Thermodynamically-consistent, reduced models of gene regulatory networks

Michael Pan[1], Peter J. Gawthrop[2], Matthew Faria[2], Stuart T. Johnston[1]*

[1]School of Mathematics and Statistics, The University of Melbourne, Parkville 3010, Victoria, Australia
[2]Department of Biomedical Engineering, The University of Melbourne, Parkville 3010, Victoria, Australia

*Corresponding author: stuart.johnston@unimelb.edu.au

Abstract

Synthetic biology aims to engineer novel functionalities into biological systems. While the approach has to date been predominantly applied to single cells, a richer set of biological phenomena can be engineered by applying synthetic biology to cell populations. To rationally design cell populations, we require mathematical models that link between intracellular biochemistry and intercellular interactions. In this study, we develop a kinetic model of gene expression that is suitable for incorporation into agent-based models of cell populations. To be scalable to large cell populations, models of gene expression should be both computationally efficient and compliant with the laws of physics. We satisfy the first requirement by applying a model reduction scheme to translation, and the second requirement by formulating models using bond graphs. Our reduced model is significantly faster to simulate than the full model, and faithfully reproduces important behaviours of the full model. We couple separate models of gene expression to build models of the toggle switch and repressilator. With these models, we explore the effects of resource availability and cell-to-cell heterogeneity on circuit behaviour. The modelling approaches developed in this study are a bridge towards engineering collective cell behaviours such as synchronisation and division of labour.

1 Introduction

Synthetic biology is the engineering of novel functionalities into biological systems. This enables biological solutions to problems that are difficult or impossible to solve with conventional engineering technologies, such as the detection and breakdown of toxic environmental pollutants and the production of biological therapeutics [1, 2]. Synthetic biology has been applied with great success to single cells, making use of synthetic genes as the fundamental building block. The field first gained attention through early biological implementations of two simple genetic circuits; the genetic *toggle switch*, which is capable of flipping between two stable states [3]; and the *repressilator*, a negative feedback loop composed of three repressor genes capable of oscillations [4]. Since then, synthetic biology has been used to design more complex systems such as biosensors and cell computers [5, 6]. To scale up from simple to complex systems, synthetic biology employs a modular approach from engineering. Thus, standard modules (genetic constructs that have been designed according to specifications) can be re-used as part of a more complex system. With this approach, even relatively simple genetic circuits can be assembled into biological logic gates, and ultimately biological devices capable of making complex decisions [6].

There are several known limitations to single-cell synthetic biology. Firstly, there may be undesired cross-talk between synthetic genes, leading to unpredictable function [7]. Secondly,

the newly inserted genes consume the same resources (energy, ribosomes and amino acids) as other pathways in the cell [8]. Thus, a resource-intensive gene may compete with the cell's endogenous pathways, leading to poor growth [9]. Finally, it cannot be assumed that upstream modules are completely isolated from the effects of downstream modules. In general, downstream pathways will influence and in some cases disrupt the behaviour of upstream pathways. This is known as retroactivity, and can lead to poor modularity even in the absence of cross-talk [10].

To address the known limitations of single-cell synthetic biology, synthetic biology of cell populations and consortia has been proposed [9, 11]. Synthetic modules can be inserted into different cells of a consortium, physically isolating them from each other [5]. This reduces the risk of cross-talk, competition and retroactivity between modules. Furthermore, it is possible to design population-level behaviours that would not be possible within homogeneous populations. For example, the quorum sensing system in bacteria is a frequent target of synthetic biology [12]. This system senses local cell density via the accumulation of signalling molecules, and can be engineered to reduce noise and synchronise oscillations [9]. It is also possible to engineer cell populations to form various spatial patterns when grown on media [13].

In spite of the clear potential of cell populations for synthetic biology, it is still an emerging field compared to single-cell synthetic biology. A major challenge is predicting how changes at the single-cell level affect behaviour at the population scale. Because it is infeasible to physically test every possible implementation of a synthetic biological system, synthetic biologists use design-build-test-learn cycles to optimise performance with limited time and resources [1]. In this context, mathematical models are essential to ruling out infeasible designs prior to implementation. Mathematical models have been widely used to describe cellular biochemistry (systems biology) and interactions between cells (agent-based models) [14]. While it is possible to model cell populations without agent-based models, coupling systems biology models to agent-based models is essential to understanding phenomena such as cell heterogeneity, synchronisation and cell migration. Such population-level dynamics pose a unique set of challenges for modelling engineered cell populations. In particular, cell populations can easily consist of millions of cells [15]. Thus, while complex models of gene regulatory networks are sufficient for single-cell synthetic biology, computational efficiency becomes a major consideration in engineering cell populations [16].

Because synthetic biology is focussed on engineering gene circuits, modelling gene regulatory networks is a key point of interest. The landmark studies implementing the toggle switch and repressilator used simple ordinary differential equation models of gene regulation, ignoring the dynamics of transcription and translation [3, 4]. This approach has been generalised and expanded upon to model large-scale gene regulatory networks [17]. However, the above models do not explicitly model transcription and translation, nor do they account for ribosome and energy availability. More recent models have sought to incorporate this level of detail. Whole-cell models of *Mycoplasma genitalium* and *Escherichia coli* account for the use of ribosomes, amino acids and energy in translation [18, 19]. Weiße et al. [20] take a simpler approach, using generic molecules for energy and monomers as variables in a differential equation model of transcription and translation. There are also more detailed models of translation that account for the position of ribosomes on mRNA and can predict the occurrence of ribosomal 'traffic jams'. These include the stochastic totally asymmetric exclusion process (TASEP) and its deterministic approximation the ribosome flow model (RFM) [21]. However, a consideration for the thermodynamics of gene expression has for the most part been lacking. Many models completely ignore the energy molecule adenosine triphosphate (ATP) and therefore do not account for the dependence of expression dynamics on energy. This is surprising given that translation consumes four ATP molecules per amino acid [22], making it one of the most energy-consuming processes in a cell

(30-75%) [18, 23]. If we are to adequately model gene regulatory networks for the synthetic biology of cell populations, a new modelling approach is needed that accounts for energy while also resulting in models simple enough to be incorporated into large cell populations.

In this study, we propose a new, simplified, and thermodynamically consistent model of gene expression. The model satisfies three key criteria: it is (i) computationally feasible to simulate; (ii) dependent on the availability of energy and ribosomes; and (iii) thermodynamically consistent. We primarily focus on translation as it is a major consumer of energy. A model reduction scheme is developed to simplify the chains of reactions involved in translation elongation. This ensures that the model is simple but biochemically plausible, i.e. it is a limiting case of a more complex reaction network. To ensure thermodynamic consistency, we use the bond graph approach, which is a powerful, modular framework for modelling biological systems in a physically consistent manner [24]. Previous work has shown that thermodynamic constraints automatically fall out from the parameterisation inherent to bond graph models, and that the approach is scalable to large biochemical networks [25–29]. Furthermore, bond graphs provide a natural methodology for accounting for retroactivity [30]. We exploit the graphical, modular nature of bond graphs to couple models of gene expression together, enabling the construction of common synthetic biology motifs such as the toggle switch and repressilator.

In Section 2, we introduce the bond graph formalism, which is then used to construct thermodynamically consistent models of gene regulatory networks. A model reduction scheme for translation is developed to ensure the model is simple but still retains important features of a more detailed model (§ 2.4). By coupling models of gene expression together, models of synthetic circuits are developed (§ 2.6). We first compare the full and reduced models, and show that their steady-state behaviours match well (§ 3.1). We next show that the bond graph model is able to reproduce experimentally observed behaviours in common synthetic biology motifs. In § 3.2–3.3, we investigate the effects of energy and ribosome availability on synthetic circuits. We then examine the potential effects of parameter heterogeneity in synthetic circuits (§ 3.4). Finally, Section 4 contains a discussion of the results and concluding remarks.

2 Methods

2.1 Overview

The gene is a key building block of synthetic biology. Here, we consider two key processes in gene expression: *transcription*, which produces mRNA from a DNA template; and *translation*, which produces protein from an mRNA template (Figure 1). This section will detail the methods required to model gene expression in a thermodynamic framework. Inspired by Weiße et al. [20], we adopt a simplistic approach for modelling these processes. In particular, we use generic species to model cell resources rather than accounting for all their constituent biomolecules. Thus, both energy (A) and ribosomes (R) are represented as individual species. Figure 1 depicts an overview of the model. It is assumed that the reactions take place in a constant volume, i.e. in the absence of cell expansion.

We first introduce the bond graph approach for modelling biochemical systems, and explain how it ensures adherence to the laws of thermodynamics as well as its relation to kinetic modelling (§ 2.2). We then develop bond graph models of transcription (§ 2.3) and translation (§ 2.4). To ensure that the model of translation is computationally tractable, we apply a model reduction scheme (§ 2.4.2). We combine the models of transcription and translation into a model of gene expression (§ 2.5) and detail how models of gene expression can be composed into models of

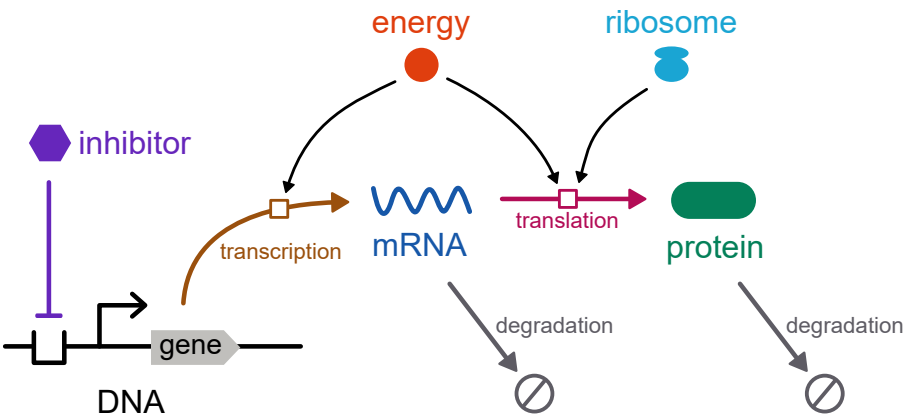

Figure 1: Schematic of gene expression model.

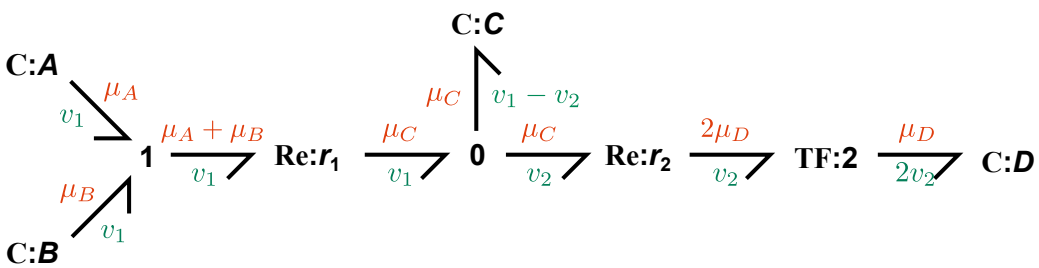

Figure 2: The bond graph of the reaction network $A + B \rightleftharpoons C \rightleftharpoons 2D$**.** Each bond is labelled with a potential μ and a flow v. The components follow the notation *component type*: *name*.

common synthetic biology motifs (§ 2.6).

2.2 Bond graph modelling

The bond graph is a powerful and general methodology for modelling biological systems while adhering to the laws of physics and thermodynamics. We outline the application of bond graphs to biochemistry specifically here; the reader is referred to Gawthrop and Pan [24] for more general applications to biology.

As an illustrative example, the bond graph of the reaction network

$$A + B \rightleftharpoons C \rightleftharpoons 2D$$

is given in Figure 2, with details explained in the following paragraphs.

Bond graphs are a generalisation of directed graphs, with physical *components* as the nodes and *bonds* as the edges. Bonds are the fundamental unit of connection in bond graphs, and are depicted using harpoons. Each bond carries two physical covariables: the chemical potential μ [J/molecule] and the molecular flow rate v [molecules/s]. Since the product of μ and v is power [J/s], bond graphs track the movement of energy between components and ensure thermodynamic consistency.

2.2.1 Chemical species

Chemical species are labelled as **C** components because they store energy in their concentration. This can be thought of in the context of diffusion: particles in a region of higher concentration have higher potentials and therefore tend to move to regions of lower concentration. In dilute

solutions, the chemical potential μ_i of a chemical species i is given by

$$\mu_i = k_B T \ln(K_i x_i),$$

where $k_B = 1.381 \times 10^{-23}$ J/K is the Boltzmann constant, T [K] is the absolute temperature, K_i [molecules^{-1}] is a thermodynamic parameter and $x_i = \int_0^t v_i(\tau) d\tau + x_i(0)$ is the molar amount of species, defined as the integral of molecular flow [31]. For convenience, we define the dimensionless chemical potential as

$$\bar{\mu} = \frac{\mu_i}{k_B T} = \ln\left(K_i x_i\right). \tag{1}$$

In some cases, species are fixed at constant concentrations due to an external supply. These are represented as **Se** components whose concentrations remain constant.

2.2.2 Reactions

Reactions convert chemical species into other chemical species, and are labelled as **Re** components. Due to the principle of microscopic reversibility, all reactions are reversible in reality and therefore are modelled as reversible reactions [32]. Assuming mass-action kinetics, the flux through a reaction j is given by the Marcelin-de Donder equation

$$v_j = r_j \left(e^{\mu_j^f/(k_B T)} - e^{\mu_j^r/(k_B T)} \right) = r_j \left(e^{\bar{\mu}_j^f} - e^{\bar{\mu}_j^r} \right), \tag{2}$$

where r_j is the reaction rate parameter, μ_j^f is the sum of reactant chemical potentials and μ_j^r is the sum of product chemical potentials [25]. The bar notation indicates dimensionless chemical potentials.

2.2.3 Network topology

When a species is involved in multiple reactions, they are connected to **0** junctions. This ensures that the species accumulate at a rate equal to the sum of contributions from each reaction (mass balance). The junction also ensures that the species contributes the same chemical potential to each reaction. For example, the **0** junction in Figure 2 ensures that the flow of the bond into the species C is equal to the difference between flows of the other two connected bonds ($\dot{x}_C = v_1 - v_2$).

When either side of a reaction involves multiple species, the species are connected to a **1** junction prior to the reaction. This junction ensures that all species are consumed or produced at the same rate, and that the contributions of the chemical potentials are summed appropriately. For example, the **1** junction in Figure 2 ensures that the chemical potential on the outgoing bond, $\mu_A + \mu_B$, is equal to the sum of chemical potentials of incoming bonds.

2.2.4 Stoichiometry

If a species contributes multiple stoichiometry to a reaction, it is connected through a **TF** component. This component encodes the relationship

$$\mu_2 = s\mu_1,$$
$$v_1 = s v_2,$$

where s is the stoichiometry, the subscript 1 indicates the bond closer to the species and the subscript 2 indicates the bond closer to the reaction. The **TF** component ensures that the chemical potential is counted multiple times prior to calculating the contribution of a species to a reaction, and that flux is scaled up by the stoichiometric constant.

2.2.5 Relationship to kinetic modelling

A general methodology for constructing bond graphs of biochemical systems is given in Gawthrop and Crampin [25]. Using this approach, the potential balance constraints of the **1** junctions give rise to the equations

$$\mu_j^f = \sum_i s_{ij}^f \mu_i,$$
$$\mu_j^r = \sum_i s_{ij}^r \mu_i,$$

where s_{ij}^f is the stoichiometric contribution of species i to the forward side of reaction j, and similarly for s_{ij}^r on the reverse side.

By using Eqs. (1)–(2), we find that

$$v_j = r_j \left[\exp\left(\sum_i s_{ij}^f \bar{\mu}_i \right) - \exp\left(\sum_i s_{ij}^r \bar{\mu}_i \right) \right], \tag{3}$$
$$= r_j \left[\exp\left(\sum_i s_{ij}^f \ln(K_i x_i) \right) - \exp\left(\sum_i s_{ij}^r \ln(K_i x_i) \right) \right], \tag{4}$$
$$= r_j \left[\prod_i (K_i x_i)^{s_{ij}^f} - \prod_i (K_i x_i)^{s_{ij}^r} \right]. \tag{5}$$

This is the well-known law of mass action

$$v_j = k_{f,j} \prod_i x_i^{s_{ij}^f} - k_{r,j} \prod_i x_i^{s_{ij}^r}, \tag{6}$$

where the forward and reverse kinetic parameters are defined as

$$k_{f,j} = r_j \prod_{i \in \mathcal{R}_j} K_i^{s_{ij}^f},$$
$$k_{r,j} = r_j \prod_{i \in \mathcal{P}_j} K_i^{s_{ij}^r},$$

where $\mathcal{R}_j$ is the set of all reactants and $\mathcal{P}_j$ the set of all products in reaction j.

Finally, the **0** junctions ensure mass balance, thus

$$\dot{x}_i = \sum_j (s_{ij}^r - s_{ij}^f) v_j. \tag{7}$$

Eqs. (5) and (7) give rise to a set of differential equations for the biochemical system. Thus, the bond graph in Figure 2 encodes the differential equations

$$\dot{x}_A = -v_1,$$
$$\dot{x}_B = -v_1,$$
$$\dot{x}_C = v_1 - v_2,$$
$$\dot{x}_D = 2v_2,$$

where

$$v_1 = r_1 K_A K_B x_A x_B - r_1 K_C x_C,$$
$$v_2 = r_2 K_C x_C - r_2 K_D^2 x_D^2.$$

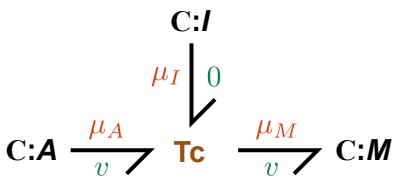

Figure 3: Bond graph model of transcription. The **Tc** component has the rate law described in Eq. (8).

2.3 Transcription

Following Weiße et al. [20], we model transcription as an enzyme-catalysed reaction with the energy (A) as the substrate and mRNA (M) as the product, under the effect of an inhibitor I (such as a transcription factor). We assume non-competitive inhibition, so the net rate of reaction under the quasi-steady-state approximation is

$$v = \bar{r}\frac{e^{\bar{\mu}_A} - e^{\bar{\mu}_M}}{(1 + \frac{e^{\bar{\mu}_A}}{R_{bA}} + \frac{e^{\bar{\mu}_M}}{R_{bM}})(1 + \frac{e^{\bar{\mu}_I}}{R_{bI}})}, \tag{8}$$

where $\bar{r}$ is a rate parameter and R_{bA}, R_{bM} and R_{bI} are affinity parameters (see § A.1 of the electronic supplementary material for further detail). Thus, Eq. (8) is represented as the three-port **Tc** component presented in Figure 3, which is a bond graph representation of transcription that reproduces the energy-saturating behaviour in Weiße et al. [20]. While we only consider the inhibition of transcription in this study, promoters can be modelled by employing a similar approach on an enzyme activation reaction scheme [33].

2.4 Translation

2.4.1 Full model

Because we explicitly model the use of energy and ribosomes in translation, a greater level of detail is required compared to the modelling of transcription. We first describe a detailed model, which is used as the basis for deriving a reduced model that is necessary for modelling large cell populations. Our model of translation considers the following processes:

1. **Binding** of ribosomes (R) to mRNA (M) to create a complex (C_0).
2. **Elongation**, the movement of ribosomes along the mRNA. For every site that the ribosome moves, a molecule of energy (A) is consumed, increasing the length of the nascent protein.
3. **Termination**, the release of ribosome and the completed protein (P) from the mRNA.

The reaction scheme that corresponds to these processes is:

$$M + R \underset{k_u}{\overset{k_b}{\rightleftharpoons}} C_0 \tag{9a}$$

$$C_0 + A \underset{k_{r1}}{\overset{k_{f1}}{\rightleftharpoons}} C_1 + M \tag{9b}$$

$$C_1 \underset{k_r}{\overset{k_f A}{\rightleftharpoons}} C_2 \underset{k_r}{\overset{k_f A}{\rightleftharpoons}} \dots \underset{k_r}{\overset{k_f A}{\rightleftharpoons}} C_n \tag{9c}$$

$$C_n \underset{k_z}{\overset{k_t}{\rightleftharpoons}} R + P \tag{9d}$$

where n is the number of ATP molecules required to synthesise the protein. Consistent with the TASEP model [21], we assume that multiple ribosomes can bind to the same transcript, thus

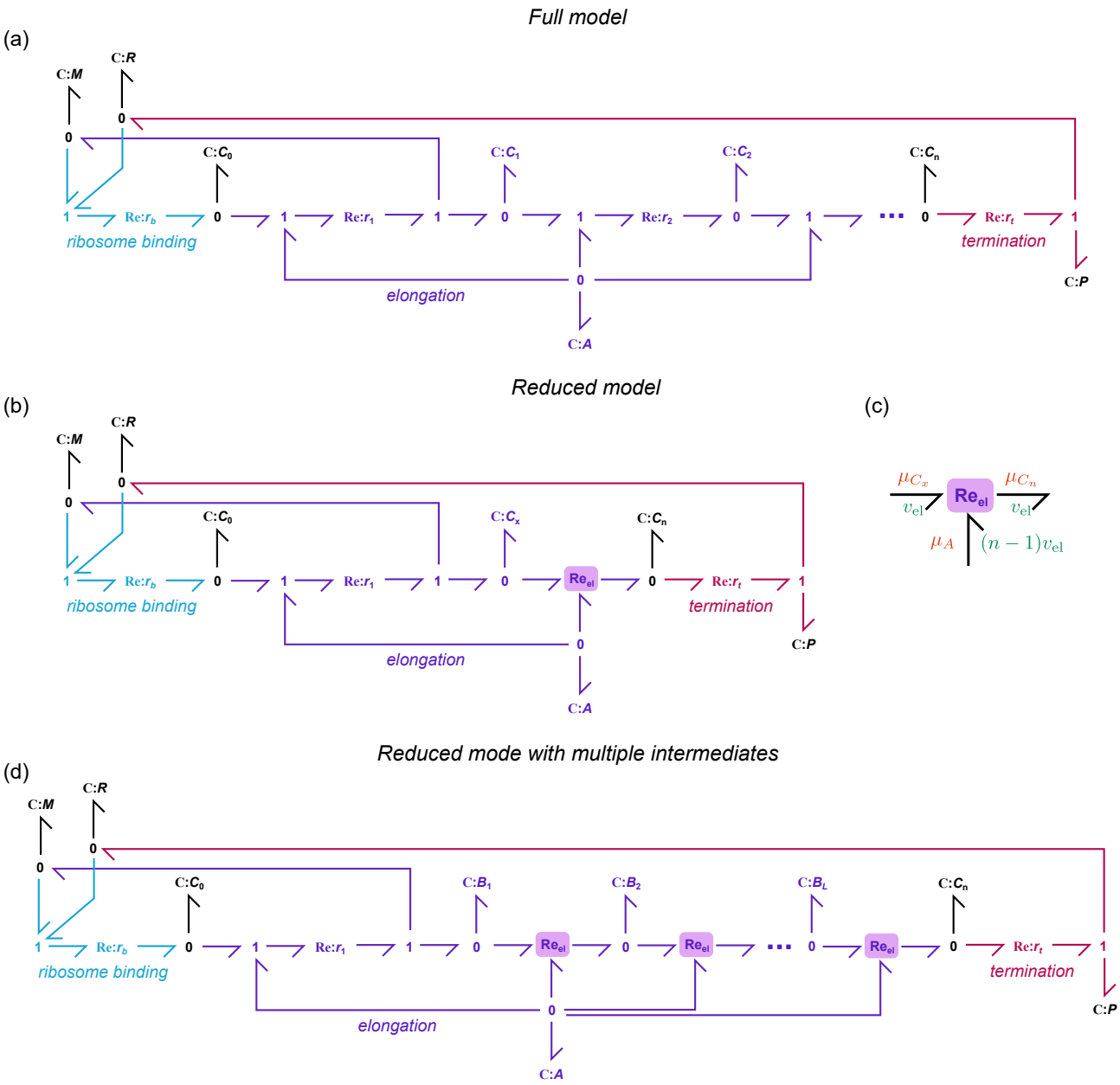

Figure 4: Bond graph models of translation. (a) Bond graph of the full translation model (§ 2.4.1); (b) bond graph of the reduced model (§ 2.4.2); (c) the $\mathrm{Re}_{\mathrm{el}}$ component used to model elongation in (b); (d) bond graph of a reduced model with multiple intermediates.

the mRNA molecule is released after the first elongation step (Eq. (9b)). However, for simplicity, we assume that apart from the first binding site, ribosomes do not obstruct each other, so that the complexes may be treated independently. While both binding and termination kinetics are dependent on the mRNA sequence [34, 35], we assume for simplicity that both initiation and termination dynamics are independent of sequence. A bond graph of the full model is shown in Figure 4a.

2.4.2 Model reduction

Since the number of ATP molecules per protein n can be on the order of thousands, it is computationally expensive to use the full model of translation in models of complex gene regulatory networks. Here, we outline a method for simplifying the full model while retaining

its essential features. We group the complexes C_1 to C_{n-1} into a lumped complex C_x, giving rise to the reaction scheme

$$M + R \underset{k_u}{\overset{k_b}{\rightleftharpoons}} C_0 \underset{k_r M}{\overset{k_f A}{\rightleftharpoons}} C_x \overset{v_{\text{el}}}{\rightleftharpoons} C_n \underset{k_z}{\overset{k_t}{\rightleftharpoons}} R + P, \tag{10}$$

where v_{el} is a non-mass-action reaction representing the elongation rate. A bond graph of the system is shown in Figure 4b, where the elongation reaction is represented as a three-port reaction component (Figure 4c). In § A.2.1 of the electronic supplementary material, we show that under conditions where the dynamics of C_1 and C_n are slower than those of the complexes in between, a quasi-steady-state approximation can be applied and the elongation rate is

$$v_{\text{el}} = k_r \left(C_1 \hat{A}^{n-1} - C_n \right) \frac{1 - \hat{A}}{1 - \hat{A}^{n-1}}. \tag{11}$$

We show in § A.2.2 of the electronic supplementary material that this equation can be represented as a three-port bond graph component (Figure 4c) with the rate law

$$v_{\text{el}} = r \left(e^{\bar{\mu}_{C_1}} - e^{\bar{\mu}_{C_n} - m\bar{\mu}_A} \right) \frac{1 - e^{\bar{\mu}_A} \alpha^{-1}}{\left(\alpha e^{-\bar{\mu}_A}\right)^m - 1}.$$

The lumping of intermediate complexes also requires the scaling of parameters and initial conditions, discussed further in § A.2.3 of the electronic supplementary material. In cases where the dynamics need to be better matched, it is also possible group the complexes into multiple lumps (Figure 4d; see § A.2.4 in the electronic supplementary material for details).

2.5 Gene expression and regulation

We construct a model of gene expression and regulation by coupling together the earlier models of transcription and translation (Figure 5). Since bond graphs are graphical by nature, they enable a modular approach to modelling biological systems. Using this approach, models of subsystems – in this case the transcription and translation modules – can be re-used in larger models without the need to manually re-derive the equations. Here, we couple bond graph modules using the approach in Gawthrop and Pan [24], where **C** components corresponding to shared species are exposed as ports within the transcription and translation modules. Degradation processes for mRNA and proteins are added to avoid overaccumulation (Figure 5c); the details are discussed in § A.3 of the electronic supplementary material. The full model equations are given in § A.4 of the electronic supplementary material.

2.6 Synthetic biology motifs

Bond graphs of these modules are shown in Figure 6. These consist of an instance of the gene expression module **GE** for each gene, but with additional connections between the genes. In particular, the protein of one gene expression system acts as the repressor for another gene expression system. These connections are shown by the coloured bonds in Figure 6. Also, since all genes require energy and ribosomes to be expressed, these resources must be shared between modules. This is modelled by connecting the modules to **0** junctions (grey bonds in Figure 6) to ensure mass conservation.

Gene regulatory networks often require binding cooperativity to enhance their sensitivity to inputs [36]. This is incorporated by using a **TF** component (§ 2.2.4) with modulus h to model the binding of multiple inhibitors to prevent transcription, resulting in a transcription rate of

$$v_{\text{Tc}} = \bar{r}_{\text{Tc}} \frac{K_A x_A - K_M x_M}{(1 + \frac{K_A x_A}{R_{bA}} + \frac{K_M x_M}{R_{bM}})(1 + \frac{K_I^h x_I^h}{R_{bI}})},$$

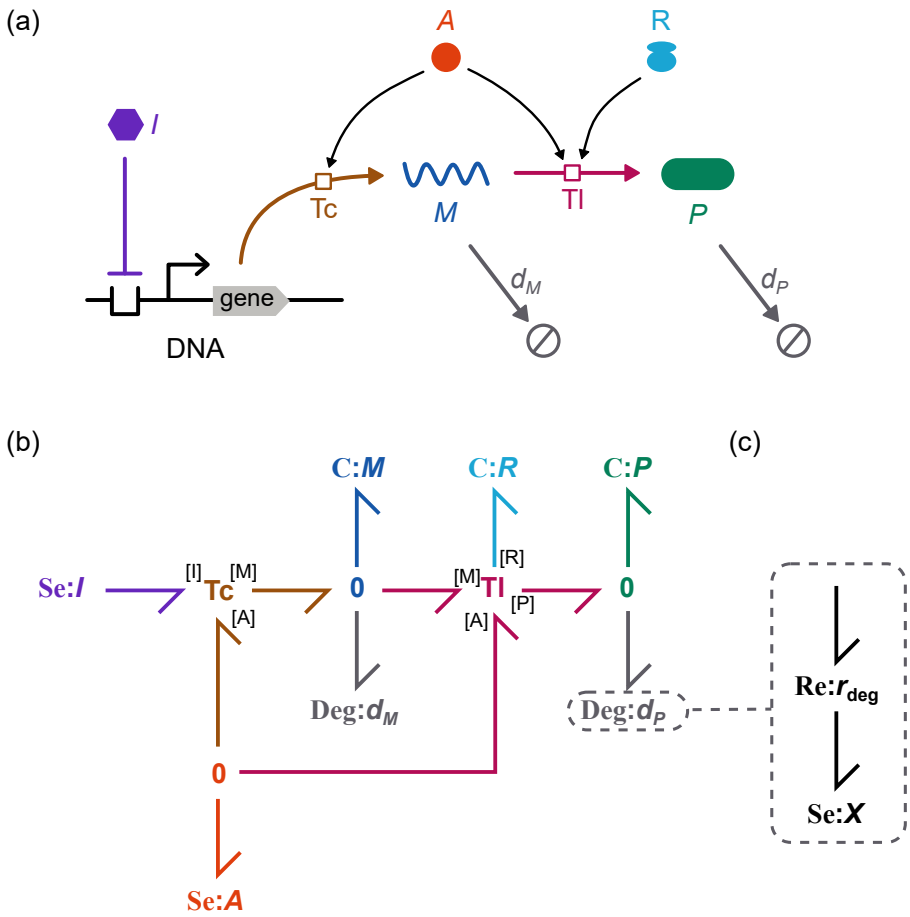

Figure 5: Model of gene expression. (a) Schematic and (b) bond graph model of gene expression. Degradation for mRNA and protein are modelled using the modules in (c). The transcription (Tc) module is shown in Figure 3 and the translation (Tl) module is shown in Figure 4b.

A Hill coefficient of $h = 2$ is used, consistent with previous models [4, 20].

3 Results

3.1 Comparison of full and reduced models

3.1.1 Elongation

To ensure the validity of the model reduction scheme in § 2.4.2, we use dynamic simulations to compare the output of the full and reduced models of translation. We first examine the elongation reactions in isolation to verify that the simple model matches the elongation rate of the full model at steady state. Figure 7a shows the results obtained from simulating the full model of elongation (Eq. (S3) of the electronic supplementary material), compared to the analytical form of the elongation rate (Eq. (11)). For simplicity, we use a chain length of $m = 5$ for these simulations, but the results generalise to longer chain lengths (see below). As expected, the full and reduced models match very well. The steady-state amounts of the intermediate complexes X_1, X_2, X_3 and X_4 are plotted in Figure 7b, and also match closely. We next investigate whether the model reduction scheme holds for different chain lengths m (Figure 7c). The elongation rate decreases, and ultimately plateaus at high m; and the reduced model is able to predict this behaviour. Thus, the reduced model is able to accurately match the steady-state behaviour of the full model under a wide range of conditions.

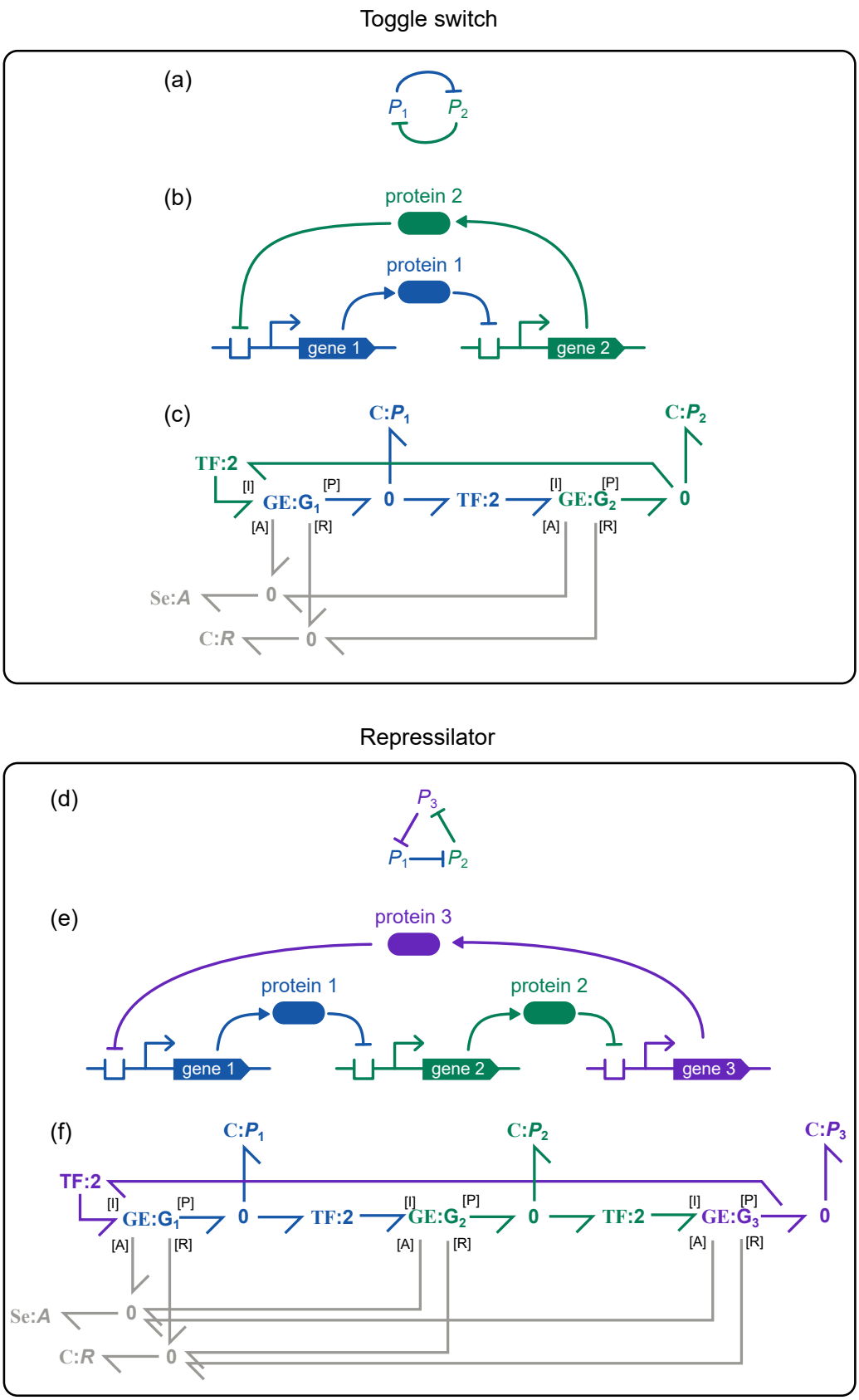

Figure 6: Models of synthetic biology motifs. The bond graph model of (a–c) the toggle switch; and (d–f) repressilator. For each panel, the relationships between proteins are shown on top (a,d), the mechanisms of interaction are displayed in the middle (b,e) and bond graphs are shown on the bottom (c,f). The **GE** modules refer to the model in Figure 5b.

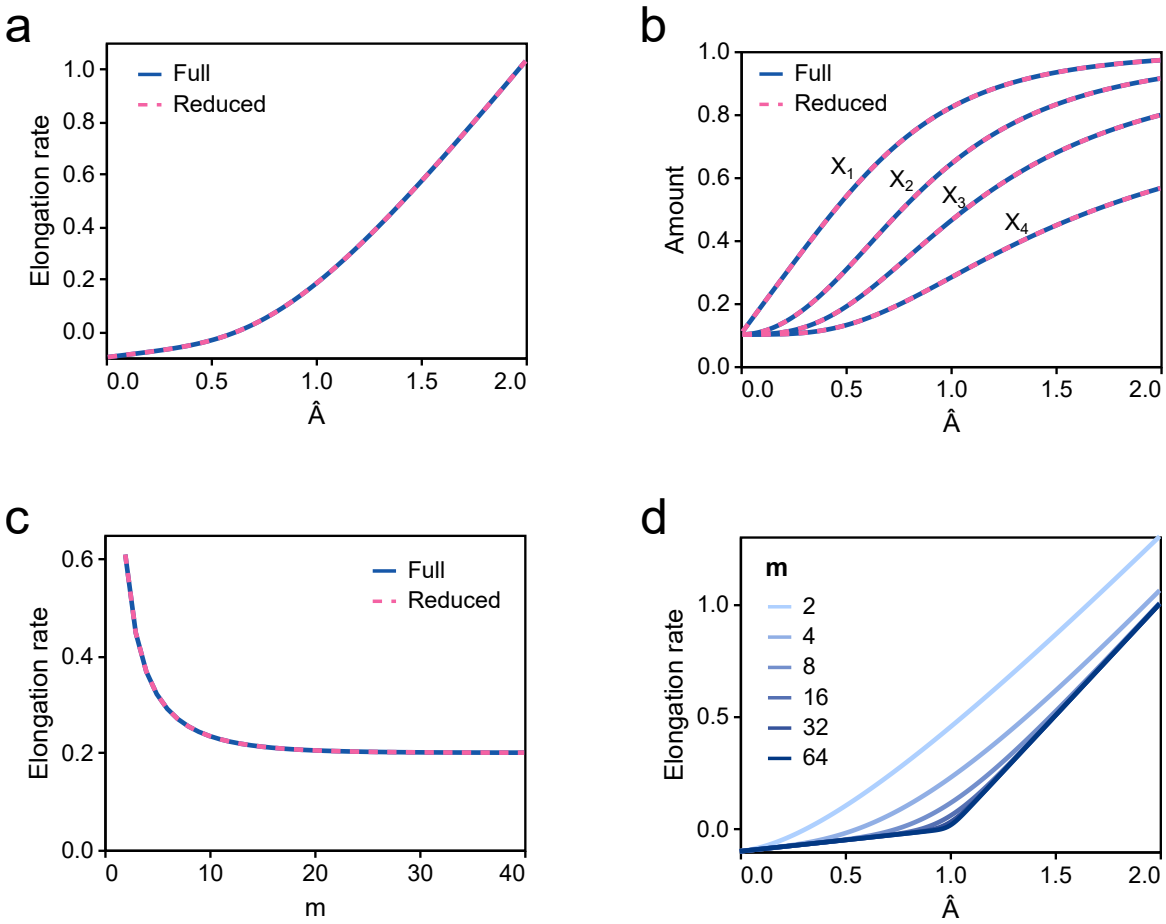

Figure 7: A simplified model of elongation. Steady-state comparisons for (a) elongation rates and (b) complex amounts are plotted for different values of $\hat{A}$. (c) Steady-state comparison of elongation rates for different values of m. The simulations in (a)–(c) use the kinetic parameters $k_f = 1$, $k_r = 1$, $C_0 = 1.0$ and $C_n = 0.1$, with A set to 1.2 for (c). (d) A plot of the analytical solution for different values of m. Models are run for $t = 50$ minutes to achieve steady state.

Figure 7d shows the relationship between elongation rate (Eq. (11)) and $\hat{A}$ for different chain lengths m. For shorter chain lengths, the elongation rate is smooth with respect to $\hat{A}$. However, as predicted by equation Eq. (S7) of the electronic supplementary material, for large m, the curve is almost piecewise linear with different slopes on either side of $\hat{A} = 1$.

3.1.2 Translation

While the above results confirm similarities between the full and reduced models at steady state, we are interested in comparing their behaviours when included in a dynamic model. Thus, we next investigate whether the dynamics of translation are similar between the full and reduced models. To simulate the models under biological conditions, the models are parameterised based on previous experimental and modelling work (see Appendix B in the electronic supplementary material). We simulate the dynamics of the full (Figure 4a) and reduced models (Figure 4b), and compare them for different energy levels and protein lengths (Figure 8). Models are initialised from an inactivated state with very small amounts of ribosome-mRNA complexes, and the evolution to an activated state is compared between models. We test the match between models for a range of energetic states A, with $A = 5.8 \times 10^6$ molecules/cell corresponding most closely to an energy-rich environment (§ B.1 in the electronic supplementary material). We also test the quality of match for different chain lengths n.

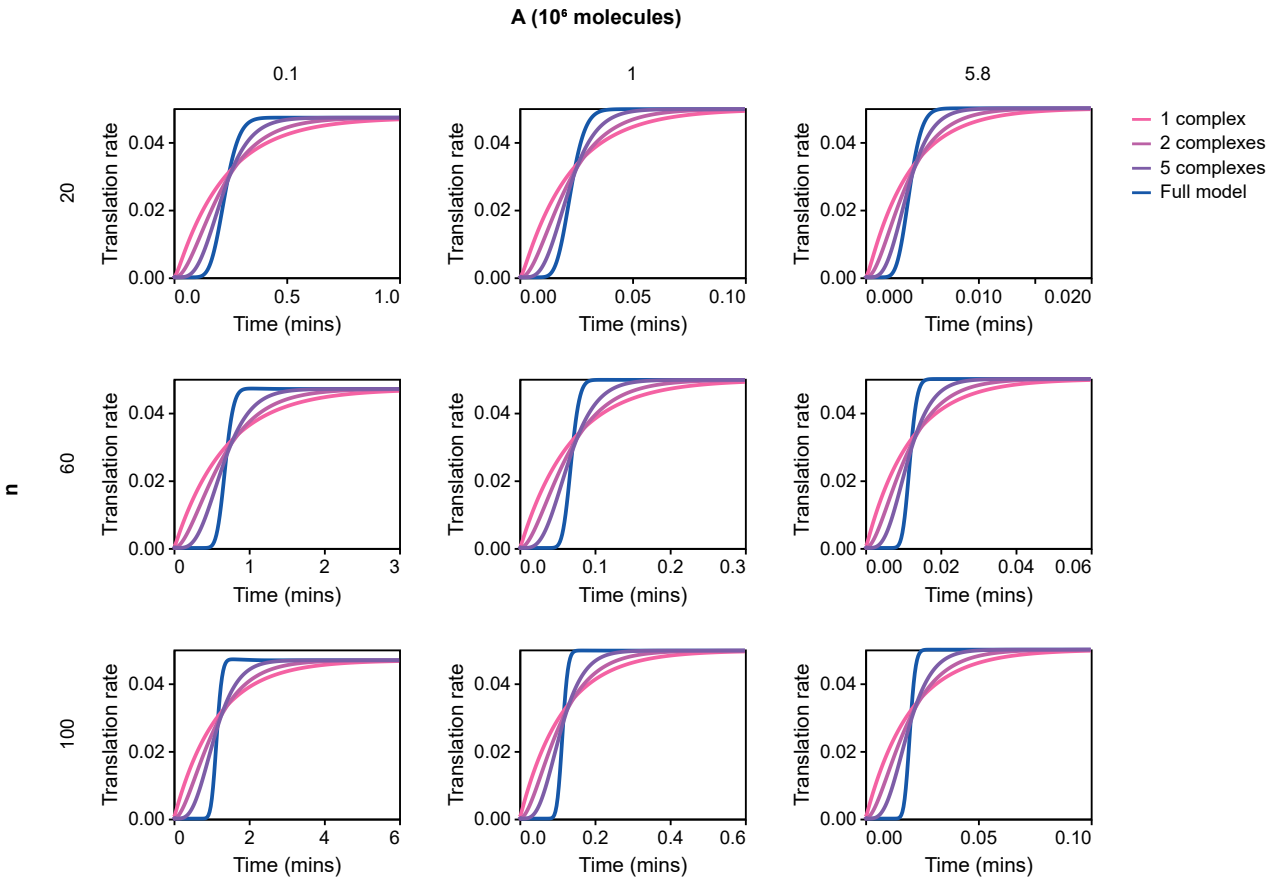

Figure 8: Comparison of the dynamics of translation rate. The energetic state A is varied horizontally, and the chain length n is varied vertically. The translation rate is given in proteins per minute. Simulations are run with 10 mRNA molecules, 5000 ribosomes and $\bar{\mu}_P = -15$ to approximate the immediate removal of proteins.

In all cases, the reduced model approaches the same translation rate as the full model. However, the models exhibit different dynamics. Whereas the full model has a more step-like response to a sudden increase in mRNA, the reduced model approaches the steady-state translation rate more gradually. As one would expect, the match between dynamics becomes poorer for larger values of n. Nonetheless, both the full and reduced models approach their steady states at similar timescales for a wide range of chain lengths. Due to the similarities in steady-state behaviour and timescales, we believe that the simple model is an adequate approximation, and we indeed show in § 3.2 that the differences in dynamics result in very minor differences at the level of gene expression. In circumstances where better matching with the dynamic behaviour is required, one could lump elongation into multiple intermediate states rather than a single complex. This approach is analogous to epidemiological models with multiple exposed or infectious compartments [37] and models of multi-stage transport processes [38]. Simulations of two and five lumped complexes are displayed in Figure 8, and show a gradual improvement of fit with the number of intermediates.

We also find that the reduced model is computationally efficient compared to the full model. A benchmark of runtimes is presented in Figure 9. As expected, the full model takes longer to run as the number of state variables increases; the approximately linear gradient at high n indicates a rapidly increasing runtime with protein length. In contrast, the time taken to run the reduced model remains stable even for higher protein lengths. At the longest tested protein

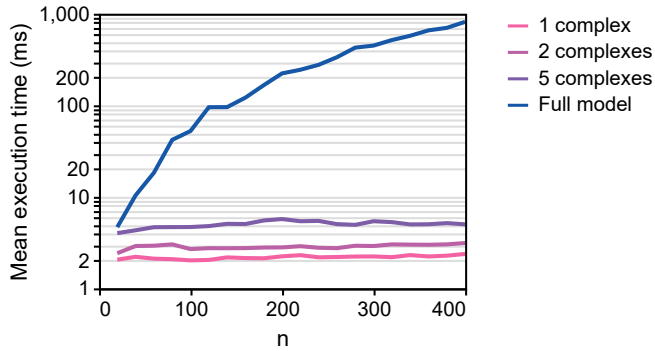

Figure 9: Benchmark of full and reduced models. Simulations are run on the translation model and the mean execution time is plotted for different protein lengths n. Models were solved using the CVODE Backward Differentiation Formula method [39] with a relative tolerance of 10^{-9} and absolute tolerance of 10^{-12}. Models were repeatedly simulated over either 1000 samples, or until 1 minute had elapsed in real time. The code was simulated on a Windows 10 server with 4 CPU cores (AMD EPYC 7702 processor).

length of $n = 400$, a 300× speedup is observed for the reduced model, relative to the full model. We expect even greater speedup for realistic protein lengths of over 1000 ATP molecules per protein.

In Appendix D of the electronic supplementary material, we verify that the bond graph model is thermodynamically consistent and analyse the energetics of gene expression under different energy availabilities.

3.2 Gene expression

Models of gene expression are constructed by coupling models of transcription and translation (Figure 5). We simulate the induction of a gene, where an inhibitor is initially present at high concentrations before being removed (Figure 10a). As expected, following the removal of the inhibitor, mRNA levels rise followed by a rise in the the protein. The counts of these molecules are within the ranges observed experimentally [40]. In § C.1 of the electronic supplementary material, we show that under the assumption of excess energy, the steady-state mRNA and protein numbers are

$$x_M = \frac{1}{r_{\mathrm{deg},M} K_M} \frac{\bar{r}_{\mathrm{Tc}} R_{bA}}{1 + e^{\bar{\mu}_I}/R_{bI}},$$
$$x_P = \frac{r_b K_R x_R}{r_{\mathrm{deg},M} r_{\mathrm{deg},P} K_P} \frac{\bar{r}_{\mathrm{Tc}} R_{bA}}{1 + e^{\bar{\mu}_I}/R_{bI}}.$$

With the parameters in the model, $x_M \approx 41$ and $x_P \approx 10351$, which is in good agreement with Figure 10b–c. The power dissipation rate is plotted against time in Figure S3 of the electronic supplementary material, which indicates that protein synthesis and degradation account for the majority of energy dissipated within the circuit.

To ensure that the model reduction scheme in § 2.4.2 is valid, the same models are simulated with the full model of translation (Figure 4a) in place of the reduced models. The resulting mRNA and protein levels are virtually identical to that of Figure 10 (Figure 11), and the quality of match is similar for the models of the toggle switch and repressilator. Thus, we expect that translation operates near steady state under biological conditions, and that the loss of dynamics from model reduction is unimportant under these conditions.

In Figure 12, the translation rate is plotted against the amount of energy and ribosomes.

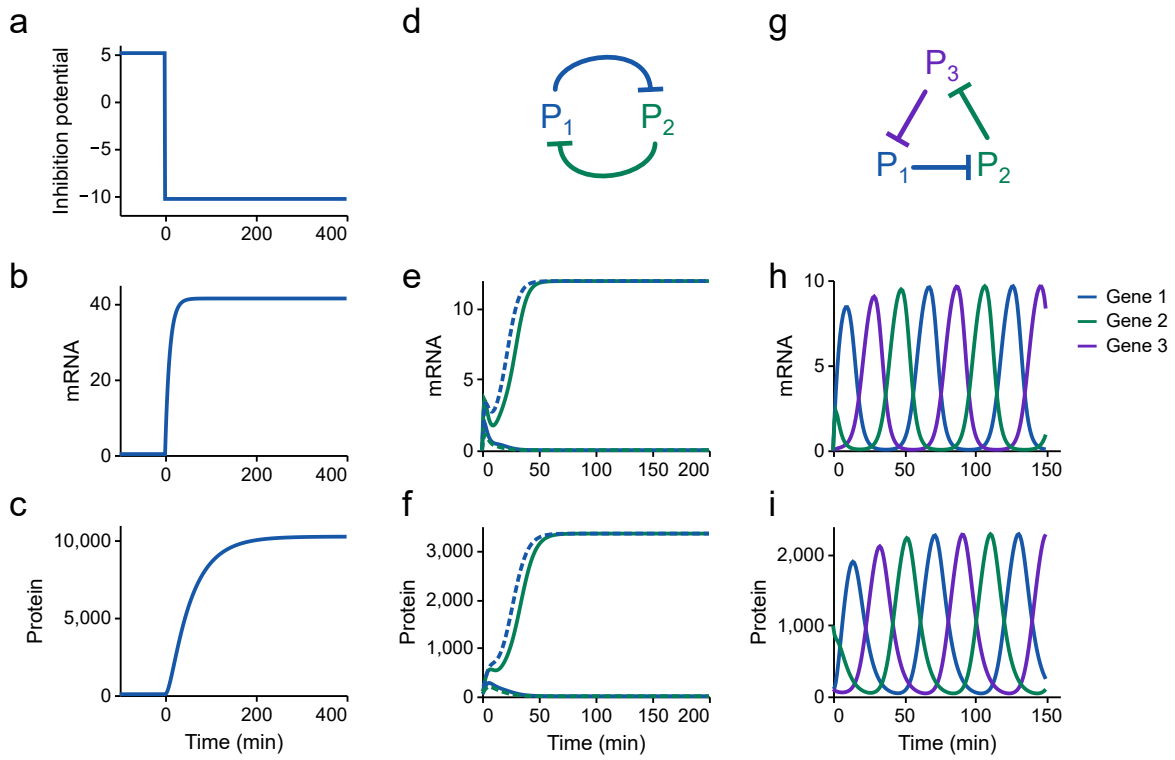

Figure 10: Simulations of synthetic circuits. (a–c) Simulations of a single gene. The inhibition potential $\mu_I - \ln(R_{bI})$ is reduced from a high to low value at $t = 0$, indicating a switch from an inactive to active state. (d–f) Simulations of the toggle switch. Two simulations (represented using solid and dashed lines) are run with different initial conditions. The steady states are dependent on the initial condition, indicating bistability. (g–i) Simulations of the repressilator circuit.

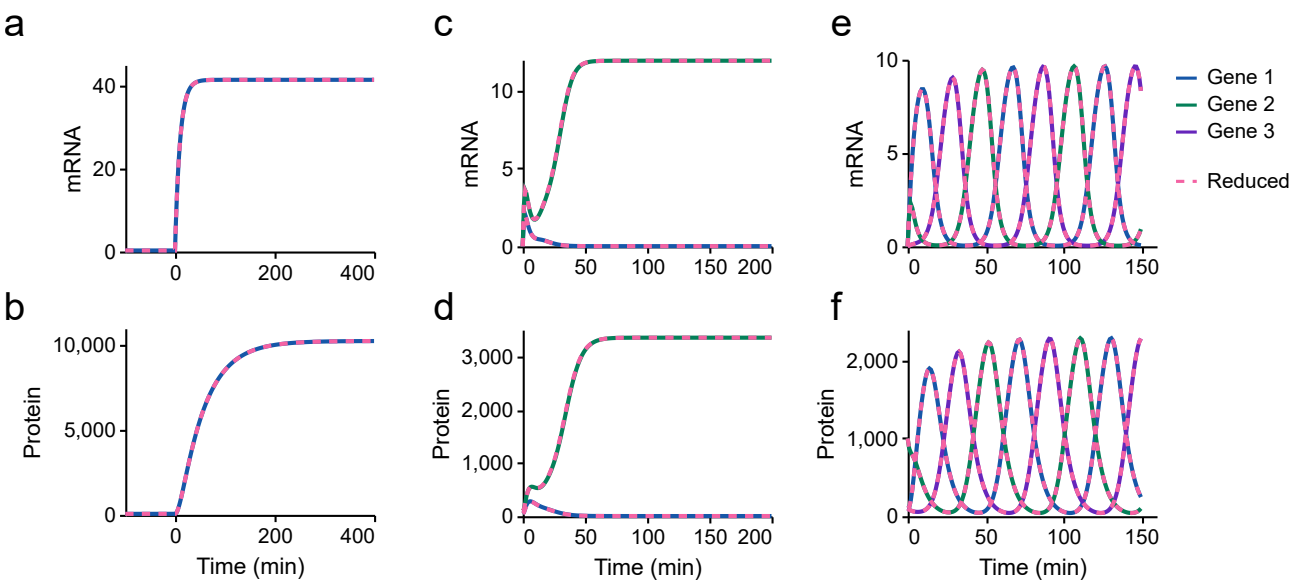

Figure 11: Comparison of full and reduced models of synthetic circuits. Models of synthetic circuits using the full translation model (solid lines) are compared to the corresponding models using the reduced translation model (dashed lines). (a–b) Induction of a single gene; (c–d) toggle switch; (e–f) repressilator.

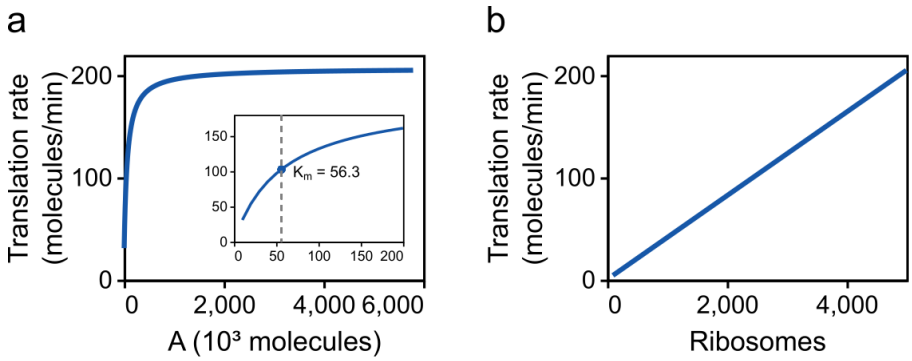

Figure 12: The effect of resource abundance on translation rate. The translation rate is plotted against (a) amount of energy; and (b) the number of ribosomes. The inset in (a) shows the translation rate for low amounts of energy.

Decreasing the amount of energy has little effect on the translation rate (Figure 12a) until the amount of energy is very low, reflecting energy starvation conditions. The translation rate reaches 50% of its maximum value at $A = 5.63 \times 10^4$ molecules (Figure 12a; inset), which corresponds to a concentration of $K_m^{\text{ATP}} = 93\ \mu$M. It is reassuring that this value is on the same order of magnitude as experimental measurements ($K_m^{\text{ATP}} \approx 27\ \mu$M; Jewett et al. [41]) despite the fact that the model has not been fitted to such data. Decreasing the number of ribosomes reduces the translation rate linearly (Figure 12b). This appears to be consistent with studies that hypothesise ribosomes are a limiting factor for translation [42]. We also show in § C.1 of the electronic supplementary material that the steady-state translation rate is linearly dependent on ribosome number at high energy levels.

3.3 Resource dependence of synthetic circuits

3.3.1 Toggle switch

We next explore the behaviours of synthetic biology circuits when modelled with thermodynamic dependencies in mind. A model of the toggle switch (Figure 6a–c) is constructed by coupling together two mutually inhibiting genes. We perform simulations of the toggle switch under two different initial conditions (Figure 10d–f), corresponding to the activation of protein 1 (dashed lines) and protein 2 (solid lines). As expected, the model reproduces the multistability of the switch observed experimentally [3].

As discussed in Section 1, synthetic circuits do not operate in isolation, and are affected by background processes such as the allocation of energy and ribosomes. We examine the effects of limiting either energy (A) or ribosome (R) availability on the performance of the toggle switch. The performance of the toggle switch is characterised using two measures:

1. the *bistability index*, defined as the steady-state amount of P_1 in the active steady-state relative to its amount in the inactive state [43].
2. the *switching time*, defined as the time taken for both proteins to reach their steady state [3]. To quantify the switching time, we define the distance from the steady state as

$$d = \sqrt{\left(\frac{P_1 - P_1^*}{P_1^*}\right)^2 + \left(\frac{P_2 - P_2^*}{P_2^*}\right)^2}, \tag{12}$$

where P_1^* and P_2^* are the steady-state amounts of proteins 1 and 2 respectively. The switching time is the time taken for the model to reach a distance of $d < 0.01$.

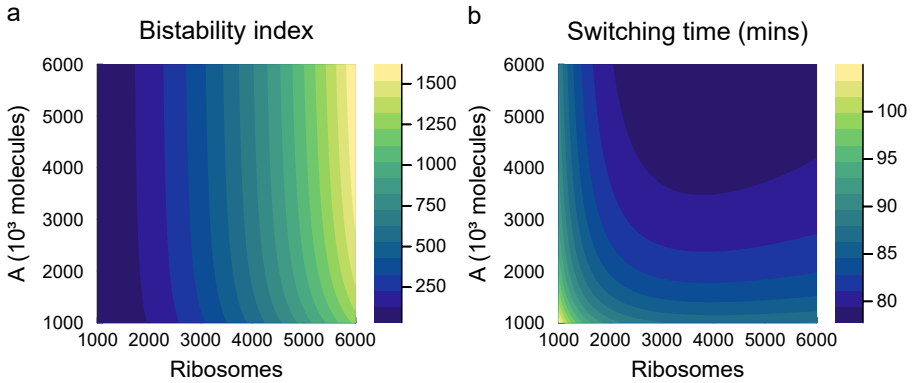

Figure 13: Simulations of the toggle switch under varying energy and ribosome levels. Simulations are run with an initially high amount of protein for gene 1, resulting in a steady state favouring this gene. (a) The bistability index, defined as the steady-state amount of protein 1 in its active state relative to its inactive state; (b) switching time, defined as the time required for the system to be within 1% of its steady-state value (Eq. (12)).

Figure 13 shows the effect of energy and ribosomes on the behaviour of the toggle switch. We find that ribosomes are the primary determinant of toggle switch bistability, as the bistability index varies by approximately 30-fold over the range of ribosome amounts considered. The steady states of the toggle switch under energy-rich conditions are analysed in § C.2 of the electronic supplementary material, verifying this result. In contrast, both the energy and ribosomes affect the switching time. A higher energy availability reduces the switching time as expected, since energy accelerates the dynamics of translation. Surprisingly, the number of ribosomes has a nonmonotonic effect on switching time, and achieves a minimum value between 3000–6000 ribosomes. This leads to potential trade-offs in performance at high numbers of ribosomes; increasing ribosomal availability increases the bistability index, but also increases the switching time.

3.3.2 Repressilator

We next study the dynamic behaviour of the repressilator. Simulations of the repressilator show oscillations between the three inhibitory proteins, in line with the expected function of the circuit (Figure 10g–i). In Figure 14, we explore the behaviour of the circuit under different energy and ribosome availabilities, characterising the amplitude and period of the steady-state oscillations. The amplitude of oscillations is primarily determined by the number of ribosomes; a reduction of ribosomes from 5000 to 1000 results in an approximately 7-fold decrease in oscillation amplitude (Figure 14a). Both energy and ribosome availability contribute to the period of oscillations. Increasing energy availability accelerates the dynamics of translation, increasing the frequency of oscillations, whereas increasing the number of ribosomes decreases the frequency of oscillations. Interestingly, similarities can be drawn with results from the toggle switch, where measures associated with the amplitude of protein changes are mostly impacted by ribosome availability, whereas temporal behaviours are determined by both energy and ribosome availability.

3.4 Parameter heterogeneity

A characteristic of cell populations is that they may experience different local environments and therefore display heterogeneity in their behaviour [44]. Because our simple model of gene expression is computationally efficient, it is well-suited to exploring the effects of heterogeneity within cell populations. Accordingly, we explore the effects of parameter heterogeneity on the behaviour of the repressilator (Figure 6d–f). Note that these investigations are intended to be a

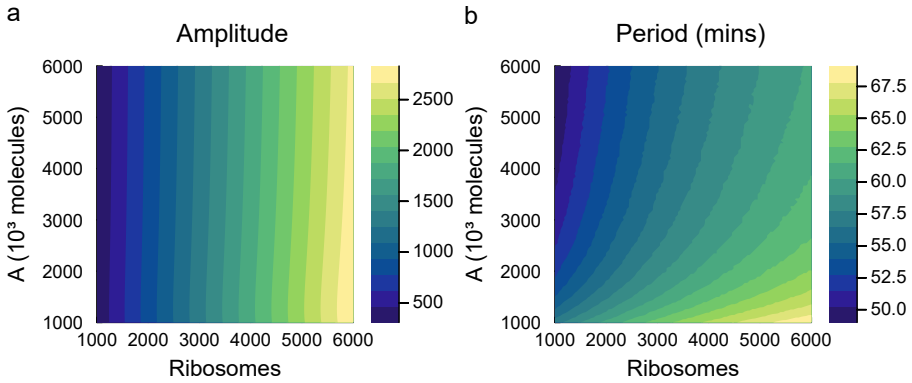

Figure 14: Simulations of the repressilator under varying energy and ribosome levels. Simulations are run until a limit cycle is reached, and the (a) amplitude and (b) period of oscillations are recorded.

proof-of-concept for simulating cell populations since they assume cells operate independently of each other. In reality, cells influence one another through the transmission of extracellular signals and the competition for nutrients and resources.

We first investigate the impact of heterogeneity in energy availability on repressilator behaviour. Nakatani et al. [45] find that ATP concentrations vary substantially in populations of *E. coli*. To simulate the effect of this heterogeneity, we fit a normal distribution (mean = 3.76 mM, std = 1.64 mM) to Fig. 1A of Nakatani et al. [45]. These concentrations are converted to a number of molecules, assuming a cell volume of 1 μm^3 [20]. We obtain samples of energy levels using a truncated normal distribution with a minimum of $A = 10^4$ molecules per cell. The resulting behaviour for 10000 samples is summarised in Figure 15a–d. Due to the relative insensitivity of the repressilator model to energy levels, the effect on the amplitude of oscillations was small, with the vast majority of samples having amplitudes between 2200–2400 (Figure 15b). The effect on the period of oscillations was similarly small, with most samples having a period of between 60–66 mins (Figure 15c).

Next, we consider heterogeneity in ribosome amounts. We assume that the number of free ribosomes per cell follows a similar distribution to protein counts in cells. Milo and Phillips [40] estimated that protein counts have a coefficient of variation of approximately 0.4 in bacteria. Thus, we sample ribosome counts from a log-normal distribution around a median free ribosome count of 5000 (Figure 15e). This heterogeneity has a marked impact on the amplitudes of oscillation ($\sim$8-fold variation; Figure 15f), and a mild impact on the period ($\sim$ 25% variation; Figure 15g).

We finally explore the effect of variations in protein degradation rates. Since the protein half-life has been shown to vary 2–3 fold among populations of *E. coli* [46], we assume the protein half-life is log-normally distributed with a median half-life of 4 minutes and a 95% confidence interval spanning from $1/\sqrt{2}$ to $\sqrt{2}$ times the median. We also assume that protein half-lives are perfectly correlated between proteins, caused by global fluctuations in cell state [44]. Using these proof-of-concept simulations, heterogeneity in protein degradation results in a wide range of variation in both the amplitudes and periods of oscillations, with approximately 2-fold variation in their values (Figure 15i–l).

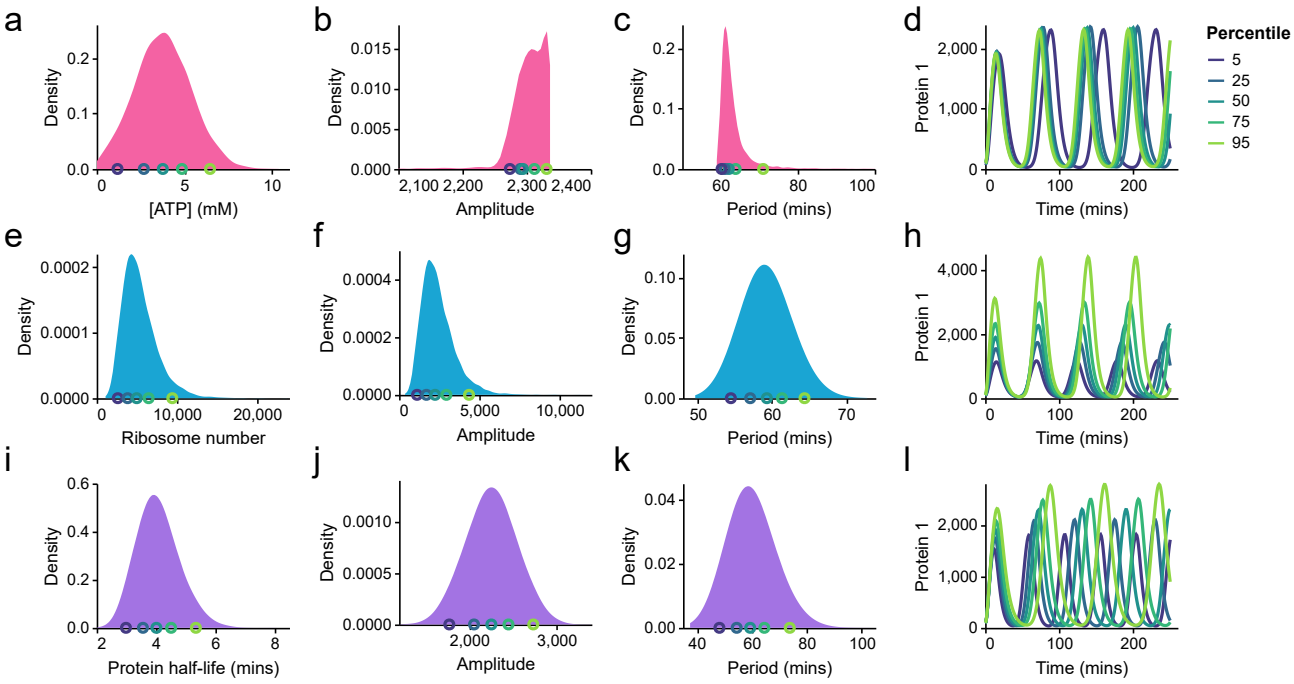

Figure 15: Effects of parameter heterogeneity on repressilator function. Each row corresponds to distributions extracted from running 10000 simulations, with a parameter varied by sampling from a distribution. (a–d) Effects of heterogeneity in ATP concentration. Changes in ATP concentration were modelled by adjusting the parameter μ_A (14 samples were excluded from analysis as they resulted in no oscillations due to insufficient energy levels). (e–h) Effects of heterogeneity of ribosome amount, which are modelled by changing the initial ribosome number. (i–l) Effects of heterogeneity in protein decay rates, incorporated by adjusting the parameter $r_{\text{deg},P}$. For each row, simulations corresponding to the 5th, 25th, 50th, 75th and 95th percentiles of the parameters are plotted on the right (d,h,l). The colours of the timecourses correspond to the circles of the same colour in the distributions.

4 Discussion

In this study, we develop a new bond graph model of gene expression. In contrast to previous models in the literature [17, 18, 20, 21], our model is both simple and thermodynamically consistent, making it suitable for understanding and designing synthetic circuits in cell populations. By coupling together multiple instances of our gene expression module, we are able to reproduce behaviours of two foundational synthetic biological circuits: the toggle switch and repressilator [3, 4]. However, the approach is generalisable to other synthetic circuits involving protein synthesis, such as antithetic feedback [47] and could be used to study the effects of energy burden in such control systems. Our approach is based on bond graphs, a general framework for modelling biophysical systems. Due to their physical consistency, commonly neglected phenomena such as resource competition and retroactivity are automatically built into bond graph models [8, 10]. Additionally, their modular and hierarchical structure allows modellers to systematically couple models of synthetic circuits into larger circuits [28]. Thus, we believe that bond graphs hold great potential in aiding the design of circuits in synthetic biology.

Our model of gene expression is based on a model reduction scheme to simplify the elongation reactions in translation (§ 2.4.2). It is common for other simple models of translation to assign rate laws based on empirical fits to data [4, 17]. By contrast, our model reduction approach is biophysically motivated, based on a detailed reaction network containing molecules and processes known to be involved in translation. As a result, our model reduction scheme preserves important

constraints such as mass and energy conservation. We show that the reduced model behaves similarly to the full model at steady state and reaches the steady state on similar timescales (Figures 7–8). Furthermore, we find that the loss of dynamics in translation is unimportant when simulating synthetic gene circuits (Figure 11). The reduced model is also computationally efficient, able to achieve runtimes hundreds of times (and possibly up to thousands of times) faster than the full model (Figure 9). This speedup is quite significant when scaling up to large gene regulatory networks or agent-based models of large cell populations. Given that the addition of monomers to macromolecules is prevalent in biology, we believe that our model reduction scheme would also be applicable to processes such as transcription, glycogenesis and fatty acid polymerisation [48].

In spite of the constraints that energy places on cells, we still have an incomplete picture of how energy constrains cellular processes and how the cell distributes its energy budget between processes [49]. While previous works have studied the effects of ribosome limitation on the function of synthetic circuits [50, 51], to the best of our knowledge, our model is the first to incorporate the known thermodynamics of ATP hydrolysis into a lumped parameter dynamic model of translation. Our model predicts a relative insensitivity of gene expression to energy, which appears to be supported by experimental measurements [41]. Our model is also able to analyse the thermodynamic costs associated with translation. We find that the majority of energy supplied through ATP hydrolysis is dissipated during elongation rather than being stored in the peptide bonds in proteins (Appendix D in the electronic supplementary material). This inefficiency has previously been attributed to proofreading mechanisms [52]; more detailed modelling of these mechanisms is required to understand performance-cost trade-offs in translation.

Our model also makes predictions about how energy and ribosome availability affect the behaviour of synthetic circuits. Interestingly, the number of free ribosomes is the primary determinant of measures associated with protein counts, whereas both energy and ribosome availability contribute to the dynamic behaviour of the circuits. We also explore the potential effects of parameter heterogeneity in cell populations. Informed by empirical measurements, we find that variations in energy levels have a limited impact on the behaviour of the repressilator (Figure 15). However, variability in both ribosome availability and protein degradation rates result in high degrees of variation in both the amplitude and period of oscillations. Furthermore, each parameter influences different aspects of the function of the repressilator. While ribosome availability greatly influences the amplitude of oscillations, the protein degradation rate has a greater impact on the frequency of oscillations. Thus, the parameter variability that exists between cells may be responsible for cell-to-cell variability in the performance of synthetic circuits, in the absence of mechanisms of intercellular interaction [3, 53].

While we explore parameter heterogeneity as a proof-of-concept for modelling cell populations, mathematical questions remain on how to incorporate mechanisms of cell-to-cell communication. Future work will focus on incorporating our model of gene expression into an agent-based model. While the model in this study is broadly applicable to synthetic biology, the bioluminescence pathway in bacteria is a promising model system to engineer for several reasons, including (i) the ease of measuring light as the output, (ii) the involvement of quorum sensing in coordinating population-level behaviours and (iii) the use of two competing pathways for its induction [54]. Furthermore, the multi-physical nature of bond graphs allows them to naturally account for the conversion of chemical energy to light [24]. Modelling the intracellular regulatory mechanisms associated with quorum sensing alongside intercellular communication has the potential to enable the engineering of synchronised behaviour of cell populations [15, 55].

Due to the simple nature of the model, certain details of gene regulation have been omitted. While we assume that elongation involves a single reaction per ATP molecule, tRNA is in reality charged by amino acids through elongation factors, which involves a more complex series of reactions [56]. The elongation rate is dependent on tRNA availability, and therefore specific to each codon. We also lump energy into a single molecule, but in reality, it is composed of the hydrolysis of ATP into ADP, Pi and H^+. Accounting for such details will be the subject of future work.

Acknowledgments

S.T.J. is supported by the Australian Research Council (Project No. DE200100988). P.J.G. would like to thank the Faculty of Engineering and Information Technology, University of Melbourne, for its support via a Professorial Fellowship. M.P. was supported by a Postdoctoral Research Fellowship from the School of Mathematics and Statistics at the University of Melbourne.

Data availability

The bond graph models were simulated in Julia using the BondGraphs.jl package [57]. The code is available on GitHub at https://github.com/mic-pan/bg_gene_expression.

Competing Interests

None declared.

References

1. Gray P, *et al.*, 2018. Synthetic biology in Australia: An outlook to 2030. Tech. Rep. 978-0-6483303-1-8, Australian Council of Learned Academies, Melbourne, Victoria, Australia

2. Church GM, Elowitz MB, Smolke CD, Voigt CA, Weiss R, 2014. Realizing the potential of synthetic biology. *Nature Reviews Molecular Cell Biology* **15**(4), 289–294. doi:10.1038/nrm3767

3. Gardner TS, Cantor CR, Collins JJ, 2000. Construction of a genetic toggle switch in Escherichia coli. *Nature* **403**(6767), 339–342. doi:10.1038/35002131

4. Elowitz MB, Leibler S, 2000. A synthetic oscillatory network of transcriptional regulators. *Nature* **403**(6767), 335–338. doi:10.1038/35002125

5. Shaw WM, *et al.*, 2019. Engineering a model cell for rational tuning of GPCR signaling. *Cell* **177**(3), 782–796. doi:10.1016/j.cell.2019.02.023

6. Gander MW, Vrana JD, Voje WE, Carothers JM, Klavins E, 2017. Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. *Nature Communications* **8**(15459), 1–11. doi:10.1038/ncomms15459

7. Slusarczyk AL, Lin A, Weiss R, 2012. Foundations for the design and implementation of synthetic genetic circuits. *Nature Reviews Genetics* **13**, 406–420. doi:10.1038/nrg3227

8. Boo A, Ellis T, Stan GB, 2019. Host-aware synthetic biology. *Current Opinion in Systems Biology* **14**, 66–72. doi:10.1016/j.coisb.2019.03.001

9. Gorochowski TE, 2016. Agent-based modelling in synthetic biology. *Essays in Biochemistry* **60**(4), 325–336. doi:10.1042/EBC20160037

10. Del Vecchio D, 2013. A control theoretic framework for modular analysis and design of biomolecular networks. *Annual Reviews in Control* **37**(2), 333–345. doi:10.1016/j.arcontrol.2013.09.011

11. Caringella G, Bandiera L, Menolascina F, 2023. Recent advances, opportunities and challenges in cybergenetic identification and control of biomolecular networks. *Current Opinion in Biotechnology* **80**, 102893. doi:10.1016/j.copbio.2023.102893

12. Boo A, Ledesma Amaro R, Stan GB, 2021. Quorum sensing in synthetic biology: A review. *Current Opinion in Systems Biology* **28**, 100378. doi:10.1016/j.coisb.2021.100378

13. Basu S, Gerchman Y, Collins CH, Arnold FH, Weiss R, 2005. A synthetic multicellular system for programmed pattern formation. *Nature* **434**, 1130–1134. doi:10.1038/nature03461

14. Machado D, *et al.*, 2011. Modeling formalisms in systems biology. *AMB Express* **1**(1), 45. doi:10.1186/2191-0855-1-45

15. Prindle A, *et al.*, 2012. A sensing array of radically coupled genetic 'biopixels'. *Nature* **481**, 39–44. doi:10.1038/nature10722

16. Nagarajan K, Ni C, Lu T, 2022. Agent-based modeling of microbial communities. *ACS Synthetic Biology* **11**(11), 3564–3574. doi:10.1021/acssynbio.2c00411

17. Ryall KA, *et al.*, 2012. Network reconstruction and systems analysis of cardiac myocyte hypertrophy signaling. *Journal of Biological Chemistry* **287**(50), 42259–42268. doi:10.1074/jbc.M112.382937

18. Karr JR, *et al.*, 2012. A whole-cell computational model predicts phenotype from genotype. *Cell* **150**(2), 389–401. doi:10.1016/j.cell.2012.05.044

19. Macklin DN, *et al.*, 2020. Simultaneous cross-evaluation of heterogeneous E. coli datasets via mechanistic simulation. *Science* **369**(6502)

20. Weiße AY, Oyarzún DA, Danos V, Swain PS, 2015. Mechanistic links between cellular trade-offs, gene expression, and growth. *Proceedings of the National Academy of Sciences* **112**(9), E1038–E1047. doi:10.1073/pnas.1416533112

21. Raveh A, Margaliot M, Sontag ED, Tuller T, 2016. A model for competition for ribosomes in the cell. *Journal of The Royal Society Interface* **13**(116), 20151062. doi:10.1098/rsif.2015.1062

22. Kempes CP, Wolpert D, Cohen Z, Prez-Mercader J, 2017. The thermodynamic efficiency of computations made in cells across the range of life. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **375**(2109), 20160343. doi:10.1098/rsta.2016.0343

23. Shallom D, Naiger D, Weiss S, Tuller T, 2021. Accelerating whole-cell simulations of mRNA translation using a dedicated hardware. *ACS Synthetic Biology* **10**(12), 3489–3506. doi:10.1021/acssynbio.1c00415

24. Gawthrop PJ, Pan M, 2022. Network thermodynamics of biological systems: A bond graph approach. *Mathematical Biosciences* **352**, 108899. doi:10.1016/j.mbs.2022.108899

25. Gawthrop PJ, Crampin EJ, 2014. Energy-based analysis of biochemical cycles using bond graphs. *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences* **470**(2171), 20140459. doi:10.1098/rspa.2014.0459

26. Pan M, Gawthrop PJ, Tran K, Cursons J, Crampin EJ, 2018. Bond graph modelling of the cardiac action potential: implications for drift and non-unique steady states. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* **474**(2214), 20180106. doi:10.1098/rspa.2018.0106

27. Pan M, Gawthrop PJ, Tran K, Cursons J, Crampin EJ, 2019. A thermodynamic framework for modelling membrane transporters. *Journal of Theoretical Biology* **481**, 10–23. doi:10.1016/j.jtbi.2018.09.034

28. Pan M, Gawthrop PJ, Cursons J, Crampin EJ, 2021. Modular assembly of dynamic models in systems biology. *PLOS Computational Biology* **17**(10), e1009513. doi:10.1371/journal.pcbi.1009513

29. Gawthrop PJ, Pan M, Crampin EJ, 2021. Modular dynamic biomolecular modelling with bond graphs: the unification of stoichiometry, thermodynamics, kinetics and data. *Journal of the Royal Society Interface* **18**(181), 20210478. doi:10.1098/rsif.2021.0478

30. Gawthrop PJ, 2021. Energy-based modeling of the feedback control of biomolecular systems with cyclic flow modulation. *IEEE Transactions on NanoBioscience* **20**(2), 183–192. doi:10.1109/TNB.2021.3058440

31. Atkins P, de Paula J, 2006. *Physical Chemistry for the Life Sciences.* Oxford, England, UK: Oxford University Press

32. Polettini M, Esposito M, 2014. Irreversible thermodynamics of open chemical networks. i. emergent cycles and broken conservation laws. *Journal of Chemical Physics* **141**(2), 024117. doi:10.1063/1.4886396

33. Turberville A, Semple H, Davies G, Ivanov D, Holdgate GA, 2022. A perspective on the discovery of enzyme activators. *SLAS Discovery* **27**(8), 419–427. doi:10.1016/j.slasd.2022.09.001

34. Salis HM, 2011. The ribosome binding site calculator. In *Methods in Enzymology*, vol. 498, pp. 19–42. Cambridge, MA, USA: Academic Press. doi:10.1016/B978-0-12-385120-8.00002-4

35. Prabhakar A, *et al.*, 2023. Dynamics of release factor recycling during translation termination in bacteria. *Nucleic Acids Research* **51**(11), 5774–5790. doi:10.1093/nar/gkad286

36. Keener J, Sneyd J, 2009. *Mathematical Physiology.* New York: Springer

37. Hurtado PJ, Kirosingh AS, 2019. Generalizations of the 'Linear Chain Trick': incorporating more flexible dwell time distributions into mean field ODE models. *Journal of Mathematical Biology* **79**(5), 1831–1883. doi:10.1007/s00285-019-01412-w

38. Dowling CV, Cevaal PM, Faria M, Johnston ST, 2022. On predicting heterogeneity in nanoparticle dosage. *Mathematical Biosciences* **354**, 108928. doi:10.1016/j.mbs.2022.108928

39. Hindmarsh AC, *et al.*, 2005. Sundials: Suite of nonlinear and differential/algebraic equation solvers. *ACM Transactions on Mathematical Software* **31**(3), 363–396. doi:10.1145/1089014.1089020

40. Milo R, Phillips R, 2015. *Cell Biology by the Numbers.* Andover, England, UK: Taylor & Francis. doi:10.1201/9780429258770

41. Jewett MC, Miller ML, Chen Y, Swartz JR, 2009. Continued protein synthesis at low [ATP] and [GTP] enables cell adaptation during energy limitation. *Journal of Bacteriology*

42. Vind J, Sørensen MA, Rasmussen MD, Pedersen S, 1993. Synthesis of proteins in Escherichia coli is limited by the concentration of free ribosomes: Expression from reporter genes does not always reflect functional mRNA levels. *Journal of Molecular Biology* **231**(3), 678–688. doi:10.1006/jmbi.1993.1319

43. Chen JR, 2006. Improving the speed of the genetic toggle switch without sacrificing its dynamic stability. *Physical Review E* **73**(4), 041901. doi:10.1103/PhysRevE.73.041901

44. Elowitz MB, Levine AJ, Siggia ED, Swain PS, 2002. Stochastic gene expression in a single cell. *Science* **297**(5584), 1183–1186. doi:10.1126/science.1070919

45. Nakatani RJ, Itabashi M, Yamada TG, Hiroi NF, Funahashi A, 2022. Intercellular interaction mechanisms promote diversity in intracellular ATP concentration in Escherichia coli populations. *Scientific Reports* **12**(17946), 1–11. doi:10.1038/s41598-022-22189-x

46. Alber AB, Paquet ER, Biserni M, Naef F, Suter DM, 2018. Single live cell monitoring of protein turnover reveals intercellular variability and cell-cycle dependence of degradation rates. *Molecular Cell* **71**(6), 1079–1091.e9. doi:10.1016/j.molcel.2018.07.023

47. Briat C, Gupta A, Khammash M, 2016. Antithetic integral feedback ensures robust perfect adaptation in noisy biomolecular networks. *Cell Systems* **2**(1), 15–26. doi:10.1016/j.cels.2016.01.004

48. Alberts B, 2007. *Molecular Biology of the Cell.* New York: Garland Science, 5th edn.

49. Yang X, *et al.*, 2021. Physical bioenergetics: Energy fluxes, budgets, and constraints in cells. *Proceedings of the National Academy of Sciences* **118**(26), e2026786118. doi:10.1073/pnas.2026786118

50. Qian Y, Huang HH, Jiménez JI, Del Vecchio D, 2017. Resource competition shapes the response of genetic circuits. *ACS Synthetic Biology* **6**(7), 1263–1272. doi:10.1021/acssynbio.6b00361

51. McBride CD, Del Vecchio D, 2021. Predicting composition of genetic circuits with resource competition: Demand and sensitivity. *ACS Synthetic Biology* **10**(12), 3330–3342. doi:10.1021/acssynbio.1c00281

52. Hopfield JJ, 1974. Kinetic proofreading: A new mechanism for reducing errors in biosynthetic processes requiring high specificity. *Proceedings of the National Academy of Sciences* **71**(10), 4135–4139. doi:10.1073/pnas.71.10.4135

53. Portle S, Iadevaia S, San KY, Bennett GN, Mantzaris N, 2009. Environmentally-modulated changes in fluorescence distribution in cells with oscillatory genetic network dynamics. *Journal of Biotechnology* **140**(3), 203–217. doi:10.1016/j.jbiotec.2009.01.011

54. Kuttler C, Hense BA, 2008. Interplay of two quorum sensing regulation systems of vibrio fischeri. *Journal of Theoretical Biology* **251**(1), 167–180. doi:10.1016/j.jtbi.2007.11.015

55. Wu S, Liu J, Liu C, Yang A, Qiao J, 2020. Quorum sensing for population-level control of bacteria and potential therapeutic applications. *Cellular and Molecular Life Sciences* **77**(7), 1319–1343. doi:10.1007/s00018-019-03326-8

56. Hu XP, Dourado H, Schubert P, Lercher MJ, 2020. The protein translation machinery is expressed for maximal efficiency in Escherichia coli. *Nature Communications* **11**(5260), 1–10. doi:10.1038/s41467-020-18948-x

57. Forrest J, Rajagopal V, Stumpf MPH, Pan M, 2023. BondGraphs.jl: composable energy-based modelling in systems biology. *Bioinformatics* **39**(10), btad578. doi:10.1093/bioinformatics/btad578