

Review

Tuning the dials of Synthetic Biology

James A. J. Arpino,^{1,2,3} Edward J. Hancock,⁴ James Anderson,^{4,5}
Mauricio Barahona,³ Guy-Bart V. Stan,^{1,6} Antonis Papachristodoulou⁴
and Karen Polizzi^{1,2}

Correspondence

Antonis Papachristodoulou
antonis@eng.ox.ac.uk

¹Centre for Synthetic Biology and Innovation, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

²Department of Life Sciences, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

³Department of Mathematics, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

⁴Department of Engineering Science, University of Oxford, Parks Road, Oxford OX1 3PJ, UK

⁵St John's College, St Giles, Oxford OX1 3JP, UK

⁶Department of Bioengineering, Imperial College London, South Kensington Campus, London SW7 2AZ, UK

Synthetic Biology is the 'Engineering of Biology' – it aims to use a forward-engineering design cycle based on specifications, modelling, analysis, experimental implementation, testing and validation to modify natural or design new, synthetic biology systems so that they behave in a predictable fashion. Motivated by the need for truly plug-and-play synthetic biological components, we present a comprehensive review of ways in which the various parts of a biological system can be modified systematically. In particular, we review the list of 'dials' that are available to the designer and discuss how they can be modelled, tuned and implemented. The dials are categorized according to whether they operate at the global, transcriptional, translational or post-translational level and the resolution that they operate at. We end this review with a discussion on the relative advantages and disadvantages of some dials over others.

Introduction

The primary goal of Synthetic Biology is to create new or add additional functionality to biological systems by constructing new parts, or modifying existing biological systems (Purnick & Weiss, 2009). Central to this goal is the idea that the synthetic organism is designed following a systematic design framework with a specific objective in mind designed a priori. Ideally such design objectives can be formulated in a quantitative manner so that the performance of the designed component can be quantified and compared to the original design specification. This design framework is required both to improve reliability of individual biological components and to build functioning genetic systems with a larger number of interconnected parts (Purnick & Weiss, 2009), both considered to be current challenges of Synthetic Biology. Currently, one of the main efforts of Synthetic Biology is on building genetic systems in micro-organisms, not only because of their relative simplicity but as it is envisioned that small genetic circuits can potentially be used as a foundation for building more complex systems (Andrianantoandro *et al.*, 2006).

Although Synthetic Biology has been described as the 'Engineering of Biology', a systematic design cycle is still not realized to its full potential, limiting the advancement of the field in terms of functionality, reliability and size of the genetic systems (Purnick & Weiss, 2009). A design framework involves design specifications, modelling, conceptual and detailed design, as well as implementation and testing (Fig. 1). In Synthetic Biology, carrying out conceptual design (e.g. choosing the basic genetic system layout) is currently relatively simple due to the limited size of present-day synthetic genetic systems, but this will become more involved as more complicated systems can be built (Purnick & Weiss, 2009; Slusarczyk *et al.*, 2012). Similarly, methods are being developed to design modules for spatial organization of the cell (Chau *et al.*, 2012; Lim *et al.*, 2012), metabolic pathways and microbial communities (Shong *et al.*, 2012). At the same time, the present design framework needs to be improved with respect to how specifications, more detailed design and robust implementation are performed. An improved forward-engineering framework would consist of a mathematical model of the system chosen in the conceptual design stage,

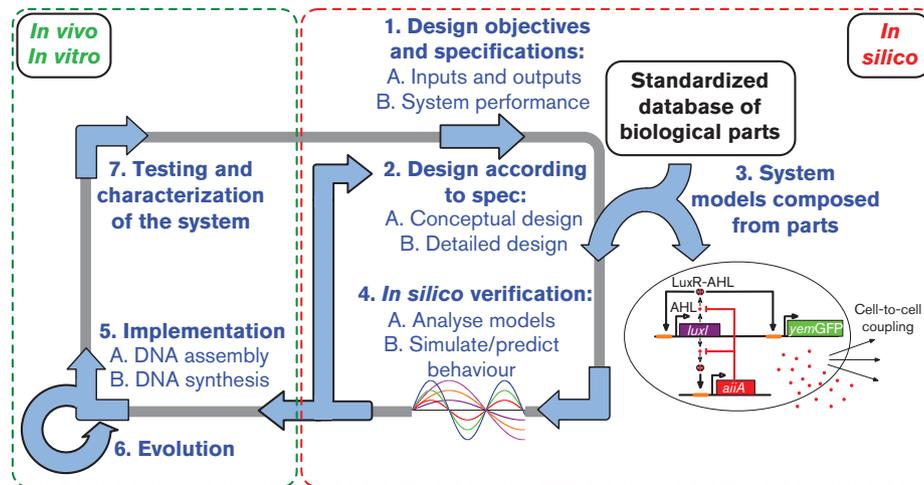


Fig. 1. A proposed forward engineering design cycle. Steps 1–4 take place *in silico* and follow a classical engineering design approach: specification, design, modelling and analysis. Steps 5, 6 and 7 take place in the laboratory where the system is assembled, might be evolved for tuned biological function, and is characterized. The cycle can be iterated if the design does not perform to the specifications. Adapted from MacDonald *et al.* (2011).

which can provide a basis for the design, construction, characterization and testing of the developed system. The parameters in this model can then be ‘tuned’ in a systematic manner in order to ensure that the resulting model meets the design specifications. The model with the chosen parameters and predicted performance can be built and its behaviour can then guide subsequent design, implementation and testing.

However, this is easier said than done. Indeed, when ‘tuning’ the different biological dials it is important to fully understand the relationship between specifications, model parameters, biological parts and implementation in order to carry out the design process. The dials used to redesign a biological system can include tuning global parameters or transcriptional, translational and post-translational parameters in the mathematical models. Experimentally this can be achieved by using different plasmid replicons for controlling gene copy number, different promoters to control the rate of transcription initiation, different ribosome-binding sites (RBSs), or different synonymous codons for controlling translation levels or degradation rates of all the species in the systems. The models used for the basic design of genetic systems often contain parameters representing multiple biological parts and so tuning a parameter in a mathematical model can be implemented experimentally in different ways. For example, in the simplest models of gene expression, one parameter often represents many different biological characteristics, such as a ‘transcription’ parameter representing promoter strength, transcription rate and gene copy number. Each of these has different tuning ranges, uncertainties and ease of experimental modification.

In this paper, we present a comprehensive review of ways in which the various parts of a biological system can be

modified systematically, focussing in particular on building genetic systems. We first discuss design and modelling of genetic systems, before reviewing in detail the typical dials that can be modified in a Synthetic Biology project. We then present various ways to tune these dials in order to achieve a desirable objective and show how tuning the parameters for each of these dials affects the output of a simple genetic system.

System design and parameter tuning

Synthetic Biology aims to be the ‘Engineering of Biology’, where an engineering design cycle is used to systematically improve existing biological systems and create new ones (Anderson *et al.*, 2012). A traditional engineering example is the design of a chemical plant. In this case specifications may include the concentrations of the final products, a conceptual design may determine the order of processes and reactions, while a more detailed design may set variables such as concentrations and flow rates in these processes, followed by further component details based on these variables such as sizes of pipes and reaction vessels (Perry & Green, 1999). Similarly, in a biological system, the specifications may be based on protein concentrations and their response characteristics, while a conceptual design determines the layout of a genetic system needed to achieve the specifications. A more detailed design may tune some of the parameters in the mathematical model(s), such as biochemical rate constants, followed by the design of individual biological parts fulfilling these parameters such as the design of a RBS to achieve a particular translation rate. In this framework, relevant models are developed and analysed at the different design stages in order to evaluate the candidate designs and predict whether they will meet

specifications. Once a detailed design is completed and verified, the system can be built and then tested to validate the design, with the design cycle repeated if at any stage the resulting functionality is not acceptable or requires improvement (RAEng, 2009).

The first step in the design of a genetic system is to specify its functionality for defined inputs and outputs. For example, the system may be required to act as a memory device or a switch (Gardner *et al.*, 2000) where the input is the concentration of an inducer and the output is the concentration of a protein. Performance specifications are required in order to determine whether the functionality is met for a particular design (Sen & Murray, 2012). These specifications can be composed of several metrics (Canton *et al.*, 2008; Del Vecchio *et al.*, 2008; Sen & Murray, 2012). For a switch, for example, there could be a requirement for the (time) mean protein concentration to be between set limits when the switch is 'on' or 'off'. Retroactivity specifications, or insulation, may also be required. This ensures that the functionality of the genetic system is not negatively affected when it is connected to another system upstream or downstream (Del Vecchio *et al.*, 2008). Another metric could be the response time: the time it takes for the protein concentration to switch from low to high with a change in inducer concentration (Canton *et al.*, 2008). Also, there may be requirements on the limits of variability or noise around the mean of a protein concentration level (Lestas *et al.*, 2010). Finally, a genetic system design should meet all performance metrics despite noise and uncertainty associated with the components and chassis of the system, as well as the uncertainty in cell size due to growth.

Once specifications are set, the design of a genetic system consists of a conceptual phase (e.g. determining genetic system topology) and then using appropriate models to complete a more detailed design. The latter involves determining model parameters to meet the design specifications set. In the conceptual phase, different system topologies can be used to obtain a desired behaviour, e.g., oscillators (Purcell *et al.*, 2010; Slusarczyk *et al.*, 2012; Strelkowa & Barahona, 2010), switches (Pfeuty & Kaneko, 2009) and adaptive systems (Ma *et al.*, 2009), and more complicated systems can be built to produce more advanced behaviour (Slusarczyk *et al.*, 2012). In the detailed design phase, a mathematical model must first be built and analysed. This model will guide the design but also be used for predicting whether a proposed design meets the required specifications. The same model will also be used for steps after design such as comparison with data from the testing phase of the engineering design cycle. Such models usually take the form of differential equations based on the biochemical reactions defining the designed system (Wilkinson, 2011). These differential equations can be deterministic or stochastic. The design proceeds by using standard optimization and control engineering approaches on the deterministic models to find the best parameter choice that achieves a desired objective. A combination of both simulations (Wilkinson, 2011) and

analytical methods (Murray, 2002; Tyson *et al.*, 2003) can then be used to verify the behaviour of the models. In particular, stochastic simulations are very useful in testing the variability of the system due to noise, and to ensure that stochastic effects do not substantially change the system behaviour for low biochemical species numbers (Tian & Kevin, 2006; Wilkinson, 2011). Once design parameters are selected, further models may be required for component design, such as for designing a RBS to match a tuneable parameter (Na *et al.*, 2010). The design process may need iteration, so that if no feasible choice of parameters for a particular system can meet the specifications, then a different topology can be used. Furthermore, once the system is implemented and tested, the process may need to be iterated. In particular, further detailed design or 'tuning of the dials' may be necessary for the circuit to function and meet specifications.

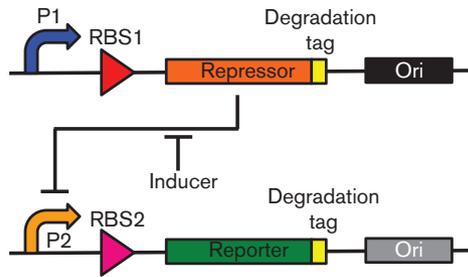
Discussion of the genetic system design leads naturally to the question of implementation, the main focus of this review. Which biological components should be modified in order to implement different genetic systems?

A simple genetic system

Before we discuss this question, let us consider a simple, illustrative genetic system and its associated model as an example (Fig. 2). This system will be used throughout the paper to illustrate the engineering design cycle and how 'tuning dials' can be performed during the design and implementation stages. The example genetic system contains a constitutive promoter driving the expression of a repressor protein, which in turn represses the expression of a reporter gene from a regulated promoter. The measured output of the system is the concentration of the reporter protein while the input is the concentration of an inducer, which binds to the repressor protein thereby sequestering it away and allowing transcription initiation. The biochemical equations used to model this system are shown in Fig. 2.

The biochemical equations are the mathematical description of the underlying biochemical reactions of the system. From a biological viewpoint, the reactions that must be described are: transcription, translation, repressor–promoter and repressor–inducer interactions, and degradation of species within the system. Equations (1) and (2) describe RNA polymerase binding to a promoter followed by transcription initiation for the repressor and reporter genes, respectively. Initiation of transcription is a reversible reaction (as denoted by the double arrows and forward and reverse reaction rate constants in the equations), whereas extension is considered to be irreversible. Equation (3) is included to reflect the biological reality that most promoters have some basal level of transcription in the absence of an inducer (also called leakiness). Taken together, these equations describe the generation of mRNA species in the system.

Equations (4) and (5) describe the binding of ribosomes to a RBS on mRNA, before translation is initiated for the

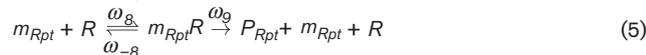
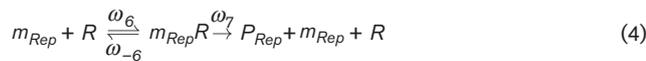


Biochemical reactions:

Transcription



Translation



Repressor dimerization



Repressor-operator binding



Inducer-repressor binding



Degradation

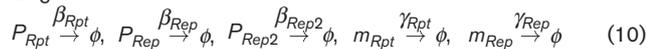


Fig. 2. A simple genetic network with model equations. A schematic representation of a simple genetic network highlighting dials that can be tuned experimentally; promoters (P1 and P2: blue and pale orange arrows), ribosome-binding sites (RBS1 and RBS2: red and pink triangles), repressor gene (orange rectangle), reporter gene (green rectangle), protein degradation tags (yellow) and origins of replication (black and grey). Expression of the repressor results in repression of P2 whilst addition of inducer activates P2. Biochemical reactions: a set of biological reactions describing the processes that take place in the simple genetic network, where g_{Con} is the constitutive promoter driving expression of the repressor gene, g_{Rep} is the regulated promoter driving expression of the reporter gene, Pol is the RNA polymerase, m_{Rpt} , m_{Rep} are the reporter and repressor mRNA, P_{Rpt} , P_{Rep} are the reporter and repressor proteins, I is the inducer, R is the ribosome, and all other species are complexes resulting from biochemical interactions between the above listed elements. The symbol ϕ denotes degradation of a species.

repressor and reporter, respectively. These steps have been described as two reactions so that the RBS strength (which is easy to modify experimentally – see below) can be

accounted for separately from the translation rate, which is usually taken as a constant number of amino acids per unit time. Equations (4) and (5) together describe the rate of generation of protein species in the system.

The interactions of the repressor with the promoter and the inducer control the number of free promoters available for RNA polymerase binding. These interactions are described in equations (6)–(9). Equation (6) describes dimerization of the repressor protein, based in this example on TetR, to produce its functional form, which is capable of binding the operator region of a promoter and repressing transcription. Other repressors form different functional multimers (e.g. LacI acts as a tetramer) and would require additional equations to reflect the further multimerization steps where necessary. Equation (7) describes the binding of the functional repressor protein to the operator, while equation (8) describes inducer binding to a repressor that is already bound to an operator, followed by dissociation of the inducer-repressor complex from the operator, allowing transcription to proceed.

Finally, equation (10) describes the degradation of the mRNA and protein species in the system. The degradation contributes to the steady state concentration of the species by ensuring its removal.

From this set of biochemical reactions, mass-action kinetics can be used to produce a deterministic model from the biochemical equations (Cornish-Bowden, 2004) while the chemical master equation can be used for a stochastic model (Gillespie, 1992). For the deterministic model, the mass-action kinetics can be used to describe the different reaction rates, while differential equations describe the rates of change of the concentrations due to the reactions. For the stochastic model, the equations describe the probability of a reaction occurring, e.g. an increased reaction constant makes that reaction more likely to occur at higher concentrations.

Characterizing dials

What sets Synthetic Biology apart from traditional molecular biology is that an engineering design approach is taken to systematically design or redesign a biological system, for quantifiably improved or different functionality. This assumes that a design objective is set, against which the performance of the synthetic component will be quantified (Anderson *et al.*, 2012). To achieve the design objective, a set of biological ‘dials’ need to be tuned. In this section, we describe the set of possible dials available to the designer, and how they can be described mathematically.

Global parameters

Chassis. When designing, constructing and characterizing simple single gene expression constructs through to complex genetic networks, it is important to consider the impact of strain-to-strain variation of the behaviour of

these networks. It has been shown that the abundance of RNA polymerases and ribosomes is dependent on cellular growth rates, affecting downstream processes such as transcription rate, translation rate, gene copy number, mRNA degradation, protein dilution, protein degradation and cell mass (Klumpp *et al.*, 2009). Regulated processes such as promoter repression or activation have also shown growth rate dependence, affecting genetic networks as a result (Klumpp *et al.*, 2009; Scott *et al.*, 2010; Tan *et al.*, 2009). There are a wide range of strains and organisms that can be used to harbour synthetic genetic networks, and in some cases, the networks work predictably across different strains (Prindle *et al.*, 2012) whilst in other cases the behaviour of the network can be drastically altered by changing the host cell (Egbert & Klavins, 2012). In this latter reference, two strains of *E. coli* were used that both contained a LacI Δ mutation but with slightly different genotypes, exemplifying that even similar strains can have profound effects on network function. Some tuneable elements may also be strain specific and whilst they function in one strain they may be inactive in another, e.g. the DIAL strains used to modulate plasmid copy number (Kittleson *et al.*, 2011).

Gene copy number. The number of gene copies can be used to increase or decrease the amount of available protein over a partially linear range. Gene copy number can be controlled either by changing the origin of replication of a plasmid-bound gene of interest or by increasing the number of chromosomally integrated copies. Plasmid copy number has been shown to play a key role in the behaviour of genetic networks (Atkinson *et al.*, 2003; Mileyko *et al.*, 2008). The range of accessible values is limited to a maximum value that does not saturate the cellular transcription and translation machinery and does not cause undue metabolic burden on cells ('run-away' replication) (Nordström & Uhlin, 1992).

Gene copy number is not a continuous number. Chromosomal integration has been shown to tolerate up to five copies of the same gene (Choi *et al.*, 2006), although this could potentially be increased further, and plasmids have a few discrete values that can be accessed (Table 1). Engineering endeavours have produced plasmids with inducible copy number that can be controlled either by the binding of a ligand (Panayotatos, 1984) or through changes in temperature (Sternberg, 1990), allowing dynamic shifts in copy number to be used as a design variable. Alternatively, multiple bacterial strains have been developed (DIAL strains) that maintain the same plasmid at different steady state copy numbers (Kittleson *et al.*, 2011). These techniques give another level of control and tuneability of plasmid copy number in genetic systems.

The potential to maintain multiple plasmids, encoding different components from genetic networks, at different copy numbers within a cell is also possible. This is, however, dependent on the incompatibility group of the plasmid (Table 1) (Tolia & Joshua-Tor, 2006). In addition,

multi-copy plasmids are maintained at a mean value in each cell, leading to population variations (Ebersbach & Gerdes, 2005) that will need to be accounted for in modelling efforts. Although plasmid copy numbers have been described here, they are dependent on the context within which they are used. The copy number of plasmids has been shown to have an inverse relationship to the plasmid size (Ebersbach & Gerdes, 2005; Zhong *et al.*, 2011) and cellular growth rate (Klumpp, 2011), and their maintenance within a cell is dependent on the toxicity of the genes encoded on the plasmid (Mileyko *et al.*, 2008).

Transcription level design

Promoter type. Promoters are regions of DNA containing consensus sequences for the recruitment of the transcriptional machinery (sigma factors; RNA polymerase, RNAP) to initiate transcription (Gruber & Gross, 2003). Core promoter sequences are recognized by different sigma factors which are active under different environmental conditions, for example, *Escherichia coli* σ^{70} , σ^S , σ^{32} and σ^{54} (Table 2) (Gruber & Gross, 2003). The number of σ factors used in bacteria can vary from 1 (in *Mycoplasma genitalium*) to 63 (in *Streptomyces coelicolor*) and provides an intricate way for organisms to regulate gene expression in response to specific environmental conditions. Promoters can also be constitutively active or regulated (Fig. 3). Constitutive promoters are active without the need for any transcription factors, whilst regulated promoters have operator-binding sites for repressors or activators that block or assist RNAP binding respectively in the presence of a small molecule (inducer) or under certain environmental conditions (Lloyd *et al.*, 2001) (Table 2). Combinatorial promoter design has been implemented to generate hybrid promoters (also sometimes called logic gates) that are conditionally activated in the presence of multiple inducer signals (Fig. 3) (Cox *et al.*, 2007).

Inducer concentration. Each transcriptional repressor/activator will respond to one or more small molecules known as inducers. There are natural inducers (e.g. allolactose for the Lac repressor (Lewis *et al.*, 1996) or tetracycline for the Tet repressor (Orth *et al.*, 2000)), and in some instances non-metabolizable chemical analogues that cause gratuitous induction (e.g. isopropyl- β -thiogalactoside, IPTG, for the Lac repressor (Lewis *et al.*, 1996) or anhydrotetracycline, aTc, for the Tet repressor (Lederer *et al.*, 1996)). The advantage of the chemical analogues is that their concentration level remains roughly constant. The level of transcription follows a sigmoidal response to the inducer concentration, which, over a certain range, can be approximated as linear (Table 3). Often the slope of this linear approximation is very large, which may make tuning difficult. Mutations in the small molecule binding site of the repressor could shift the range over which the response is linear (Satya Lakshmi & Rao, 2008), adding further control.

Table 1. Plasmid copy number and plasmid incompatibility groups

Plasmid incompatibility groups are highlighted.

Plasmid	Replicon	Copy number*	Reference
ColE1	ColE1	15–20	Schmidt & Inselburg (1982); Shizuya <i>et al.</i> (1992)
pBR322	pMB1	15–20	Balbás <i>et al.</i> (2013); Lin-Chao <i>et al.</i> (1992); Sternberg (1990)
pBR3722	pMB1	30–40†	Hakkaart <i>et al.</i> (1985); Kool & Nijkamp (1974); Lin-Chao <i>et al.</i> (1992)
pUC	pMB1	>200†	Figurski <i>et al.</i> (1979); Lin-Chao <i>et al.</i> (1992)
pACYC	p15A	18–22	Chang & Cohen (1978); Hasunuma & Sekiguchi (1979)
F1	F1	1–2	Peterson & Phillips (2008); Shizuya <i>et al.</i> (1992)
pNS358	P1 lytic <i>ori</i>	1–>25	Peterson & Phillips (2008); Sternberg (1990)
pCDF	CloDF13	10–90	Hakkaart <i>et al.</i> (1985); Kool & Nijkamp (1974); Kües & Stahl (1989)
RK2	RK2	4–7	Figurski <i>et al.</i> (1979); Kittleson <i>et al.</i> (2011)
pSC101	pSC101	~5	Hasunuma & Sekiguchi (1979); Kittleson <i>et al.</i> (2011)
pJPA12	pSC101	27	Deuschle <i>et al.</i> (1986); Gruber & Gross (2003); Peterson & Phillips (2008)
pJPA13	pSC101	~240	Becker & Hengge-Aronis (2001); Peterson & Phillips (2008)
pRSF	RSF1030	10–60	Grossman <i>et al.</i> (1987); Kües & Stahl (1989); Wang & deHaseth (2003)
pBjk2992-jtk2541	ColE2	1–60	Bernardo <i>et al.</i> (2009); Buck <i>et al.</i> (2000); Kittleson <i>et al.</i> (2011)
pBjk2993-jtk2541	R6K	5–250	Kittleson <i>et al.</i> (2011); Schleif (2000)

*Plasmid copy number is dependent on plasmid size, gene toxicity and, in some instances (indicated by †), by temperature.

Promoter leakiness and basal expression. The function of promoters is often thought of encompassing two states, either on or off. However, this is very rarely the case, with transcription initiation still occurring even in the repressed state (Lanzer & Bujard, 1988). Often this behaviour is not desirable for inducible promoters. However, it can be beneficial if a particular level of basal expression is required in the repressed state but can be further increased with the addition of an inducer. The leakiness of a promoter can be tuned through the location (Lanzer & Bujard, 1988; Murphy *et al.*, 2007), number (Murphy *et al.*, 2007) and sequence (Alper *et al.*, 2005; Lanzer & Bujard, 1988) of the operator binding sites (Fig. 4) and core promoter elements. A combinatorial TetR repressed promoter library has been described previously with varying operator multiplicity and location resulting in promoters with leaky expression

in the repressed state spanning more than two orders of magnitude (Murphy *et al.*, 2007). The strength of repressor binding to its operator sites can affect promoter leakiness and can be tuned through mutation of the operator DNA sequence or the DNA-binding domain of the repressor protein (Nichols & Matthews, 1997; Penumetcha *et al.*, 2010). Another way with which basal transcription can be tuned is through the co-expression of the target gene from a constitutive promoter as well as a regulated promoter, although this will also result in an increase in gene expression level in the induced state.

Promoter strength. Promoter strength can be tuned by creating promoter libraries randomized in the RNAP-binding regions affecting RNAP-binding affinity (Alper *et al.*, 2005; Brewster *et al.*, 2012; Jensen & Hammer, 1998),

Table 2. Constitutive, negative, positive and hybrid regulated promoter types

Promoter	σ factor	Regulatory type*	Regulator	Inducer	Reference
P _{bla}	σ_{70}	Con.	NA	NA	Deuschle <i>et al.</i> (1986); Gruber & Gross (2003)
csiD	σ_S	Con.	NA	Starvation/hyperosmolarity	Becker & Hengge-Aronis (2001)
P _{GroE}	σ_{32}	Con.	NA	Heat shock/nutrient starvation	Grossman <i>et al.</i> (1987); Wang & DeHaseth (2003)
P _o	σ_{54}	Con.	NA	Nitrogen/nutrient starvation	Bernardo <i>et al.</i> (2009); Buck <i>et al.</i> (2000)
P _{Bad}	σ_{70}	Pos.	AraC	Arabinose	Schleif (2000)
P _{LlacO-1}	σ_{70}	Neg.	LacR	IPTG	Lutz & Bujard (1997)
P _{LtetO-1}	σ_{70}	Neg.	TetR	aTc	Lutz & Bujard (1997)
P _{lac/ara-1}	σ_{70}	Neg./Pos.	LacR/AraC	IPTG/arabinose	Lutz & Bujard (1997)

*Con., constitutive; Pos., positive regulation; Neg., negative regulation.

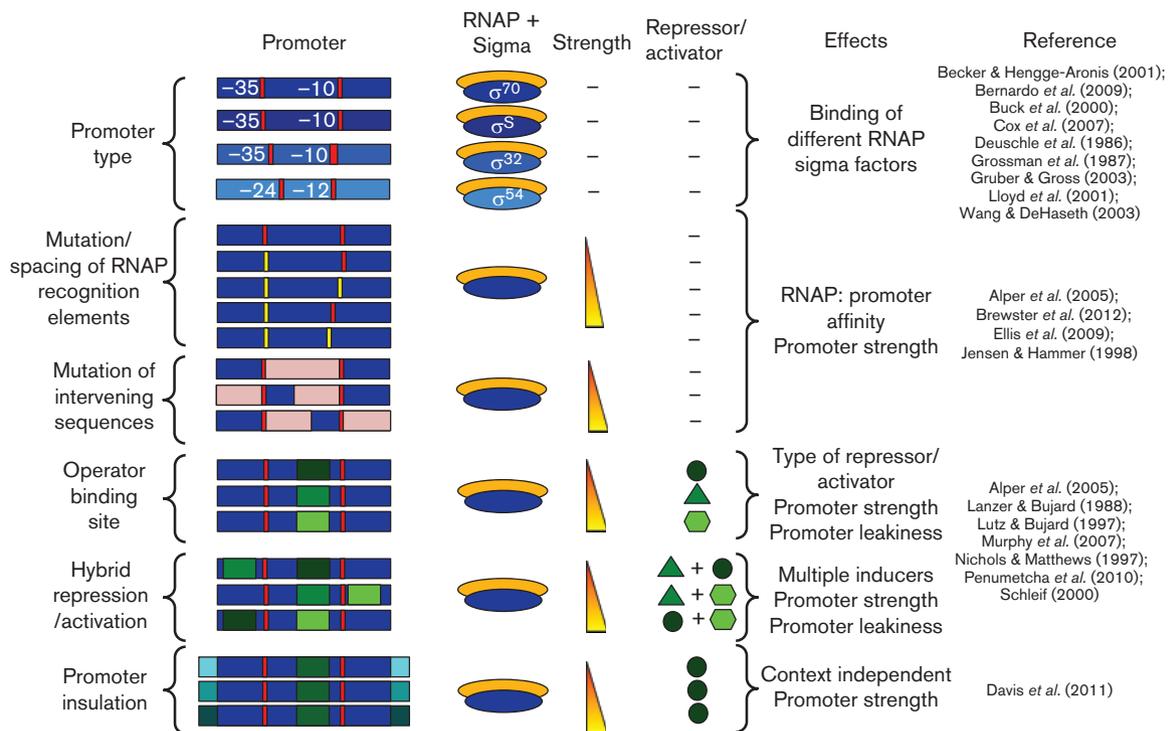


Fig. 3. The effect of promoter architecture on promoter strength, regulation and basal transcription (leakiness). A schematic representation of different promoter architectures that can modulate the strength, regulation type and leakiness of a promoter by changing the core sigma factor (orange and blue ovals), binding sequences (blue rectangles with -35 and -10 or -24 and -12 recognition sequences in red and yellow), intervening sequences (pink), operator recognition sequences, multiplicity and location (green rectangles), repressor/activator types (green circles, triangles or hexagons) or flanking sequences (shades of turquoise).

in the operator region, which changes the strength of the interaction of the repressor/activator with the DNA (Alper *et al.*, 2005), or in the DNA sequences flanking the promoter, reducing any upstream/downstream context dependent effects on promoter strength (Davis *et al.*, 2011) (Fig. 4). Mutation of the DNA sequences between important binding motifs has also shown that a linear variation in promoter strength can be achieved (Ellis *et al.*, 2009; Jensen & Hammer, 1998). Although discrete values

will be obtained through the creation of a randomized promoter library, the promoter strength can typically be treated as a continuous variable for modelling purposes. Randomized promoter libraries have elicited numerous constitutive and regulated promoters with activities spanning over two orders of magnitude (Alper *et al.*, 2005; Davis *et al.*, 2011; Ellis *et al.*, 2009; Jensen & Hammer, 1998). A limitation in the use of these promoters as predictable parts is the lack of a standardized promoter

Table 3. Inducers and their working conditions

Promoter type	Repressor	Inducer	Working condition	Reference
Lac	LacI	IPTG	10 ⁻⁴ -10 ⁻³ M	Lutz & Bujard (1997)
Lac	LacI*	IPTG	10 ⁻⁵ -10 ⁻⁴ M	Satya Lakshmi & Rao (2008)
Lac	LacI*	Temperature	30-40 °C	Yabuta <i>et al.</i> (1995)
Tet	TetR	aTc	5-50 ng ml ⁻¹	Lutz & Bujard (1997)
P _{BAD}	AraC	Arabinose	10 ⁻⁴ -10 ⁻¹ M	Mirasoli <i>et al.</i> (2002)
P _L /P _R	cI 857	Temperature	30-42 °C	Menart <i>et al.</i> (2003); Villaverde <i>et al.</i> (1993)
P _{Lux}	LuxR	AHL	10 ⁻⁹ -10 ⁻⁶ M	Koch (2005); Urbanowski <i>et al.</i> (2004)

*LacI mutants developed to respond to decreased IPTG concentration or alternative stimulus.

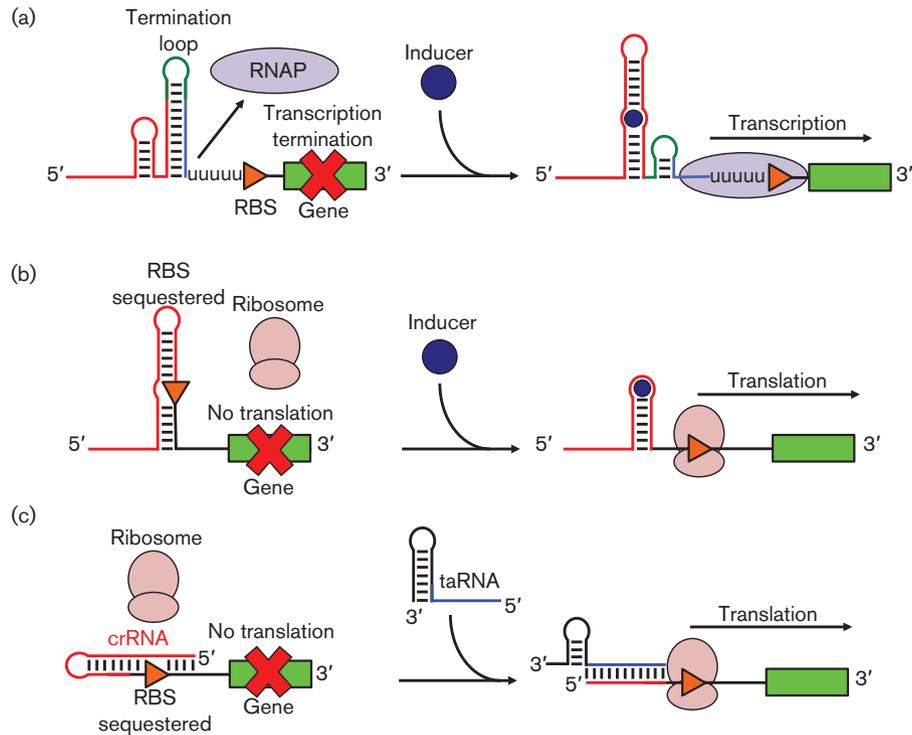


Fig. 4. Transcriptional and translational control by riboregulators. A schematic representation of transcriptional control by a riboswitch (a), and translational control by a riboswitch (b) or a *trans*-activating RNA (taRNA) (c).

strength metric. Promoters can often perform differently from how their original characterization would suggest, due to differences in experimental conditions and measurement equipment. Therefore predicting the behaviour of a gene regulatory network component such as a promoter across different laboratories can be difficult. The need for a promoter strength metric for the accurate comparison of promoters produced from different libraries, experimental conditions and laboratories has resulted in the development of a technique to standardize a promoter strength with respect to a reference promoter, and quantifying this relative strength in terms of relative promoter units (Kelly *et al.*, 2009).

Placement of genes in a multi-gene construct or operon. The length of time it takes to transcribe a gene depends upon the gene length (RNAP transcribes at a rate of $\sim 50 \text{ bp s}^{-1}$ (Gotta *et al.*, 1991; Vogel & Jensen, 1994)) and its distance from the transcription initiation site. Therefore, by placing one or more genes in front of the gene of interest it is possible to introduce a delay in the transcription process. It is important to note, however, that there may be unintended effects on the translation process if the secondary structure of the mRNA molecule produced causes ribosomes to disengage before translation of the gene of interest, or if secondary structure elements are introduced which alter the stability of the mRNA molecule (Carrier & Keasling, 1997a; Newbury *et al.*,

1987). In principle, this transcription delay increases linearly with the length of the superfluous genes added in front of the gene of interest and can be approximated as a continuous variable – although, strictly speaking, this is a discrete variable whose values are multiples of the time it takes to transcribe a single base (though very long mRNA constructs will tend to have larger translational effects). An increase in the length of a transcript also has a positive influence on the amount of translation from the first gene in an operon (Lim *et al.*, 2011). This is due to the fact that transcription and translation take place simultaneously in prokaryotes. Therefore, the first genes in an operon have a longer period for translation during transcription before RNAP dissociation and mRNA degradation (Lim *et al.*, 2011).

Translation level design

Ribosome-binding site (RBS) strength. The efficiency of translation initiation depends on the recruitment of ribosomes to the Shine–Dalgarno sequence, which can be stronger or weaker depending on variations in this sequence and the distance from the translation initiation codon (Chen *et al.*, 1994; Osterman *et al.*, 2013). The RBS strength is also dependent on the upstream (Komarova *et al.*, 2005) and downstream mRNA sequence (Salis *et al.*, 2009) due to the formation of local secondary structures that can influence or inhibit ribosome binding. Prediction

of the strength of prokaryotic Shine–Dalgarno sequences can therefore be facilitated by the use of Chris Voigt's simulation prediction program (RBS calculator) (Salis *et al.*, 2009) or Gyoo Yeol Jung's UTR Designer (Seo *et al.*, 2013). Over 100 predicted RBSs have been experimentally tested showing that the translation initiation rate can be controlled over a 100 000-fold range (Salis *et al.*, 2009). The Ouyang lab used the RBS calculator to design RBSs with predicted strengths for use in a predetermined bistable toggle switch exemplifying the usefulness of this tool (Chen *et al.*, 2012). Fine-tuning of a genetic toggle switch has also been demonstrated by altering the length of the spacer between the Shine–Dalgarno sequence and the start codon (Egbert & Klavins, 2012). Comparisons of experimental data with RBS calculator predictions were in fairly good agreement dependent on the spacer sequence makeup (Egbert & Klavins, 2012).

Codon optimization. Because of the degeneracy of the genetic code, it is possible to create mRNA transcripts with differing sequence that encode the same protein, eliminating rare codons and increasing translational efficiency. An altered coding sequence can also contribute to different mRNA secondary structures and, therefore, translational efficiency. Whilst standard codon optimization techniques aim to maximize protein production through using the most abundant codons observed for highly expressing native host proteins (codon adaptation index, CAI) (Angov, 2011), this technique does not take into account several factors that influence translational efficiency: translational pausing (Angov, 2011), local mRNA secondary structure (Kudla *et al.*, 2009) and tRNA abundances (Welch *et al.*, 2009). Kudla *et al.* have shown a correlation between codon optimization and the secondary structure of the mRNA at the beginning of a gene (regions -4 to $+37$) with the translational efficiency in *E. coli*, with a 250-fold variation in GFP expression across the constructs they tested (Kudla *et al.*, 2009). Progress has been made in predictive algorithms that take into account codon usage and tRNA abundance to optimize a gene's coding sequence to give a desired translation efficiency (Welch *et al.*, 2009). This codon optimization algorithm could potentially be combined with RNA secondary structure prediction programs in order to facilitate a more accurate prediction in the resulting efficiency of translation.

mRNA decay rate. The longevity of the mRNA transcript is controlled by its secondary structure in the untranslated regions, which protect it (Bouvet & Belasco, 1992; Carrier & Keasling, 1997b; Mackie, 2012) or make it more vulnerable (Bouvet & Belasco, 1992) to degradation by RNases, and through efficient binding and translation by ribosomes blocking RNase action (Carrier & Keasling, 1997b; Komarova *et al.*, 2005; Osterman *et al.*, 2013). The half-life for most mRNAs in *E. coli* is relatively short at ~ 1 – 2 min (Mackie, 2012). The longer-lived an mRNA molecule is, the more translation will occur from each transcript. Appending 5' stem–loop structures of varying

sizes to mRNAs has been shown to increase the mRNA half-life between 5- and 10-fold up to a half life of ~ 24 min (Arnold *et al.*, 1998; Hansen *et al.*, 1994). Appending 3' REP sequences or insertion of REP sequences into intercistronic regions of polycistronic operons can also stabilize upstream mRNA transcripts by ~ 3 -fold (Newbury *et al.*, 1987).

Riboregulators. Riboswitches are RNA genetic control elements that modulate gene expression in response to an inducer molecule (Vitreschak, 2004) or transacting RNA (taRNA) (Isaacs *et al.*, 2004) without the requirement of any RNA–protein interactions. Since their discovery, a number of synthetic riboswitches have been developed that control gene expression by either premature transcriptional termination (Wachsmuth *et al.*, 2013) or by translational inhibition by sequestering RBSs (Dixon *et al.*, 2010; Lynch *et al.*, 2007; Topp *et al.*, 2010) in a dose-responsive manner to specific inducers (Fig. 4). Riboswitches that control premature transcription termination have been shown to elicit up to a 3-fold change in transcription in response to an inducer (Wachsmuth *et al.*, 2013), whilst riboswitches that modulate translation initiation have been developed that span a 2- to 150-fold range in response to an inducer. A model-directed redesign of a translational riboswitch has also been used to predictively adjust its efficiency (Beisel & Smolke, 2009). The taRNA riboregulators work by the binding of the taRNA to a *cis*-repressed mRNA (crRNA) resulting in the release of the RBS, allowing translation initiation (Isaacs *et al.*, 2004) (Fig. 4). taRNA riboregulators have been utilized in controlling a metabolic pathway and showed a ~ 1 - to 200-fold increase in translation initiation in the presence of the trRNAs (Callura *et al.*, 2012; Isaacs *et al.*, 2004). Whilst the riboregulators described here do not require RNA–protein interactions for their function, the CRISPRi platform for transcriptional repression utilizes ribonucleoproteins (Qi *et al.*, 2013). Briefly, a small guide RNA (sgRNA) is expressed with complementary base pairing to a target DNA sequence and a secondary structural stem–loop that is recognized by a catalytically inactive RNA-binding protein, Cas9. Together the sgRNA-Cas9 ribonucleoprotein binds the target DNA sequence and inhibits initiation of transcription, elongation or transcription factor binding depending on where the sgRNA is targeted (Qi *et al.*, 2013).

Transcriptional, translational and post-translational design

Inteins. Inteins are the protein-splicing equivalents of introns found in eukaryotic pre-mRNAs. An intein is a genetically encoded element within a target gene and is transcribed and translated together with the target protein before it undergoes autocatalytic self-excision and splicing of the target protein exteins (Gogarten *et al.*, 2002) (Fig. 5). Inteins, therefore, work at both a transcriptional and translational level by increasing the time it takes to

transcribe and translate a target gene. Bacterial inteins range in size from 36 to 1986 amino acids (Perler, 2002), theoretically increasing transcription by ~2–120 s and translation by ~2–100 s. A comprehensive list of inteins and their sizes can be found at http://tools.neb.com/inbase/list_prop.php. Inteins would be beneficial for engineering delay into genetic networks, in particular tuning oscillators that depend on transcriptional and translational delay for their function (Mather *et al.*, 2009; Purcell *et al.*, 2010; Stricker *et al.*, 2008). Split inteins have also been described where the intein domain is transcribed and translated by two separate genes and the resulting proteins can undergo *trans*-splicing to produce a single functional protein (Elleuche & Pöggeler, 2010; Lockless & Muir, 2009) (Fig. 5). Post-translational control of inteins has been engineered to modulate intein splicing (Skretas, 2005), *trans*-splicing (Mootz & Muir, 2002) or a combination of both (Shi & Muir, 2005) in the presence of small molecule inducers (Fig. 5).

Protein level design

Protein degradation. The longevity of proteins within a cell can be tuned by addition of degradation tags at the N (Gur & Sauer, 2008) or the C terminus (Andersen *et al.*, 1998; Flynn *et al.*, 2001; Gur & Sauer, 2008; McGinness *et al.*, 2006) as well as internally (Gur & Sauer, 2008). Different degradation tags can be used to target proteins for degradation to one of several cellular degradation complexes (Kirstein *et al.*, 2009) (Table 4). Modifying the N-terminal residues of a protein can also target it for proteolytic degradation by the N-End rule mechanism (Erbse *et al.*, 2006; Kirstein *et al.*, 2009). Recent work has resulted in degradation tags of varying sequences that tune the rate of degradation by the ClpXP/AP degradation

complexes (Andersen *et al.*, 1998; Dougan *et al.*, 2002; Flynn *et al.*, 2001; Hoskins *et al.*, 2000; McGinness *et al.*, 2006; Purcell *et al.*, 2012; Wang *et al.*, 2007). However, Cookson *et al.* have shown that an abundance of protein targeted to the ClpXP machinery can lead to a queuing effect, which, in turn, leads to a slower rate of protein degradation that is dependent on the overall concentration of tagged species in the system (Cookson *et al.*, 2011). This can be detrimental to genetic network behaviour when a fast turnaround of network elements is required but can also be beneficial in coupling separate networks through the queuing effect (Cookson *et al.*, 2011). By utilizing multiple degradation pathways, the queuing effect could potentially be suppressed.

Protein activity. Manipulation of protein activity through point mutations can be used as a means of control with a few discrete values. If the system consists entirely of genetic elements (repressors or activators), then modification of the DNA-binding affinity is best achieved by manipulating the DNA sequence, rather than attempting to mutate the protein. However, for enzymic activities, mutants with altered substrate specificity (Brannigan & Wilkinson, 2002; Wilson & Agard, 1991), kinetics (Brannigan & Wilkinson, 2002) or thermostability (Lehmann & Wyss, 2001) may already be available or can sometimes be created via protein engineering (Brannigan & Wilkinson, 2002; Lehmann & Wyss, 2001; Wilson & Agard, 1991).

Extension to eukaryotic dials

The majority of the dials described above can also be used in eukaryotes. However, there are additional dials that can be used to tune genetic networks when working in eukaryotes, as follows.

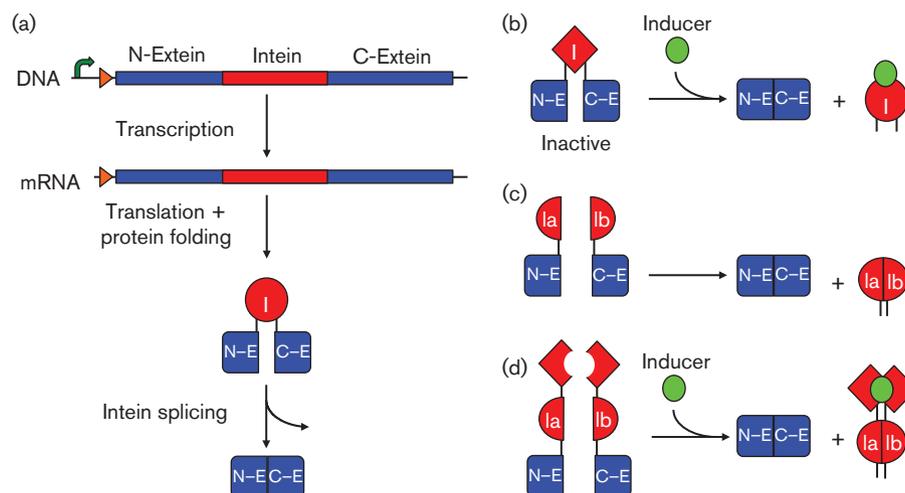


Fig. 5. Mechanism and types of inteins. Schematic representation of the mechanism of protein splicing (a) and different intein (I) types. Small molecule (green) induced splicing (b), split intein (Ia/Ib) splicing (c) and small molecule induced split intein splicing (d).

Table 4. Protein degradation pathways

Protease	Cellular localization	Degradation tag/ mechanism	Tag position	Approx. half-life (min)*	Reference
ClpXP	Cytoplasm	SsrA	C-terminus	10–110	Andersen <i>et al.</i> (1998); Flynn <i>et al.</i> (2001); McGinness <i>et al.</i> (2006); Purcell <i>et al.</i> (2012)
ClpAP	Cytoplasm	SsrA/Rep	C-terminus	10–110	Dougan <i>et al.</i> (2002); Flynn <i>et al.</i> (2001); Hoskins <i>et al.</i> (2000)
ClpAPS	Cytoplasm	N-End rule	N-terminus	2–50	Erbse <i>et al.</i> (2006); Wang <i>et al.</i> (2007)
FtsH	Inner membrane	SsrA/non-polar pentapeptide	C-terminus	2	Herman <i>et al.</i> (1998)
Tsp	Periplasm	SsrA	C-terminus	>15	Parsell <i>et al.</i> (1990); Silber <i>et al.</i> (1992)
Lon	Cytoplasm	β 20	N-terminus, internal, C-terminus	>10	Gur & Sauer (2008)

*Approximate protein half-life is dependent on protein size, stability and temperature.

Post-transcriptional modification of mRNA (splicing). In higher eukaryotic chassis, RNA splicing can be used to introduce a delay between transcription and translation (Swinburne *et al.*, 2008). Transcripts of mRNA that have no introns will not need to be spliced, whereas those that have increasing numbers of introns will require longer transcription and processing time before translation. Splicing of mRNA can also be controlled by the addition of morpholinos (synthetic molecules that base pair with target DNA sequences) to block the pre-mRNA protein splicing machinery (Morcos, 2007). In prokaryotes, group I and II self-splicing introns have been identified, but these appear to have no known specific biological function and are thought to be remnants from an ancient RNA world (Raghavan & Minnick, 2009). Thus, with the current state of understanding, it seems they would not be very easy to target as tuneable dials at the moment.

Translocation. In eukaryotes, transcription occurs in the nucleus and the resulting transcripts must then be translocated to the cytoplasm for translation (Oeffinger & Zenklusen, 2012). This will introduce a delay between transcription and translation (Grünwald & Singer, 2010). Therefore, moving a dial from a prokaryotic chassis to a eukaryotic chassis could be used as a means to introduce delay into a genetic network.

Protein trafficking. In eukaryotic chassis, signal sequences are used to localize proteins to different subcellular compartments/organelles (Alberts *et al.*, 2002; Martoglio & Dobberstein, 1998). Adding a signal sequence to a protein will introduce a delay between translation and function due to the time required for transport of the protein to its intended destination (Hirschberg & Lippincott-Schwartz, 1999). This can also be used as a means to increase the effective concentration of a protein by accumulating it in a smaller volume. Protein trafficking from the cytoplasm to the inner/outer membrane and to

the periplasmic space also takes place in prokaryotes, although far fewer subcellular compartments are available for sampling in prokaryotes (Driessen & Nouwen, 2008; Papanikou *et al.*, 2007).

Discussion and perspectives

Tuning parameters in mathematical models will modify the system behaviour; in this way, design specifications can be met. It is often the case, however, that a single design objective can be met via different parameter changes. For example, modifying the gene copy number, promoter strength, RBS strength, mRNA degradation rate or protein degradation rate can all change the steady state protein concentration expressed from a single gene (Fig. 6). To change the dynamics of protein accumulation, the degradation rates of the mRNA and protein can both be modified (Fig. 6) (Alon, 2007). It is, therefore, important to understand how different ‘dials’ affect the variability and noise of the system. For example, low transcription levels and high translation levels will produce a higher variability in protein concentration than high transcription and low translation levels with the same mean protein concentration at steady state (Raj & van Oudenaarden, 2008) (Fig. 6). For the case of high output variability, increasing promoter strength and decreasing RBS strength can be used to decrease variability in the output while keeping the mean output the same. Changing parameters can also modify system behaviour with regard to the output. As an example, negative feedback, either through auto-regulation or indirectly through intermediary genes in the system (Cosentino & Bates, 2012), can be used to improve robustness, reduce output variability and reduce response times of a system. Negative feedback often has trade-offs in the design and there are theoretical limits as to how much negative feedback can improve performance of genetic systems (Lestas *et al.*, 2010), which should be taken into account when tuning parameters. While simple design

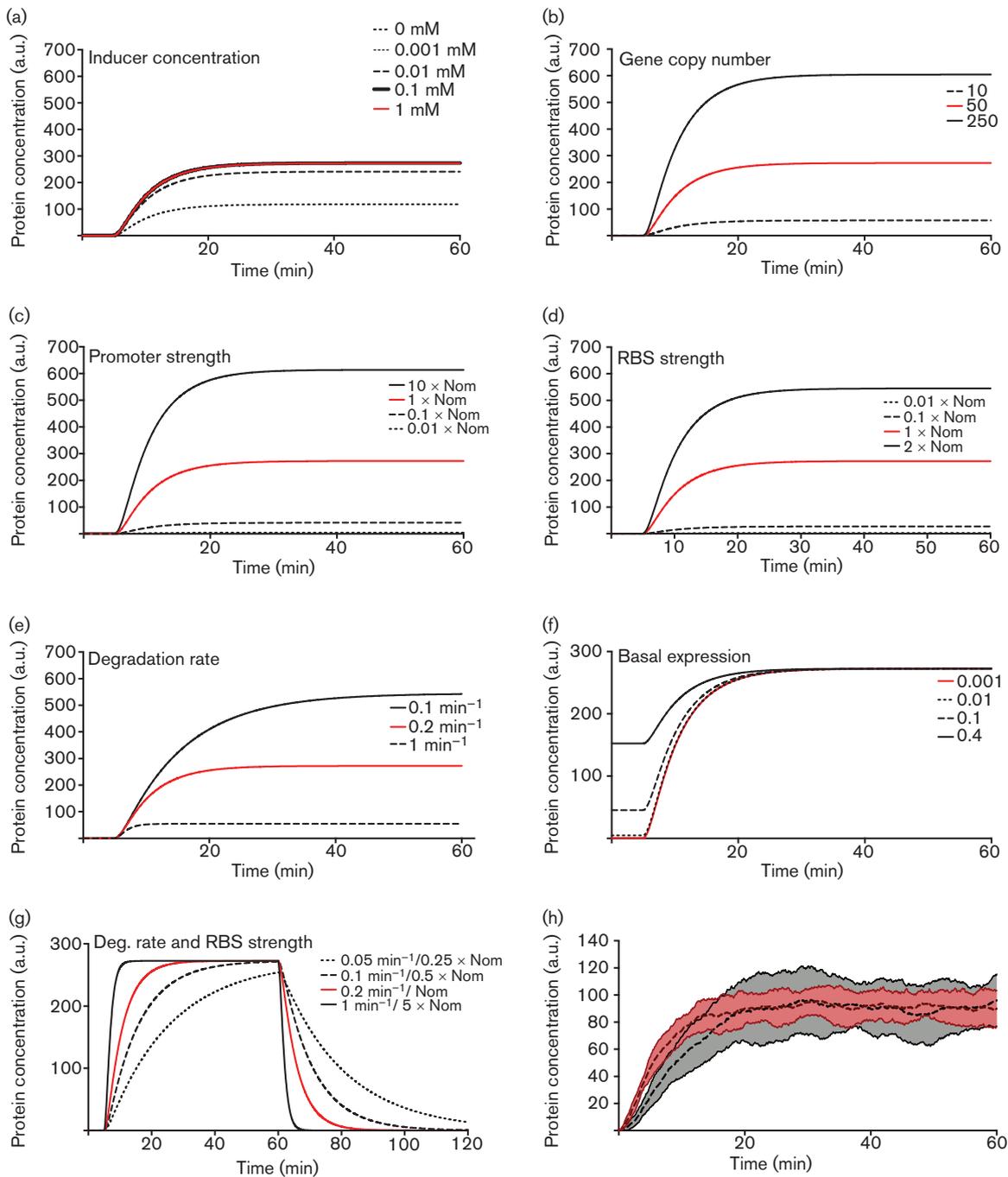
objectives (e.g. increasing the concentration of protein in the system) can often be designed rationally by an expert, other properties are more complex. For example, to change the system dynamics while keeping the expression level constant involves tuning multiple parameters, such as both the degradation rate and RBS strength (Fig. 6). In engineering design, this is generally achieved by setting such specifications as constraints and searching over all parameters that satisfy those constraints (Perry & Green, 1999). If more than one parameter choice meets all constraints and a benefit or cost is defined then the design can optimize the cost/benefit ratio. This is just one instance where the model can inform the choice between alternative redesigns.

The dials we have described span a range of scales and levels, and as such provide a series of options, which can be combined to achieve a design objective. Nonetheless, there are caveats: ultimately tuning any dial in a genetic network results in modulating the concentration of mRNAs and encoded proteins in the system. The dynamics of the system, however, depend on the type of control dials that are used. For example, if transcriptional level control dials (regulated promoters, transcriptional riboswitches) are used there will be a longer delay before a functional protein is produced since transcription, translation and protein folding must take place sequentially. On the other hand, if translational or post-translational control dials are used, the functional protein will be produced in a shorter period of time. Therefore if a genetic network with fast dynamics is required, it may be beneficial to factor post-translational control into the design process rather than transcriptional control. There will typically be a trade-off between the cell producing high levels of protein, poised to carry out its function, and gene expression inducing metabolic burden on the cell. The use of degradation tags on proteins also incurs a high metabolic burden since it decreases the steady state concentration by increasing the protein turnover rate. This results in resources being used to make a protein that is then targeted for fast degradation and is thus short-lived.

Linking multiple dials together can provide a genetic network with several avenues for tuning, providing a high level of control over network behaviour, e.g. coarse tuning through different origins of replication (modulating gene copy number), medium-level tuning through different promoters, and fine-tuning with different RBSs. However, linking different dials together often takes them out of the context under which they were initially characterized, thereby reducing the predictability of their individual behaviour and of their impact on the designed systems. For instance, it has been shown that increasing gene copy number can reduce the dynamic range and increase the leakiness of a promoter (Lutz & Bujard, 1997). Leakiness could potentially be reduced by adding a second layer of transcriptional or translational control by the addition of a third tuneable element, a riboregulator, thus combating the unwanted effects of copy number on promoter behaviour.

Addition and removal of control dials from a genetic network can be experimentally facilitated by the use of modular plasmid designs with large multiple cloning sites, allowing for the sequential addition of network components. Litcofsky *et al.* demonstrated this by constructing a simple toggle switch and a three-node or four-node feed-forward loop (Litcofsky *et al.*, 2012). Progress has also been made in the use of bio-parts in a plug-and-play methodology through the standardization of plasmid design (Silva-Rocha *et al.*, 2013).

Another factor to keep in mind is that, experimentally, some dials are easier to predictably tune than others. Altering gene copy number can be easy to achieve by replacing the origin of replication on plasmid-borne genetic networks or through single or multiple genomic integrations. Whilst the gene copy number can be controlled exactly through genomic integration, plasmid copy numbers can be harder to tune to exact levels given that many factors, described above, can affect plasmid copy numbers. Cell chassis tuning is less simple, potentially requiring genome engineering to achieve particular cell traits that impact on genetic network behaviour. As the effects of different cell chassis on network behaviour are currently not predictable, two approaches are available to aid in network redesign: (1) a genetic network can be characterized in several cell chassis to envisage the differential effects on the network with alternate chassis environments or (2) by using software such as Intermine (Smith *et al.*, 2012) or Ondex (Köhler *et al.*, 2006), developed for searching, data mining and integration of biological databases, which could help in identifying particular characteristics of different cell chassis to help direct and inform the design process. While the use of *in silico* approaches to design RBSs with predicted strengths can speed up the design and tuning process (Salis *et al.*, 2009), tuning most other dials can be time intensive due to the lack of software to help predict the effect changes on these dials may have. For example, whilst new promoters can be engineered, as described previously, there is often a trade-off between promoter strength, repressor strength, dynamic range and leakiness (Lanzer & Bujard, 1988). Trying to tune one of these parameters can often alter the others. Therefore, predictively designing a promoter with specific attributes is not straightforward. However, these trade-offs are common in engineering design for other fields, where they are typically handled using an optimization framework which considers various constraints and objective functions in the design (Boyd & Vandenberghe, 2004; Perry & Green, 1999; Dolan *et al.*, 2012). Directed evolution approaches (Lutz & Patrick, 2004; Neylon, 2004) are available to produce libraries of promoters but they often require extensive screening for desired characteristics and are thus often experimentally time consuming. Likewise, adding transcriptional level control with riboswitches can be relatively easy, whilst using a riboswitch for translational level control is more difficult as its function is often dependent on the RBS



sequence, which cannot be easily tuned without affecting the riboswitch integrity.

Two of the pioneering hallmarks for Synthetic Biology were the realization of simple designs inspired by existing electronic counterparts, i.e. a genetic toggle switch (Gardner *et al.*, 2000) and an oscillator (Stricker *et al.*, 2008). Their designs were inspired by a model-guided approach that provided an *in silico* assessment of the qualitative behaviour of these simple genetic networks. Further advancements in the field led to the use of a model-guided design (Ellis *et al.*, 2009), which allowed for the tuning of transcriptional layer

dials (promoter characteristics) in a reliable and relatively straightforward manner, to achieve a predictable genetic timer that controls yeast sedimentation (Ellis *et al.*, 2009). Within the scope of cell-based biosensing, model-guided design approaches have been used to inform the development of layered AND gates, housed in separate cell populations, which communicate through quorum sensing to detect specific combinations of metals (Beguerisse-Díaz *et al.*, 2012; Wang *et al.*, 2013). One of the most complex genetic designs achieved to date is exemplified by Moon *et al.* (2012), who used a combination of computational tools, model-guided design and directed evolution to construct a four-

Fig. 6. Deterministic simulations illustrating the effect of changing different dials on reporter protein concentration. The dot-dashed line in each graph gives a baseline case, which is identical across all simulations. In each panel except for (g) a single parameter has been changed whilst keeping the others constant. (a) Increasing the inducer concentration, added after 5 min, increases the steady state concentration until a saturation point is reached. (b) Increasing gene copy number increases the steady state concentration without changing the dynamics. The copy numbers used here are for real plasmid origins of replication and show that the range of values accessible using this method can be limited. Gene copy number is also a global change which affects all genes on the same plasmid equally. (c, d) Increasing or decreasing promoter or RBS strength respectively increases or decreases the steady state concentration, with respect to a nominal case (Nom), without changing dynamics. (e) Promoter leakiness increases the baseline level of protein and affects the ratio of steady state concentrations of protein in the presence or absence of inducer. Leaky promoters have a decreased dynamic range. (f) Decreasing the degradation rate increases both the steady state concentration and the time taken to reach steady state. (g) Simultaneously increasing RBS and protein degradation can change the dynamics while not modifying the steady state concentration. In this particular case the dynamics are changed, as the time taken to reach steady state is lower. Other combinations of parameters can lead to different behaviour. (h) Stochastic simulations highlighting the difference in noise between systems with high transcription and low translation (mean: dotted line, \pm SD shaded light red, $n=25$) or low transcription and high translation (mean: dashed line, \pm SD shaded dark grey, $n=25$).

input AND gate that consists of three circuits integrating four inducible systems within a single *E. coli* cell.

In order to realize the need for truly plug-and-play Synthetic Biology, the designer has to appreciate the types of dials they can use to achieve their design objectives: some are ‘difficult to tune’, some are ‘sensitive’ and some others are ‘uncertain’. In this review, we have described some of the possible dials that are available to the Synthetic Biologist at various organizational layers, thus opening the possibility for a design cycle that will involve mathematical modelling and optimization to produce systems with predictable, robust behaviour.

Acknowledgements

The authors were supported by EPSRC grants EP/I031944/1 and EP/I032223/1. K.P. is supported by an RCUK fellowship in Biopharmaceutical Processing. This review forms part of an overall project involving Dr Jorge Gonçalves, Dr Alex A. R. Webb and Dr Ye Yuan from the University of Cambridge and Dr George H. Wadhams from the University of Oxford.

References

- Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K. & Walter, P. (2002). *Molecular Biology of the Cell*, 4th edn. New York: Garland Science.
- Alon, U. (2007). *An Introduction to Systems Biology*. Boca Raton, FL: Chapman & Hall/CRC Press.
- Alper, H., Fischer, C., Nevoigt, E. & Stephanopoulos, G. (2005). Tuning genetic control through promoter engineering. *Proc Natl Acad Sci U S A* **102**, 12678–12683.
- Andersen, J. B., Sternberg, C., Poulsen, L. K., Bjorn, S. P., Givskov, M. & Molin, S. (1998). New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. *Appl Environ Microbiol* **64**, 2240–2246.
- Anderson, J., Strelkowa, N., Stan, G.-B., Douglas, T., Savulescu, J., Barahona, M. & Papachristodoulou, A. (2012). Engineering and ethical perspectives in synthetic biology. Rigorous, robust and predictable designs, public engagement and a modern ethical framework are vital to the continued success of synthetic biology. *EMBO Rep* **13**, 584–590.

Andrianantoandro, E., Basu, S., Karig, D. K. & Weiss, R. (2006). Synthetic biology: new engineering rules for an emerging discipline. *Mol Syst Biol* **2**, 0028.

Angov, E. (2011). Codon usage: nature’s roadmap to expression and folding of proteins. *Biotechnol J* **6**, 650–659.

Arnold, T. E. T., Yu, J. J. & Belasco, J. G. J. (1998). mRNA stabilization by the ompA 5’ untranslated region: two protective elements hinder distinct pathways for mRNA degradation. *RNA* **4**, 319–330.

Atkinson, M. R. M., Savageau, M. A. M., Myers, J. T. J. & Ninfa, A. J. A. (2003). Development of genetic circuitry exhibiting toggle switch or oscillatory behavior in *Escherichia coli*. *Cell* **113**, 597–607.

Balbás, M. D., Evans, M. J., Hosfield, D. J., Wongvipat, J., Arora, V. K., Watson, P. A., Chen, Y., Greene, G. L., Shen, Y. & Sawyers, C. L. (2013). Overcoming mutation-based resistance to antiandrogens with rational drug design. *Elife* **2**, e00499.

Becker, G. & Hengge-Aronis, R. (2001). What makes an *Escherichia coli* promoter σ^S dependent? Role of the -13/-14 nucleotide promoter positions and region 2.5 of σ^S . *Mol Microbiol* **39**, 1153–1165.

Beguerisse-Díaz, M., Wang, B., Desikan, R. & Barahona, M. (2012). Squeeze-and-breathe evolutionary Monte Carlo optimization with local search acceleration and its application to parameter fitting. *J R Soc Interface* **9**, 1925–1933.

Beisel, C. L. & Smolke, C. D. (2009). Design principles for riboswitch function. *PLoS Comput Biol* **5**, e1000363.

Bernardo, L. M., Johansson, L. U., Skärfstad, E. & Shingler, V. (2009). σ^{54} -promoter discrimination and regulation by ppGpp and DksA. *J Biol Chem* **284**, 828–838. (ASBMB).

Bouvet, P. P. & Belasco, J. G. J. (1992). Control of RNase E-mediated RNA degradation by 5’-terminal base pairing in *E. coli*. *Nature* **360**, 488–491.

Boyd, S. & Vandenberghe, L. (2004). *Convex Optimisation*. Cambridge, UK: Cambridge University Press.

Brannigan, J. A. & Wilkinson, A. J. (2002). Protein engineering 20 years on. *Nat Rev Mol Cell Biol* **3**, 964–970.

Brewster, R. C., Jones, D. L. & Phillips, R. (2012). Tuning promoter strength through RNA polymerase binding site design in *Escherichia coli*. *PLoS Comput Biol* **8**, e1002811.

Buck, M., Gallegos, M. T., Studholme, D. J., Guo, Y. & Gralla, J. D. (2000). The bacterial enhancer-dependent σ^{54} (σ^N) transcription factor. *J Bacteriol* **182**, 4129–4136.

- Callura, J. M., Cantor, C. R. & Collins, J. J. (2012).** Genetic switchboard for synthetic biology applications. *Proc Natl Acad Sci U S A* **109**, 5850–5855.
- Canton, B., Labno, A. & Endy, D. (2008).** Refinement and standardization of synthetic biological parts and devices. *Nat Biotechnol* **26**, 787–793.
- Carrier, T. A. & Keasling, J. D. (1997a).** Controlling messenger RNA stability in bacteria: strategies for engineering gene expression. *Biotechnol Prog* **13**, 699–708.
- Carrier, T. A. & Keasling, J. D. (1997b).** Controlling messenger RNA stability in bacteria: strategies for engineering gene expression. *Biotechnol Prog* **13**, 699–708.
- Chang, A. C. & Cohen, S. N. (1978).** Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J Bacteriol* **134**, 1141–1156.
- Chau, A. H., Walter, J. M., Gerardin, J., Tang, C. & Lim, W. A. (2012).** Designing synthetic regulatory networks capable of self-organizing cell polarization. *Cell* **151**, 320–332.
- Chen, H. H., Bjerknes, M. M., Kumar, R. R. & Jay, E. E. (1994).** Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. *Nucleic Acids Res* **22**, 4953–4957.
- Chen, S., Zhang, H., Shi, H., Ji, W., Feng, J., Gong, Y., Yang, Z. & Ouyang, Q. (2012).** Automated design of genetic toggle switches with predetermined bistability. *ACS Synth Biol* **1**, 284–290.
- Choi, Y. J., Bourque, D., Morel, L., Groleau, D. & Miguez, C. B. (2006).** Multicopy integration and expression of heterologous genes in *Methylobacterium extorquens* ATCC 55366. *Appl Environ Microbiol* **72**, 753–759.
- Cookson, N. A., Mather, W. H., Danino, T., Mondragón-Palomino, O., Williams, R. J., Tsimring, L. S. & Hasty, J. (2011).** Queueing up for enzymatic processing: correlated signaling through coupled degradation. *Mol Syst Biol* **7**, 561.
- Cornish-Bowden, A. (2004).** *Fundamentals of Enzyme Kinetics*, 3rd edn. London: Portland Press.
- Cosentino, C. & Bates, D. G. (2012).** *Feedback Control in Systems Biology*. Boca Raton, FL: Taylor & Francis/CRC Press.
- Cox, R. S., Surette, M. G. & Elowitz, M. B. (2007).** Programming gene expression with combinatorial promoters. *Mol Syst Biol* **3**, 145.
- Davis, J. H., Rubin, A. J. & Sauer, R. T. (2011).** Design, construction and characterization of a set of insulated bacterial promoters. *Nucleic Acids Res* **39**, 1131–1141.
- Del Vecchio, D., Ninfa, A. J. & Sontag, E. D. (2008).** Modular cell biology: retroactivity and insulation. *Mol Syst Biol* **4**, 169.
- Deuschle, U., Kammerer, W., Gentz, R. & Bujard, H. (1986).** Promoters of *Escherichia coli*: a hierarchy of in vivo strength indicates alternate structures. *EMBO J* **5**, 2987–2994.
- Dixon, N., Duncan, J. N., Geerlings, T., Dunstan, M. S., McCarthy, J. E. G., Leys, D. & Micklefield, J. (2010).** Reengineering orthogonally selective riboswitches. *Proc Natl Acad Sci U S A* **107**, 2830–2835.
- Dolan, J., Anderson, J. & Papachristodoulou, A. (2012).** A loop shaping approach for designing biological circuits. In *Proceedings of the IEEE Conference on Decision and Control*, 3614–3619.
- Dougan, D. A. D., Reid, B. G. B., Horwich, A. L. A. & Bukau, B. B. (2002).** ClpS, a substrate modulator of the ClpAP machine. *Mol Cell* **9**, 673–683.
- Driessen, A. J. M. & Nouwen, N. (2008).** Protein translocation across the bacterial cytoplasmic membrane. *Annu Rev Biochem* **77**, 643–667.
- Ebersbach, G. & Gerdes, K. (2005).** Plasmid segregation mechanisms. *Annu Rev Genet* **39**, 453–479.
- Egbert, R. G. R. & Klavins, E. E. (2012).** Fine-tuning gene networks using simple sequence repeats. *Proc Natl Acad Sci U S A* **109**, 16817–16822.
- Elleuche, S. & Pöggeler, S. (2010).** Inteins, valuable genetic elements in molecular biology and biotechnology. *Appl Microbiol Biotechnol* **87**, 479–489.
- Ellis, T., Wang, X. & Collins, J. J. (2009).** Diversity-based, model-guided construction of synthetic gene networks with predicted functions. *Nat Biotechnol* **27**, 465–471.
- Erbse, A., Schmidt, R., Bornemann, T., Schneider-Mergener, J., Mogk, A., Zahn, R., Dougan, D. A. & Bukau, B. (2006).** ClpS is an essential component of the N-end rule pathway in *Escherichia coli*. *Nature* **439**, 753–756.
- Figurski, D. H., Meyer, R. J. & Helinski, D. R. (1979).** Suppression of ColE1 replication properties by the Inc P-1 plasmid RK2 in hybrid plasmids constructed in vitro. *J Mol Biol* **133**, 295–318.
- Flynn, J. M. J., Levchenko, I. I., Seidel, M. M., Wickner, S. H. S., Sauer, R. T. R. & Baker, T. A. T. (2001).** Overlapping recognition determinants within the ssrA degradation tag allow modulation of proteolysis. *Proc Natl Acad Sci U S A* **98**, 10584–10589.
- Gardner, T. S., Cantor, C. R. & Collins, J. J. (2000).** Construction of a genetic toggle switch in *Escherichia coli*. *Nature* **403**, 339–342.
- Gillespie, D. T. (1992).** A rigorous derivation of the chemical master equation. *Physica A: Statist. Mechanics Applic* **188**, 404–425.
- Gogarten, J. P., Senejani, A. G., Zhaxybayeva, O., Olendzenski, L. & Hilario, E. (2002).** Inteins: structure, function, and evolution. *Annu Rev Microbiol* **56**, 263–287.
- Gotta, S. L., Miller, O. L., Jr & French, S. L. (1991).** rRNA transcription rate in *Escherichia coli*. *J Bacteriol* **173**, 6647–6649.
- Grossman, A. D., Straus, D. B., Walter, W. A. & Gross, C. A. (1987).** Sigma 32 synthesis can regulate the synthesis of heat shock proteins in *Escherichia coli*. *Genes Dev* **1**, 179–184.
- Gruber, T. M. & Gross, C. A. (2003).** Multiple sigma subunits and the partitioning of bacterial transcription space. *Annu Rev Microbiol* **57**, 441–466.
- Grünwald, D. & Singer, R. H. (2010).** In vivo imaging of labelled endogenous β -actin mRNA during nucleocytoplasmic transport. *Nature* **467**, 604–607.
- Gur, E. & Sauer, R. T. (2008).** Recognition of misfolded proteins by Lon, a AAA(+) protease. *Genes Dev* **22**, 2267–2277.
- Hakkaart, M. J. J., van Gemen, B., Veltkamp, E. & Nijkamp, H. J. J. (1985).** Maintenance of multicopy plasmid Clo DF13 III. Role of plasmid size and copy number in partitioning. *Mol Gen Genet* **198**, 364–366.
- Hansen, M. J., Chen, L. H., Fejzo, M. L. & Belasco, J. G. (1994).** The ompA 5' untranslated region impedes a major pathway for mRNA degradation in *Escherichia coli*. *Mol Microbiol* **12**, 707–716. (Wiley Online Library.)
- Hasunuma, K. K. & Sekiguchi, M. M. (1979).** Effect of dna mutations on the replication of plasmid pSC101 in *Escherichia coli* K-12. *J Bacteriol* **137**, 1095–1099.
- Herman, C. C., Thévenet, D. D., Boulloc, P. P., Walker, G. C. G. & D'Ari, R. R. (1998).** Degradation of carboxy-terminal-tagged cytoplasmic proteins by the *Escherichia coli* protease HflB (FtsH). *Genes Dev* **12**, 1348–1355.
- Hirschberg, K. & Lippincott-Schwartz, J. (1999).** Secretory pathway kinetics and in vivo analysis of protein traffic from the Golgi complex to the cell surface. *FASEB J* **13** (Suppl 2), S251–S256.
- Hoskins, J. R., Singh, S. K., Maurizi, M. R. & Wickner, S. (2000).** Protein binding and unfolding by the chaperone ClpA and degradation by the protease ClpAP. *Proc Natl Acad Sci U S A* **97**, 8892–8897.

- Isaacs, F. J., Dwyer, D. J., Ding, C., Pervouchine, D. D., Cantor, C. R. & Collins, J. J. (2004). Engineered riboregulators enable post-transcriptional control of gene expression. *Nat Biotechnol* **22**, 841–847.
- Jensen, P. R. & Hammer, K. (1998). The sequence of spacers between the consensus sequences modulates the strength of prokaryotic promoters. *Appl Environ Microbiol* **64**, 82–87.
- Kelly, J. R., Rubin, A. J., Davis, J. H., Ajo-Franklin, C. M., Cumbers, J., Czar, M. J., de Mora, K., Gliberman, A. L., Monie, D. D. & Endy, D. (2009). Measuring the activity of BioBrick promoters using an in vivo reference standard. *J Biol Eng* **3**, 4.
- Kirstein, J., Molière, N., Dougan, D. A. & Turgay, K. (2009). Adapting the machine: adaptor proteins for Hsp100/Clp and AAA + proteases 1–11. *Nat Rev Microbiol* **7**, 589–599.
- Kittleson, J. T., Cheung, S. & Anderson, J. C. (2011). Rapid optimization of gene dosage in *E. coli* using DIAL strains. *J Biol Eng* **5**, 10.
- Klumpp, S. (2011). Growth-rate dependence reveals design principles of plasmid copy number control. *PLoS ONE* **6**, e20403.
- Klumpp, S., Zhang, Z. & Hwa, T. (2009). Growth rate-dependent global effects on gene expression in bacteria. *Cell* **139**, 1366–1375.
- Koch, B., Liljefors, T., Persson, T., Nielsen, J., Kjelleberg, S. & Givskov, M. (2005). The LuxR receptor: the sites of interaction with quorum-sensing signals and inhibitors. *Microbiology* **151**, 3589–3602.
- Köhler, J., Baumbach, J., Taubert, J., Specht, M., Skusa, A., Rüegg, A., Rawlings, C., Verrier, P. & Philippi, S. (2006). Graph-based analysis and visualization of experimental results with ONDEX. *Bioinformatics* **22**, 1383–1390.
- Komarova, A. V., Tchufistova, L. S., Dreyfus, M. & Boni, I. V. (2005). AU-rich sequences within 5' untranslated leaders enhance translation and stabilize mRNA in *Escherichia coli*. *J Bacteriol* **187**, 1344–1349.
- Kool, A. J. A. & Nijkamp, H. J. H. (1974). Isolation and characterization of a copy mutant of the bacteriocinogenic plasmid Clo DF13. *J Bacteriol* **120**, 569–578.
- Kudla, G., Murray, A. W., Tollervey, D. & Plotkin, J. B. (2009). Coding-sequence determinants of gene expression in *Escherichia coli*. *Science* **324**, 255–258.
- Kües, U. & Stahl, U. (1989). Replication of plasmids in gram-negative bacteria. *Microbiol Rev* **53**, 491–516.
- Lanzer, M. & Bujard, H. (1988). Promoters largely determine the efficiency of repressor action. *Proc Natl Acad Sci U S A* **85**, 8973–8977.
- Lederer, T., Kintrup, M., Takahashi, M., Sum, P.-E., Ellestad, G. A. & Hillen, W. (1996). Tetracycline analogs affecting binding to Tn10-Encoded Tet repressor trigger the same mechanism of induction. *Biochemistry* **35**, 7439–7446.
- Lehmann, M. & Wyss, M. (2001). Engineering proteins for thermostability: the use of sequence alignments versus rational design and directed evolution. *Curr Opin Biotechnol* **12**, 371–375.
- Lestas, I., Vinnicombe, G. & Paulsson, J. (2010). Fundamental limits on the suppression of molecular fluctuations. *Nature* **467**, 174–178.
- Lewis, M., Chang, G., Horton, N. C., Kercher, M. A., Pace, H. C., Schumacher, M. A., Brennan, R. G. & Lu, P. (1996). Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* **271**, 1247–1254.
- Lim, H. N., Lee, Y. & Hussein, R. (2011). Fundamental relationship between operon organization and gene expression. *Proc Natl Acad Sci U S A* **108**, 10626–10631.
- Lim, W. A., Alvania, R. & Marshall, W. F. (2012). Cell biology 2.0. *Trends Cell Biol* **22**, 611–612.
- Lin-Chao, S., Chen, W. T. & Wong, T. T. (1992). High copy number of the pUC plasmid results from a Rom/Rop-suppressible point mutation in RNA II. *Mol Microbiol* **6**, 3385–3393.
- Litcofsky, K. D., Afeyan, R. B., Krom, R. J., Khalil, A. S. & Collins, J. J. (2012). Iterative plug-and-play methodology for constructing and modifying synthetic gene networks. *Nat Methods* **9**, 1077–1080.
- Lloyd, G., Landini, P. & Busby, S. (2001). Activation and repression of transcription initiation in bacteria. *Essays Biochem* **37**, 17–31.
- Lockless, S. W. & Muir, T. W. (2009). Traceless protein splicing utilizing evolved split inteins. *Proc Natl Acad Sci U S A* **106**, 10999–11004.
- Lutz, R. & Bujard, H. (1997). Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. *Nucleic Acids Res* **25**, 1203–1210.
- Lutz, S. & Patrick, W. M. (2004). Novel methods for directed evolution of enzymes: quality, not quantity. *Curr Opin Biotechnol* **15**, 291–297.
- Lynch, S. A., Desai, S. K., Sajja, H. K. & Gallivan, J. P. (2007). A high-throughput screen for synthetic riboswitches reveals mechanistic insights into their function. *Chem Biol* **14**, 173–184.
- Ma, W., Trusina, A., El-Samad, H., Lim, W. A. & Tang, C. (2009). Defining network topologies that can achieve biochemical adaptation. *Cell* **138**, 760–773.
- MacDonald, J. T., Barnes, C., Kitney, R. I., Freemont, P. S. & Stan, G.-B. V. (2011). Computational design approaches and tools for synthetic biology. *Integr Biol (Camb)* **3**, 97–108.
- Mackie, G. A. (2013). RNase E: at the interface of bacterial RNA processing and decay. *Nat Rev Microbiol* **11**, 45–57.
- Martoglio, B. & Dobberstein, B. (1998). Signal sequences: more than just greasy peptides. *Trends Cell Biol* **8**, 410–415.
- Mather, W., Bennett, M. R., Hasty, J. & Tsimring, L. S. (2009). Delay-induced degrade-and-fire oscillations in small genetic circuits. *Phys Rev Lett* **102**, 068105.
- McGinness, K. E., Baker, T. A. & Sauer, R. T. (2006). Engineering controllable protein degradation. *Mol Cell* **22**, 701–707.
- Menart, V., Jevševar, S., Vilar, M., Trobiš, A. & Pavko, A. (2003). Constitutive versus thermoinducible expression of heterologous proteins in *Escherichia coli* based on strong $P_{R_{PL}}$ promoters from phage lambda. *Biotechnol Bioeng* **83**, 181–190.
- Mileyko, Y. Y., Joh, R. I. R. & Weitz, J. S. J. (2008). Small-scale copy number variation and large-scale changes in gene expression. *Proc Natl Acad Sci U S A* **105**, 16659–16664.
- Mirasoli, M., Feliciano, J., Michelini, E., Daunert, S. & Roda, A. (2002). Internal response correction for fluorescent whole-cell biosensors. *Anal Chem* **74**, 5948–5953.
- Moon, T. S., Lou, C., Tamsir, A., Stanton, B. C. & Voigt, C. A. (2012). Genetic programs constructed from layered logic gates in single cells. *Nature* **491**, 249–253.
- Mootz, H. D. & Muir, T. W. (2002). Protein splicing triggered by a small molecule. *J Am Chem Soc* **124**, 9044–9045.
- Morcos, P. A. (2007). Achieving targeted and quantifiable alteration of mRNA splicing with Morpholino oligos. *Biochem Biophys Res Commun* **358**, 521–527.
- Murphy, K. F., Balázs, G. & Collins, J. J. (2007). Combinatorial promoter design for engineering noisy gene expression. *Proc Natl Acad Sci U S A* **104**, 12726–12731.
- Murray, J. D. (2002). *Mathematical Biology*. Berlin, Heidelberg: Springer Verlag.
- Na, D., Lee, S. & Lee, D. (2010). Mathematical modeling of translation initiation for the estimation of its efficiency to computationally design mRNA sequences with desired expression levels in prokaryotes. *BMC Syst Biol* **4**, 71.

- Newbury, S. F. S., Smith, N. H. N. & Higgins, C. F. C. (1987).** Differential mRNA stability controls relative gene expression within a polycistronic operon. *Cell* **51**, 1131–1143.
- Neylon, C. (2004).** Chemical and biochemical strategies for the randomization of protein encoding DNA sequences: library construction methods for directed evolution. *Nucleic Acids Res* **32**, 1448–1459.
- Nichols, J. C. & Matthews, K. S. (1997).** Combinatorial mutations of lac repressor. Stability of monomer-monomer interface is increased by apolar substitution at position 84. *J Biol Chem* **272**, 18550–18557.
- Nordström, K. K. & Uhlin, B. E. B. (1992).** Runaway-replication plasmids as tools to produce large quantities of proteins from cloned genes in bacteria. *Biotechnology (N Y)* **10**, 661–666.
- Oeffinger, M. & Zenklusen, D. (2012).** To the pore and through the pore: a story of mRNA export kinetics. *Biochim Biophys Acta* **1819**, 494–506.
- Orth, P., Schnappinger, D., Hillen, W., Saenger, W. & Hinrichs, W. (2000).** Structural basis of gene regulation by the tetracycline inducible Tet repressor-operator system. *Nat Struct Biol* **7**, 215–219.
- Osterman, I. A., Evfratov, S. A., Sergiev, P. V. & Dontsova, O. A. (2013).** Comparison of mRNA features affecting translation initiation and reinitiation. *Nucleic Acids Res* **41**, 474–486.
- Panayotatos, N. (1984).** DNA replication regulated by the priming promoter. *Nucleic Acids Res* **12**, 2641–2648.
- Papanikou, E., Karamanou, S. & Economou, A. (2007).** Bacterial protein secretion through the translocase nanomachine. *Nat Rev Microbiol* **5**, 839–851.
- Parsell, D. A., Silber, K. R. & Sauer, R. T. (1990).** Carboxy-terminal determinants of intracellular protein degradation. *Genes Dev* **4**, 277–286.
- Penumetcha, P., Lau, K., Zhu, X., Davis, K., Eckdahl, T. T. & Campbell, A. M. (2010).** Improving the Lac system for synthetic biology. *BIOS* **81**, 7–15.
- Perler, F. B. (2002).** InBase: the intein database. *Nucleic Acids Res* **30**, 383–384.
- Perry, R. H. & Green, D. W. (1999).** *Perry's Chemical Engineers' Handbook*, 7th edn. New York: McGraw-Hill.
- Peterson, J. & Phillips, G. J. (2008).** New pSC101-derivative cloning vectors with elevated copy numbers. *Plasmid* **59**, 193–201.
- Pfeuty, B. & Kaneko, K. (2009).** The combination of positive and negative feedback loops confers exquisite flexibility to biochemical switches. *Phys Biol* **6**, 046013.
- Prindle, A., Selimkhanov, J., Danino, T., Samayoa, P., Goldberg, A., Bhatia, S. N. & Hasty, J. (2012).** Genetic circuits in *Salmonella typhimurium*. *ACS Synth Biol* **1**, 458–464.
- Purcell, O., Savery, N. J., Grierson, C. S. & di Bernardo, M. (2010).** A comparative analysis of synthetic genetic oscillators. *J R Soc Interface* **7**, 1503–1524.
- Purcell, O., Grierson, C. S., Bernardo, M. & Savery, N. J. (2012).** Temperature dependence of ssrA-tag mediated protein degradation. *J Biol Eng* **6**, 10.
- Purnick, P. E. M. & Weiss, R. (2009).** The second wave of synthetic biology: from modules to systems. *Nat Rev Mol Cell Biol* **10**, 410–422.
- Qi, L. S., Larson, M. H., Gilbert, L. A., Doudna, J. A., Weissman, J. S., Arkin, A. P. & Lim, W. A. (2013).** Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell* **152**, 1173–1183.
- RAEng (2009).** *Synthetic Biology: Scope, Applications and Implications*. London: Royal Academy of Engineering
- Raghavan, R. & Minnick, M. F. (2009).** Group I introns and inteins: disparate origins but convergent parasitic strategies. *J Bacteriol* **191**, 6193–6202.
- Raj, A. & van Oudenaarden, A. (2008).** Nature, nurture, or chance: stochastic gene expression and its consequences. *Cell* **135**, 216–226.
- Salis, H. M., Mirsky, E. A. & Voigt, C. A. (2009).** Automated design of synthetic ribosome binding sites to control protein expression. *Nat Biotechnol* **27**, 946–950.
- Satya Lakshmi, O. & Rao, N. M. (2009).** Evolving Lac repressor for enhanced inducibility. *Protein Eng Des Sel* **22**, 53–58.
- Schleif, R. (2000).** Regulation of the L-arabinose operon of *Escherichia coli*. *Trends Genet* **16**, 559–565.
- Schmidt, L. & Inselburg, J. (1982).** ColE1 copy number mutants. *J Bacteriol* **151**, 845–854.
- Scott, M., Gunderson, C. W., Mateescu, E. M., Zhang, Z. & Hwa, T. (2010).** Interdependence of cell growth and gene expression: origins and consequences. *Science* **330**, 1099–1102.
- Sen, S. & Murray, R. M. (2012).** Performance metrics for a biomolecular step response. In *Decision and Control (CDC), 2012 IEEE 51st Annual Conference*. <http://www.cds.caltech.edu/~murray/papers/sm12-cdc.html>
- Seo, S. W., Yang, J.-S., Kim, I., Yang, J., Min, B. E., Kim, S. & Jung, G. Y. (2013).** Predictive design of mRNA translation initiation region to control prokaryotic translation efficiency. *Metab Eng* **15**, 67–74.
- Shi, J. & Muir, T. W. (2005).** Development of a tandem protein trans-splicing system based on native and engineered split inteins. *J Am Chem Soc* **127**, 6198–6206.
- Shizuya, H., Birren, B., Kim, U. J., Mancino, V., Slepak, T., Tachiiri, Y. & Simon, M. (1992).** Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in *Escherichia coli* using an F-factor-based vector. *Proc Natl Acad Sci U S A* **89**, 8794–8797.
- Shong, J., Jimenez Diaz, M. R. & Collins, C. H. (2012).** Towards synthetic microbial consortia for bioprocessing. *Curr Opin Biotechnol* **23**, 798–802.
- Silber, K. R. K., Keiler, K. C. K. & Sauer, R. T. R. (1992).** Tsp: a tail-specific protease that selectively degrades proteins with nonpolar C termini. *Proc Natl Acad Sci U S A* **89**, 295–299.
- Silva-Rocha, R., Martínez-García, E., Calles, B., Chavarría, M., Arce-Rodríguez, A., de Las Heras, A., Páez-Espino, A. D., Durante-Rodríguez, G., Kim, J. & other authors (2013).** The Standard European Vector Architecture (SEVA): a coherent platform for the analysis and deployment of complex prokaryotic phenotypes. *Nucleic Acids Res* **41** (Database issue), D666–D675.
- Skretas, G. & Wood, D. W. (2005).** Regulation of protein activity with small-molecule-controlled inteins. *Protein Sci* **14**, 523–532.
- Slusarczyk, A. L., Lin, A. & Weiss, R. (2012).** Foundations for the design and implementation of synthetic genetic circuits. *Nat Rev Genet* **13**, 406–420.
- Smith, R. N., Aleksic, J., Butano, D., Carr, A., Contrino, S., Hu, F., Lyne, M., Lyne, R., Kalderimis, A. & other authors (2012).** InterMine: a flexible data warehouse system for the integration and analysis of heterogeneous biological data. *Bioinformatics* **28**, 3163–3165.
- Sternberg, N. (1990).** Bacteriophage P1 cloning system for the isolation, amplification, and recovery of DNA fragments as large as 100 kilobase pairs. *Proc Natl Acad Sci U S A* **87**, 103–107.
- Strelkova, N. & Barahona, M. (2010).** Switchable genetic oscillator operating in quasi-stable mode. *J R Soc Interface* **7**, 1071–1082.
- Stricker, J., Cookson, S., Bennett, M. R., Mather, W. H., Tsimring, L. S. & Hasty, J. (2008).** A fast, robust and tunable synthetic gene oscillator. *Nature* **456**, 516–519.

- Swinburne, I. A., Miguez, D. G., Landgraf, D. & Silver, P. A. (2008).** Intron length increases oscillatory periods of gene expression in animal cells. *Genes Dev* **22**, 2342–2346.
- Tan, C., Marguet, P. & You, L. (2009).** Emergent bistability by a growth-modulating positive feedback circuit. *Nat Chem Biol* **5**, 842–848.
- Tian, T. & Burrage, K. (2006).** Stochastic models for regulatory networks of the genetic toggle switch. *Proc Natl Acad Sci U S A* **103**, 8372–8377.
- Tolia, N. H. & Joshua-Tor, L. (2006).** Strategies for protein coexpression in *Escherichia coli*. *Nat Methods* **3**, 55–64.
- Topp, S., Reynoso, C. M. K., Seeliger, J. C., Goldlust, I. S., Desai, S. K., Murat, D., Shen, A., Puri, A. W., Komeili, A. & other authors (2010).** Synthetic riboswitches that induce gene expression in diverse bacterial species. *Appl Environ Microbiol* **76**, 7881–7884.
- Tyson, J. J., Chen, K. C. & Novak, B. (2003).** Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. *Curr Opin Cell Biol* **15**, 221–231.
- Urbanowski, M. L., Lostroh, C. P. & Greenberg, E. P. (2004).** Reversible acyl-homoserine lactone binding to purified *Vibrio fischeri* LuxR protein. *J Bacteriol* **186**, 631–637.
- Villaverde, A. A., Benito, A. A., Viaplana, E. E. & Cubarsi, R. R. (1993).** Fine regulation of *cI857*-controlled gene expression in continuous culture of recombinant *Escherichia coli* by temperature. *Appl Environ Microbiol* **59**, 3485–3487.
- Vitreschak, A. G., Rodionov, D. A., Mironov, A. A. & Gelfand, M. S. (2004).** Riboswitches: the oldest mechanism for the regulation of gene expression? *Trends Genet* **20**, 44–50.
- Vogel, U. & Jensen, K. F. (1994).** The RNA chain elongation rate in *Escherichia coli* depends on the growth rate. *J Bacteriol* **176**, 2807–2813.
- Wachsmuth, M., Findeiß, S., Weissheimer, N., Stadler, P. F. & Mörl, M. (2013).** De novo design of a synthetic riboswitch that regulates transcription termination. *Nucleic Acids Res* **41**, 2541–2551.
- Wang, Y. & deHaseth, P. L. (2003).** Sigma 32-dependent promoter activity in vivo: sequence determinants of the *groE* promoter. *J Bacteriol* **185**, 5800–5806.
- Wang, K. H., Sauer, R. T. & Baker, T. A. (2007).** ClpS modulates but is not essential for bacterial N-end rule degradation. *Genes Dev* **21**, 403–408.
- Wang, B., Barahona, M. & Buck, M. (2013).** A modular cell-based biosensor using engineered genetic logic circuits to detect and integrate multiple environmental signals. *Biosens Bioelectron* **40**, 368–376.
- Welch, M., Govindarajan, S., Ness, J. E., Villalobos, A., Gurney, A., Minshull, J. & Gustafsson, C. (2009).** Design parameters to control synthetic gene expression in *Escherichia coli*. *PLoS ONE* **4**, e7002.
- Wilkinson, D. J. (2011).** *Stochastic Modelling for Systems Biology*, 2nd edn. Boca Raton, FL: Chapman & Hall/CRC Press.
- Wilson, C. & Agard, D. A. (1991).** Engineering substrate specificity. *Curr Opin Struct Biol* **1**, 617–623.
- Yabuta, M., Onai-Miura, S. & Ohsuye, K. (1995).** Thermo-inducible expression of a recombinant fusion protein by *Escherichia coli* lac repressor mutants. *J Biotechnol* **39**, 67–73.
- Zhong, C., Peng, D., Ye, W., Chai, L., Qi, J., Yu, Z., Ruan, L. & Sun, M. (2011).** Determination of plasmid copy number reveals the total plasmid DNA amount is greater than the chromosomal DNA amount in *Bacillus thuringiensis* YBT-1520. *PLoS ONE* **6**, e16052.