):382–388, 1994.

[198] H. Nikaido. Multidrug efflux pumps of Gram-negative bacteria. **Journal of Bacteriology**, 178(20):5853, 1996.

[199] H. Nikaido and J.-M. Pagès. Broad-specificity efflux pumps and their role in multidrug resistance of Gram-negative bacteria. **FEMS Microbiology Reviews**, 36(2):340–363, 2012.

[200] T. Nilsson, V. Schultz, P.-O. Berggren, B. E. Corkey, and T. Tornheim. Temporal patterns of changes in ATP/ADP ratio, glucose 6-phosphate and cytoplasmic free Ca^{2+} in glucose-stimulated pancreatic β-cells. **Biochemical Journal**, 314(1):91–94, 1996.

[201] K. Nishino, E. Nikaido, and A. Yamaguchi. Regulation and physiological function of multidrug efflux pumps in *Escherichia coli* and *Salmonella*. **Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics**, 1794(5):834–843, 2009.

[202] A. Nolting and H. Derendorf. Pharmacokinetic/pharmacodynamic modeling of antibiotics. **Handbook of Pharmacokinetic/Pharmacodynamic Correlation.** CRC Press, Boca Raton, Fla, 1995.

[203] P. Nordmann and M. Guibert. Extended-spectrum beta-lactamases in *Pseudomonas aeruginosa*. **The Journal of Antimicrobial Chemotherapy**, 42(2):128–131, 1998.

[204] P. G. Ntagiopoulos, E. Paramythiotou, A. Antoniadou, H. Giamarellou, and A. Karabinis. Impact of an antibiotic restriction policy on the antibiotic resistance patterns of Gram-negative microorganisms in an intensive care unit in Greece. **International Journal of Antimicrobial Agents**, 30(4):360–365, 2007.

[205] S. Okawa, S. Nicklas, S. Zickenrott, J. C. Schwamborn, and A. del Sol. A generalized gene-regulatory network model of stem cell differentiation for predicting lineage specifiers. **Stem Cell Reports**, 7(3):307–315, 2016.

[206] H. Okusu, D. Ma, and H. Nikaido. AcrAB efflux pump plays a major role in the antibiotic resistance phenotype of *Escherichia coli* multiple-antibiotic-resistance (Mar) mutants. **Journal of Bacteriology**, 178(1):306–308, 1996.

[207] A. Olliver, M. Vallé, E. Chaslus-Dancla, and A. Cloeckaert. Role of an *acrR* mutation in multidrug resistance of in vitro-selected fluoroquinolone-resistant mutants of *Salmonella enterica* serovar Typhimurium. **FEMS Microbiology Letters**, 238(1):267–272, 2004.

[208] J. O'Neill. Tackling drug-resistant infections globally: final report and recommendations. **Government of the United Kingdom**, 2016.

[209] J. O'neill. Antimicrobial resistance: tackling a crisis for the health and wealth of nations. **Review on Antimicrobial Resistance**, 20:1–16, 2014.

[210] W. H. O. *et al.* Global action plan on antimicrobial resistance. 2015.

[211] D. I. Orozco-Gómez, J. E. Sosa-Hernández, Ó. A. Gallardo-Navarro, J. Santana-Solano, and M. Santillán. Bistable behaviour and medium-dependent post-translational regulation of the tryptophanase operon regulatory pathway in *Escherichia coli*. **Scientific Reports**, 9(1):5451, 2019.

[212] E. M. Ozbudak, M. Thattai, I. Kurtser, A. D. Grossman, and A. Van Oudenaarden. Regulation of noise in the expression of a single gene. **Nature Genetics**, 31(1):69, 2002.

[213] L. Paixão, L. Rodrigues, I. Couto, M. Martins, P. Fernandes, C. C. De Carvalho, G. A. Monteiro, F. Sansonetty, L. Amaral, and M. Viveiros. Fluorometric determination of ethidium bromide efflux kinetics in *Escherichia coli*. **Journal of Biological Engineering**, 3(1):18, 2009.

[214] I.-H. Park, R. Zhao, J. A. West, A. Yabuuchi, H. Huo, T. A. Ince, P. H. Lerou, M. W. Lensch, and G. Q. Daley. Reprogramming of human somatic cells to pluripotency with defined factors. **Nature**, 451(7175):141, 2008.

[215] J. Paulsson. Summing up the noise in gene networks. **Nature**, 427(6973):415, 2004.

[216] S. Paytubi, S. Aznar, C. Madrid, C. Balsalobre, S. C. Dillon, C. J. Dorman, and A. Juárez. A novel role for antibiotic resistance plasmids in facilitating *Salmonella* adaptation to non-host environments. **Environmental Microbiology**, 16(4):950–962, 2014.

[217] J. M. Pedraza and A. van Oudenaarden. Noise propagation in gene networks. **Science**, 307(5717):1965–1969, 2005.

[218] C. Peña, G. Cabot, S. Gómez-Zorrilla, L. Zamorano, A. Ocampo-Sosa, J. Murillas, B. Almirante, V. Pomar, M. Aguilar, A. Granados, *et al.* Influence of virulence genotype and resistance profile in the mortality of *Pseudomonas aeruginosa* bloodstream infections. **Clinical Infectious Diseases**, 60(4):539–548, 2015.

[219] R. Pereira da Cruz, T. Sampaio de Freitas, S. Costa, L. dos Santos, A. Thassya, F. Ferreira Campina, R. L. S. Pereira, J. W. A. Bezerra, L. J. Quintans-Júnior, A. A. De Souza Araújo, *et al.* Effect of α-bisabolol and its β-cyclodextrin complex as TetK and NorA efflux pump inhibitors in *Staphylococcus aureus* strains. **Antibiotics**, 9(1):28, 2020.

[220] A. M. Perez, M. M. Gomez, P. Kalvapalle, E. O'Brien-Gilbert, M. R. Bennett, and Y. Shamoo. Using cellular fitness to map the structure and function of a major facilitator superfamily effluxer. **Molecular Systems Biology**, 13(12):964, 2017.

[221] R. Perez-Carrasco, P. Guerrero, J. Briscoe, and K. M. Page. Intrinsic noise profoundly alters the dynamics and steady state of morphogen-controlled bistable genetic switches. **PLoS Computational Biology**, 12(10):e1005154, 2016.

[222] L. G. Phillips, J. P. Heggers, M. C. Robson, J. A. Boertman, T. Meltzer, and D. J. Smith Jr. The effect of endogenous skin bacteria on burn wound infection. **Annals of Plastic Surgery**, 23(1):35–38, 1989.

[223] Z. Phillips and M. Strauch. *Bacillus subtilis* sporulation and stationary phase gene expression. **Cellular and Molecular Life Sciences**, 59(3):392–402, 2002.

[224] L. J. Piddock. Clinically relevant chromosomally encoded multidrug resistance efflux pumps in bacteria. **Clinical Microbiology Reviews**, 19(2):382–402, 2006.

[225] L. J. Piddock, D. G. White, K. Gensberg, L. Pumbwe, and D. J. Griggs. Evidence for an efflux pump mediating multiple antibiotic resistance in *Salmonella enterica* serovar Typhimurium. **Antimicrobial Agents and Chemotherapy**, 44(11):3118–3121, 2000.

[226] K. Poole. Efflux-mediated multiresistance in Gram-negative bacteria. **Clinical Microbiology and Infection**, 10(1):12–26, 2004.

[227] K. Poole. Efflux pumps as antimicrobial resistance mechanisms. **Annals of Medicine**, 39(3):162–176, 2007.

[228] M. Ptashne. **A Genetic Switch: Phage Lambda Revisited**, volume 3. Cold Spring Harbor Laboratory Press, 2004.

[229] Y. Pu, Z. Zhao, Y. Li, J. Zou, Q. Ma, Y. Zhao, Y. Ke, Y. Zhu, H. Chen, M. A. Baker, *et al.* Enhanced efflux activity facilitates drug tolerance in dormant bacterial cells. **Molecular Cell**, 62(2):284–294, 2016.

[230] B. Pugashetti, H. Metzger, L. Vadas, and D. Feingold. Phenotypic differences among clinically isolated mucoid *Pseudomonas aeruginosa* strains. **Journal of Clinical Microbiology**, 16(4):686–691, 1982.

[231] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. **Proceedings of the IEEE**, 77(2):257–286, 1989.

[232] M. Raeispour and R. Ranjbar. Antibiotic resistance, virulence factors and genotyping of uropathogenic *Escherichia coli* strains. **Antimicrobial Resistance and Infection Control**, 7(1):1–9, 2018.

[233] A. A. Ramadan, N. A. Abdelaziz, M. A. Amin, and R. K. Aziz. Novel bla ctx-m variants and genotype-phenotype correlations among clinical isolates of extended spectrum beta lactamase-producing escherichia coli. *Scientific reports*, 9(1):1–12, 2019.

[234] V. E. Rees, D. S. Deveson Lucas, C. López-Causapé, Y. Huang, T. Kotsimbos, J. B. Bulitta, M. C. Rees, A. Barugahare, A. Y. Peleg, R. L. Nation, *et al.* Characterization of hypermutator *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis in Australia. **Antimicrobial Agents and Chemotherapy**, 63(4):e02538–18, 2019.

[235] T.-B. Ren, S. Wen, L. Wang, P. Lu, B. Xiong, L. Yuan, and X.-B. Zhang. Engineering a reversible fluorescent probe for real-time live-cell imaging and quantification of mitochondrial ATP. **Analytical Chemistry**, 2020.

[236] V. Ricci and L. J. Piddock. Ciprofloxacin selects for multidrug resistance in *Salmonella enterica* serovar Typhimurium mediated by at least two different pathways. **Journal of Antimicrobial Chemotherapy**, 63(5):909–916, 2009.

[237] V. Ricci, P. Tzakas, A. Buckley, N. C. Coldham, and L. J. Piddock. Ciprofloxacin-resistant *Salmonella enterica* serovar Typhimurium strains are difficult to select in the absence of AcrB and TolC. **Antimicrobial Agents and Chemotherapy**, 50(1):38–42, 2006.

[238] P. A. Roberts, R. M. Huebinger, E. Keen, A.-M. Krachler, and S. Jabbari. Predictive modelling of a novel anti-adhesion therapy to combat bacterial colonisation of burn wounds. **PLoS Computational Biology**, 14(5), 2018.

[239] P. A. Roberts, R. M. Huebinger, E. Keen, A.-M. Krachler, and S. Jabbari. Mathematical model predicts anti-adhesion–antibiotic–debridement combination therapies can clear an antibiotic resistant infection. **PLoS Computational Biology**, 15(7):e1007211, 2019.

[240] G. Rodrigo, D. Bajic, I. Elola, and J. F. Poyatos. Antagonistic autoregulation speeds up a homogeneous response in *Escherichia coli*. **Scientific Reports**, 6:36196, 2016.

[241] G. Rodrigo, D. Bajić, I. Elola, and J. F. Poyatos. Deconstructing a multiple antibiotic resistance regulation through the quantification of its input function. **NPJ Systems Biology and Applications**, 3(1):1–9, 2017.

[242] M. Roghanian, S. Semsey, A. Løbner-Olesen, and F. Jalalvand. (p)ppGpp-mediated stress response induced by defects in outer membrane biogenesis and ATP production promotes survival in *Escherichia coli*. **Scientific Reports**, 9(1):1–11, 2019.

[243] V. D. Rosenthal, D. G. Maki, A. Mehta, C. Álvarez-Moreno, H. Leblebicioglu, F. Higuera, L. E. Cuellar, N. Madani, Z. Mitrev, L. Dueñas, *et al.* International nosocomial infection control consortium report, data summary for 2002-2007, issued January 2008. **American Journal of Infection Control**, 36(9):627–637, 2008.

[244] E. Rotem, A. Loinger, I. Ronin, I. Levin-Reisman, C. Gabay, N. Shoresh, O. Biham, and N. Q. Balaban. Regulation of phenotypic variability by a threshold-based mechanism underlies bacterial persistence. **Proceedings of the National Academy of Sciences**, 107(28):12541–12546, 2010.

[245] A. Saadatpour and R. Albert. Boolean modeling of biological regulatory networks: a methodology tutorial. **Methods**, 62(1):3–12, 2013.

[246] L. Sanchez, W. Pan, M. Vinas, and H. Nikaido. The *acrAB* homolog of *Haemophilus influenzae* codes for a functional multidrug efflux pump. **Journal of Bacteriology**, 179(21):6855–6857, 1997.

[247] M. A. Sánchez-Romero and J. Casadesús. Contribution of phenotypic heterogeneity to adaptive antibiotic resistance. **Proceedings of the National Academy of Sciences**, 111(1):355–360, 2014.

[248] S. M. Schieke, M. Ma, L. Cao, J. P. McCoy, C. Liu, N. F. Hensel, A. J. Barrett, M. Boehm, and T. Finkel. Mitochondrial metabolism modulates differentiation and teratoma formation capacity in mouse embryonic stem cells. **Journal of Biological Chemistry**, 283(42):28506–28512, 2008.

[249] S. Schwarz, J. Shen, K. Kadlec, Y. Wang, G. B. Michael, A. T. Feßler, and B. Vester. Lincosamides, streptogramins, phenicols, and pleuromutilins: mode of action and mechanisms of resistance. **Cold Spring Harbor Perspectives in Medicine**, 6(11):a027037, 2016.

[250] A. S. Seoane and S. B. Levy. Characterization of MarR, the repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*. **Journal of Bacteriology**, 177(12):3414–3419, 1995.

[251] Y. Shan, A. B. Gandt, S. E. Rowe, J. P. Deisinger, B. P. Conlon, and K. Lewis. ATP-dependent persister formation in *Escherichia coli*. **MBio**, 8(1):e02267–16, 2017.

[252] P. Sharma, J. R. Haycocks, A. D. Middlemiss, R. A. Kettles, L. E. Sellars, V. Ricci, L. J. Piddock, and D. C. Grainger. The multiple antibiotic resistance operon of enteric bacteria controls DNA repair and outer membrane integrity. **Nature Communications**, 8(1):1–12, 2017.

[253] M. A. Sid Ahmed, H. Abdel Hadi, S. Abu Jarir, A. L. Al Khal, M. A. Al-Maslamani, J. Jass, E. B. Ibrahim, and H. Ziglam. Impact of an antimicrobial stewardship programme on antimicrobial utilization and the prevalence of MDR *Pseudomonas aeruginosa* in an acute care hospital in Qatar. **JAC-Antimicrobial Resistance**, 2(3):dlaa050, 2020.

[254] S. O. Skinner, L. A. Sepúlveda, H. Xu, and I. Golding. Measuring mRNA copy number in individual *Escherichia coli* cells using single-molecule fluorescent *in situ* hybridization. **Nature Protocols**, 8(6):1100, 2013.

[255] W. K. Smits, O. P. Kuipers, and J.-W. Veening. Phenotypic variation in bacteria: the role of feedback regulation. **Nature Reviews Microbiology**, 4(4):259, 2006.

[256] M. L. Sobel, G. A. McKay, and K. Poole. Contribution of the MexXY multidrug transporter to aminoglycoside resistance in *Pseudomonas aeruginosa* clinical isolates. **Antimicrobial Agents and Chemotherapy**, 47(10):3202–3207, 2003.

[257] S. Song and T. K. Wood. ppGpp ribosome dimerization model for bacterial persister formation and resuscitation. **Biochemical and Biophysical Research Communications**, 2020.

[258] M. Souli, I. Galani, and H. Giamarellou. Emergence of extensively drug-resistant and pandrug-resistant Gram-negative bacilli in Europe. **Eurosurveillance**, 13(47):19045, 2008.

[259] F. M. Stewart and B. R. Levin. The population biology of bacterial plasmids: a priori conditions for the existence of conjugationally transmitted factors. **Genetics**, 87(2):209–228, 1977.

[260] M. Stitt, R. M. Lilley, and H. W. Heldt. Adenine nucleotide levels in the cytosol, chloroplasts, and mitochondria of wheat leaf protoplasts. **Plant Physiology**, 70(4):971–977, 1982.

[261] C.-C. Su, D. J. Rutherford, and W. Y. Edward. Characterization of the multidrug efflux regulator AcrR from *Escherichia coli*. **Biochemical and Biophysical Research Communications**, 361(1):85–90, 2007.

[262] G. M. Süel, J. Garcia-Ojalvo, L. M. Liberman, and M. B. Elowitz. An excitable gene regulatory circuit induces transient cellular differentiation. **Nature**, 440(7083):545, 2006.

[263] M. C. Sulavik, L. F. Gambino, and P. F. Miller. The MarR repressor of the multiple antibiotic resistance (*mar*) operon in *Escherichia coli*: prototypic member of a family of bacterial regulatory proteins involved in sensing phenolic compounds. **Molecular Medicine**, 1(4):436–446, 1995.

[264] J. Sun, Z. Deng, and A. Yan. Bacterial multidrug efflux pumps: mechanisms, physiology and pharmacological exploitations. **Biochemical and Biophysical Research Communications**, 453(2):254–267, 2014.

[265] N. S. Sundaramoorthy, P. Suresh, S. S. Ganesan, A. GaneshPrasad, and S. Nagarajan. Restoring colistin sensitivity in colistin-resistant *E. coli*: combinatorial use of MarR inhibitor with efflux pump inhibitor. **Scientific Reports**, 9(1):1–13, 2019.

[266] F. S. Taccone, F. Cotton, S. Roisin, J.-L. Vincent, and F. Jacobs. Optimal meropenem concentrations to treat multidrug-resistant *Pseudomonas aeruginosa* septic shock. **Antimicrobial Agents and Chemotherapy**, 56(4):2129–2131, 2012.

[267] E. Tacconelli, E. Carrara, A. Savoldi, S. Harbarth, M. Mendelson, D. L. Monnet, C. Pulcini, G. Kahlmeter, J. Kluytmans, Y. Carmeli, *et al.* Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. **The Lancet Infectious Diseases**, 18(3):318–327, 2018.

[268] K. Takahashi, K. Tanabe, M. Ohnuki, M. Narita, T. Ichisaka, K. Tomoda, and S. Yamanaka. Induction of pluripotent stem cells from adult human fibroblasts by defined factors. **Cell**, 131(5):861–872, 2007.

[269] M. Takaine. QUEEN-based spatiotemporal ATP imaging in budding and fission yeast. **Bio-protocol**, 9:e3320, 2019.

[270] M. Takaine, M. Ueno, K. Kitamura, H. Imamura, and S. Yoshida. Reliable imaging of ATP in living budding and fission yeast. **Journal of Cell Science**, 2019.

[271] V. H. Tam, A. N. Schilling, K. Poole, and M. Nikolaou. Mathematical modelling response of *Pseudomonas aeruginosa* to meropenem. **Journal of Antimicrobial Chemotherapy**, 60(6):1302–1309, 2007.

[272] L. Ternent, R. J. Dyson, A.-M. Krachler, and S. Jabbari. Bacterial fitness shapes the population dynamics of antibiotic-resistant and-susceptible bacteria in a model of combined antibiotic and anti-virulence treatment. **Journal of Theoretical Biology**, 372:1–11, 2015.

[273] T. Teslaa and M. A. Teitell. Pluripotent stem cell energy metabolism: an update. **The EMBO Journal**, 34(2):138–153, 2015.

[274] M. Thattai and A. Van Oudenaarden. Intrinsic noise in gene regulatory networks. **Proceedings of the National Academy of Sciences**, 98(15):8614–8619, 2001.

[275] R. Thomas. Boolean formalization of genetic control circuits. **Journal of Theoretical Biology**, 42(3):563–585, 1973.

[276] S. S. Thota and L. M. Chubiz. Multidrug resistance regulators MarA, SoxS, Rob, and RamA repress flagellar gene expression and motility in *Salmonella enterica* serovar Typhimurium. **Journal of Bacteriology**, pages JB–00385, 2019.

[277] E. B. Tikhonova and H. I. Zgurskaya. AcrA, AcrB, and TolC of *Escherichia coli* form a stable intermembrane multidrug efflux complex. **Journal of Biological Chemistry**, 279(31):32116–32124, 2004.

[278] T. Toni, D. Welch, N. Strelkowa, A. Ipsen, and M. P. Stumpf. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. **Journal of the Royal Society Interface**, 6(31):187–202, 2009.

[279] A. T. Topham, R. E. Taylor, D. Yan, E. Nambara, I. G. Johnston, and G. W. Bassel. Temperature variability is integrated by a spatially embedded decision-making center to break dormancy in *Arabidopsis* seeds. **Proceedings of the National Academy of Sciences**, 114(25):6629–6634, 2017.

[280] T. Tsuyama, J.-i. Kishikawa, Y.-W. Han, Y. Harada, A. Tsubouchi, H. Noji, A. Kakizuka, K. Yokoyama, T. Uemura, and H. Imamura. In vivo fluorescent adenosine

5'-triphosphate (ATP) imaging of *Drosophila melanogaster* and *Caenorhabditis elegans* by using a genetically encoded fluorescent ATP biosensor optimized for low temperatures. **Analytical Chemistry**, 85(16):7889–7896, 2013.

[281] J. J. Tyson, K. C. Chen, and B. Novak. Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. **Current Opinion in Cell Biology**, 15(2):221–231, 2003.

[282] M. Usui, H. Nagai, M. Hiki, Y. Tamura, and T. Asai. Effect of antimicrobial exposure on AcrAB expression in *Salmonella enterica* subspecies *enterica* serovar Choleraesuis. **Frontiers in Microbiology**, 4:53, 2013.

[283] T. van der Straaten, R. Janssen, D. J. Mevius, and J. T. van Dissel. *Salmonella* gene *rma* (*ramA*) and multiple-drug-resistant *Salmonella enterica* serovar Typhimurium. **Antimicrobial Agents and Chemotherapy**, 48(6):2292–2294, 2004.

[284] N. Vázquez-Laslop, H. Lee, and A. A. Neyfakh. Increased persistence in *Escherichia coli* caused by controlled expression of toxins or other unrelated proteins. **Journal of Bacteriology**, 188(10):3494–3497, 2006.

[285] W. L. Veal, R. A. Nicholas, and W. M. Shafer. Overexpression of the MtrC-MtrD-MtrE efflux pump due to an *mtrR* mutation is required for chromosomally mediated penicillin resistance in *Neisseria gonorrhoeae*. **Journal of Bacteriology**, 184(20):5619–5624, 2002.

[286] J.-W. Veening, L. W. Hamoen, and O. P. Kuipers. Phosphatases modulate the bistable sporulation gene expression pattern in *Bacillus subtilis*. **Molecular Microbiology**, 56(6):1481–1494, 2005.

[287] C. L. Ventola. The antibiotic resistance crisis: part 1: causes and threats. **Pharmacy and Therapeutics**, 40(4):277, 2015.

[288] B. Verd, A. Crombach, and J. Jaeger. Classification of transient behaviours in a time-dependent toggle switch model. **BMC Systems Biology**, 8(1):43, 2014.

[289] J. Vergalli, I. V. Bodrenko, M. Masi, L. Moynié, S. Acosta-Gutiérrez, J. H. Naismith, A. Davin-Regli, M. Ceccarelli, B. van den Berg, M. Winterhalter, *et al.* Porins and small-molecule translocation across the outer membrane of Gram-negative bacteria. **Nature Reviews Microbiology**, pages 1–13, 2019.

[290] L. Vinué, L. M. McMurry, and S. B. Levy. The 216-bp *marB* gene of the *marRAB* operon in *Escherichia coli* encodes a periplasmic protein which reduces the transcription rate of *marA*. **FEMS Microbiology Letters**, 345(1):49–55, 2013.

[291] G. Von Dassow and G. M. Odell. Design and constraints of the *Drosophila* segment polarity module: robust spatial patterning emerges from intertwined cell state switches. **Journal of Experimental Zoology**, 294(3):179–215, 2002.

[292] C. H. Waddington. **The Strategy of the Genes**. Allen and Unwin, London, 1957.

[293] A. Walkty, M. DeCorby, K. Nichol, M. R. Mulvey, D. Hoban, G. Zhanel, Canadian Antimicrobial Resistance Alliance, *et al.* Antimicrobial susceptibility of *Pseudomonas aeruginosa* isolates obtained from patients in Canadian intensive care units as part of the Canadian National Intensive Care Unit study. **Diagnostic Microbiology and Infectious Disease**, 61(2):217–221, 2008.

[294] A. Walkty, P. Lagace-Wiens, H. Adam, M. Baxter, J. Karlowsky, M. R. Mulvey, M. McCracken, and G. G. Zhanel. Antimicrobial susceptibility of 2906 *Pseudomonas aeruginosa* clinical isolates obtained from patients in Canadian hospitals over a period of 8 years: results of the Canadian Ward surveillance study (CANWARD), 2008–2015. **Diagnostic Microbiology and Infectious Disease**, 87(1):60–63, 2017.

[295] C. Walsh *et al.* **Antibiotics: Actions, Origins, Resistance**. American Society for Microbiology, 2003.

[296] Y. Wan, M. Wang, E. W. C. Chan, and S. Chen. Membrane transporters of the major facilitator superfamily are essential for long-term maintenance of phenotypic tolerance to multiple antibiotics in *E. coli*. **Microbiology Spectrum**, 9(3):e01846–21, 2021.

[297] R.-S. Wang, A. Saadatpour, and R. Albert. Boolean modeling in systems biology: an overview of methodology and applications. **Physical Biology**, 9(5):055001, 2012.

[298] Y. Wang, M. S. Bojer, S. E. George, Z. Wang, P. R. Jensen, C. Wolz, and H. Ingmer. Inactivation of TCA cycle enhances *Staphylococcus aureus* persister cell formation in stationary phase. **Scientific Reports**, 8(1):1–13, 2018.

[299] P. B. Warren and P. R. ten Wolde. Chemical models of genetic toggle switches. **The Journal of Physical Chemistry B**, 109(14):6812–6823, 2005.

[300] L. Wasserman. **All of Statistics: A Concise Course in Statistical Inference**, volume 26. Springer, 2004.

[301] M. A. Webber and L. J. Piddock. Absence of mutations in *marRAB* or *soxRS* in *acrB*-overexpressing fluoroquinolone-resistant clinical and veterinary isolates of *Escherichia coli*. **Antimicrobial Agents and Chemotherapy**, 45(5):1550–1552, 2001.

[302] M. A. Webber, A. Talukder, and L. J. Piddock. Contribution of mutation at amino acid 45 of AcrR to *acrB* expression and ciprofloxacin resistance in clinical and veterinary *Escherichia coli* isolates. **Antimicrobial Agents and Chemotherapy**, 49(10):4390–4392, 2005.

[303] B. Weisblum. Inducible resistance to macrolides, lincosamides and streptogramin type B antibiotics: the resistance phenotype, its biological diversity, and structural elements that regulate expression–a review. **Journal of Antimicrobial Chemotherapy**, 16(suppl$_A$):63–90, 1985.

[304] M. H. Wells-Bennik, R. T. Eijlander, H. M. Den Besten, E. M. Berendsen, A. K. Warda, A. O. Krawczyk, M. N. Nierop Groot, Y. Xiao, M. H. Zwietering, O. P. Kuipers, *et al.* Bacterial spores in food: survival, emergence, and outgrowth. **Annual Review of Food Science and Technology**, 7:457–482, 2016.

[305] X. Wen, A. M. Langevin, and M. J. Dunlop. Antibiotic export by efflux pumps affects growth of neighboring bacteria. **Scientific Reports**, 8(1):1–9, 2018.

[306] M. F. Wernet, E. O. Mazzoni, A. Çelik, D. M. Duncan, I. Duncan, and C. Desplan. Stochastic spineless expression creates the retinal mosaic for colour vision. **Nature**, 440(7081):174, 2006.

[307] B. P. Williams, I. G. Johnston, S. Covshoff, and J. M. Hibberd. Phenotypic landscape inference reveals multiple evolutionary paths to C4 photosynthesis. **eLife**, 2:e00961, 2013.

[308] Y. Wu, M. Vulić, I. Keren, and K. Lewis. Role of oxidative stress in persister tolerance. **Antimicrobial Agents and Chemotherapy**, 56(9):4922–4926, 2012.

[309] H. Yaginuma, S. Kawai, K. V. Tabata, K. Tomiyama, A. Kakizuka, T. Komatsuzaki, H. Noji, and H. Imamura. Diversity in ATP concentrations in a single bacterial cell population revealed by quantitative single-cell imaging. **Scientific Reports**, 4:6522, 2014.

[310] J. Yamada, S. Yamasaki, H. Hirakawa, M. Hayashi-Nishino, A. Yamaguchi, and K. Nishino. Impact of the RNA chaperone Hfq on multidrug resistance in *Escherichia coli*. **Journal of Antimicrobial Chemotherapy**, 65(5):853–858, 2010.

[311] A. Yamagishi, S. Nakano, S. Yamasaki, and K. Nishino. An efflux inhibitor of the MacAB pump in *Salmonella enterica* serovar Typhimurium. **Microbiology and Immunology**, 2019.

[312] S. Yamasaki, E. Nikaido, R. Nakashima, K. Sakurai, D. Fujiwara, I. Fujii, and K. Nishino. The crystal structure of multidrug-resistance regulator RamR with multiple drugs. **Nature Communications**, 4(1):1–7, 2013.

[313] T. Yoshida, A. Kakizuka, and H. Imamura. BTeam, a novel BRET-based biosensor for the accurate quantification of ATP concentration within living cells. **Scientific Reports**, 6:39618, 2016.

[314] G. H. Youlden, V. Ricci, X. Wang-Kan, L. J. Piddock, S. Jabbari, and J. R. King. Time dependent asymptotic analysis of the gene regulatory network of the AcrAB-TolC efflux pump system in Gram-negative bacteria. **Journal of Mathematical Biology**, 82(4):1–48, 2021.

[315] E. A. Zalis, A. S. Nuxoll, S. Manuse, G. Clair, L. C. Radlinski, B. P. Conlon, J. Adkins, and K. Lewis. Stochastic variation in expression of the tricarboxylic acid cycle produces persister cells. **mBio**, 10(5):e01930–19, 2019.

[316] H. I. Zgurskaya, G. Krishnamoorthy, A. Ntreh, and S. Lu. Mechanism and function of the outer membrane channel TolC in multidrug resistance and physiology of enterobacteria. **Frontiers in Microbiology**, 2:189, 2011.

[317] Y. Zhang, W. W. Yew, and M. R. Barer. Targeting persisters for tuberculosis control. **Antimicrobial Agents and Chemotherapy**, 56(5):2223–2230, 2012.

[318] J. X. Zhou and S. Huang. Understanding gene circuits at cell-fate branch points for rational cell reprogramming. **Trends in Genetics**, 27(2):55–62, 2011.

[319] S. Zimmermann, M. Klinger-Strobel, J. A. Bohnert, M. Pletz, B. Löffler, and L. Tuchscherr. Clinically approved drugs inhibit the *Staphylococcus aureus* multidrug NorA efflux pump and reduce biofilm formation. **Frontiers in Microbiology**, 10:2762, 2019.