Recognizing and engineering digital-like logic gates and switches in gene regulatory networks

Robert W Bradley¹,², Martin Buck¹ and Baojun Wang²,³

A central aim of synthetic biology is to build organisms that can perform useful activities in response to specified conditions. The digital computing paradigm which has proved so successful in electrical engineering is being mapped to synthetic biological systems to allow them to make such decisions. However, stochastic molecular processes have graded input-output functions, thus, bioengineers must select those with desirable characteristics and refine their transfer functions to build logic gates with digital-like switching behaviour. Recent efforts in genome mining and the development of programmable RNA-based switches, especially CRISPR, have greatly increased the number of parts available to synthetic biologists. Improvements to the digital characteristics of these parts are required to enable robust predictable design of deeply layered logic circuits.

Addresses
¹ Department of Life Sciences, Faculty of Natural Sciences, Imperial College London, London SW7 2AZ, UK
² School of Biological Sciences, University of Edinburgh, Edinburgh EH9 3FF, UK
³ Centre for Synthetic and Systems Biology, University of Edinburgh, Edinburgh EH9 3FF, UK

Corresponding authors: Wang, Baojun (baojun.wang@ed.ac.uk) and Buck, Martin (m.buck@imperial.ac.uk)

Introduction

Electronic computers contain powerful decision-making circuits, built using switches with well-defined digital characteristics that are connected to produce Boolean logic operators. Synthetic biologists are making progress at replicating digital decision making in living organisms, aiming to program cells for applications in areas such as environmental sensing and medicine [1–3].

Digital-like behaviour in natural and synthetic biological systems is used to produce in effect all-or-nothing responses: the output signal from digital-like modules switches between low and high output levels (OFF/ON; binary 0/1) over a short range of input signal. Biology is inherently analogue due to the stochastic nature of the molecular interactions that propagate information flow, and so biological switches possess digital characteristics to greater or lesser degrees. Strongly digital-like characteristics are desirable when implementing biological switches in bio-computing circuits as Boolean logic gates (Figure 1a). A steep, ultrasensitive transition between OFF and ON states is key, minimising signal degradation when logic gates are layered (a condition where the output of one logic gates acts as the input for another) [4–6]. A large difference between output levels in the OFF and ON states also reduces noise propagation through the circuit, maintaining signal fidelity.

The inputs and outputs from connected gates in a circuit must be composable both in terms of signal type — so information can be transferred — and amplitude — so that the OFF and ON output levels of an upstream gate are below and above the switching threshold for the downstream gate (Figure 1b). Ideally the switching threshold and output level of a gate should be tunable. Decision-making also requires that logic gates receive inputs from multiple upstream gates, whilst remaining orthogonal to signals from all other host and synthetic components in the system [5,6].

Here we review efforts that have been made to identify parts for digital bio-computation, with an emphasis on large part families and those that are amenable to rational redesign, as these will form the basis of future large-scale genetic logic circuits. Improvements to the digital characteristics of existing biological logic gates are necessary to maintain signal fidelity in deeply layered circuits, and we discuss engineering strategies for making these enhancements.

Identifying modules with digital characteristics

Characterisation of a component’s switching properties allows key properties such as dynamic range, activation threshold, and transfer function steepness to be determined [7,8]. The nonlinear, ultrasensitive response to an input signal that characterises digital-like biological parts is usually quantified by fitting the Hill function to the curve, with ultrasensitive mechanisms having an apparent Hill coefficient greater than one [9]. Fundamental knowledge of a biological part’s mechanisms of action allows
probable candidates for logic gates to be selected: Components with known cooperative mechanisms, such as the TetR repressor’s ligand-induced weakening of DNA binding affinity [10], can be chosen to provide sensitive switching; High ON:OFF ratios can be found in part types with low intrinsic leakiness, for example when a part is absolutely required for output such as a phage RNA polymerase [11]; The requirement for integration of multiple signals can be fulfilled by choosing components with activating or repressing partners, for example transcription factors which need activating chaperones [12]. Our lab has investigated the Pseudomonas syringae hypersensitive response pathway regulatory components as a model for engineering orthogonal digital-like control of transcription in Escherichia coli [2,5,13,14**] (Figure 2). A great number of similar regulatory modules exist in many different bacterial species, offering a largely untapped resource to construct versatile orthogonal genetic logic devices.

Sophisticated digital genetic circuits require a large number of composable parts that act with minimal crosstalk and cause low toxicity to the host. Genomic mining strategies can be employed to screen for orthogonal homologs of useful parts. Stanton et al. produced a set of 16 orthogonal TetR repressor homologs and cognate operators which was used to build NOT and NOR gates [15] (NOR gates are desirable because they are functionally complete). Whilst the design of a single repressor binding site within a strong constitutive promoter was
appropriate for library construction and screening, the authors note that this configuration produces gates with a high OFF state and low cooperativity. Future versions could use multiple operators to improve digital characteristics, as will be discussed in the section “Motifs for ultrasensitivity”.

Modifying logic gate characteristics

Biological components require some modification from their native configuration to allow them to connect properly and retain signal fidelity in the context of a large synthetic gene circuit. Whilst largely irrational modification of individual components has been shown to be an effective strategy for isolating variants with enhanced ON:OFF ratios and altered thresholds [16], and improved orthogonality [12,17], more rational approaches, often using in silico models, enable efficient and systematic optimisation.

Tuning for composition

The output and activation threshold of a switch may be tuned to facilitate composition with neighbouring gates,
sensors, or analogue synthetic circuitry. This is usually performed by altering the concentration of a gate’s constituent components, for example, higher concentrations of an activating transcription factor will decrease the activation threshold of a switch [7]. Tuning can be achieved using a number of mechanisms [18,19], though usually via changes to the transcription and translation initiation sequences. For bacteria especially, part libraries [20,21] and computational tools for ribosome binding site design [22,23] enable efficient screening of sequences to achieve desired component levels. Transcription and translation initiation sequences suffer from context dependencies which must be minimised to enable predictive design of synthetic gene circuits [21,24**]. Repressive antisense transcription is another technique that could be widely applied to fine-tune transcriptional logic gate activation thresholds [25].

**Motifs for ultrasensitivity**

Many part types do not display ultrasensitive responses, so this property must be engineered. Steep switching transitions can occur due to various molecular mechanisms [26,27], but some of these are more amenable to intervention by design: Whilst introducing cooperative binding of ligand molecules to a receptor would be a difficult (and probably unique) protein engineering problem, building gene circuits with motifs that create an ultrasensitive response — such as sequestration, multi-step mechanisms, and positive feedback — is a widely applicable strategy, and one that allows for tuning of the transfer function.

The threshold and profile of a transfer function can be modified to have more digital-like characteristics using a ‘sequestration’ or ‘titration’ strategy, where high affinity sequestration of a signal-carrying factor by a buffer of decoy binding sites must be overcome before its effect on the output is observed (Figure 3a). This strategy also has the effect of lowering the OFF state, and shifting the activation threshold to a higher level [28]. The degree of sensitivity and shape of the response may be modified by using decoys with different binding affinities, or different concentrations of decoy [29,30*]. Sequestration can be performed by a constitutively expressed binding partner: Rhodius et al. identified twenty highly orthogonal extracytoplasmic function (ECF) and their corresponding promoters, plus cognate anti-σ factors, using genomic part mining [6]. The simple buffer gate that results from inducible ECF expression does not exhibit good digital characteristics, but using low-level expression of the anti-σ to sequester its partner improves the sigmoidicity of the response. Using RNA-RNA interactions for sequestration is an appealing strategy as binding partners can be easily designed [31]. Similarly, it is simple to add decoy binding sites for transcription factors into synthetic DNA [32].

Ultrasensitivity can also arise from multi-step mechanisms, which use an input to regulate multiple levels of a signal cascade, resulting in a steeper multiplicative output response (Figure 3b). Implementation of a cascade also allows for signal amplification, increasing the ON:OFF ratio. Xie et al. made use of the programmability of nucleic acid components when applying this motif in their HeLa cell classifier, adding miRNA target sites to the mRNAs of cascading transcription factors [1].

Positive feedback loops have also been successfully employed to increase the steepness of transfer functions, and amplify the output signal (Figure 3c) [33,34]. Palani and Sarkar made use of a dual-feedback motif which amplified both receptor and transcription factor components of a cascade (i.e. also a multi-step mechanism) to improve and tune the threshold, sensitivity, and output of their transfer function [34]. Unwanted bistability is a potential downside of using positive feedback motifs: because the ‘reset’ transfer function is offset in a bistable system, the range of input concentration over which effective bi-directional switching occurs increases, possibly obscuring the improvements made by steepening the transfer function (Figure 3d). The bistable region can however be tuned (minimised) through sequestration [27,30*].

**Integrating signals**

Logic gates need to assimilate multiple input signals. For transcriptional logic gates, it is often possible to simply combine promoter or operator sequences to control transcription of the output [4**,*15]. Similarly at the RNA level, some cis-acting sequences can be concatenated to allow multiple trans-acting elements to control translation, for example small transcription activating RNA cis-elements [35] or micro RNA target sequences [1].

Another general strategy for creating AND or NAND logic gates is to split the carrier of an input signal into parts that are individually inactive. Split parts might recombine to form the active component spontaneously [17], or can be fused to domains that (inducibly [36]) promote association. Addition of split-intein domains to divided protein components allows the native polypeptide to be reformed, which is useful if there is weak spontaneous association or the activity is sensitive to fusions [11,37].

**Scaling-up logic circuits**

Large-scale circuits require large orthogonal sets of switches that are composable, retain signal fidelity, and are functionally complete. Part mining is a promising approach for discovering such sets, but using parts that have programmable specificities enables their creation in a rational manner. Protein tools with customisable DNA-binding specificity, such as transcription activator-like repressors, have been used successfully to build logic gates [38], but their repeated structure makes them difficult to synthesise. Nucleic acids are facile to produce with current cloning and synthesis techniques, but many
RNA-based part families suffer from weaker binding interactions compared to proteins. In recent years a number of new RNA-based tools have been developed which have overcome previous limitations in dynamic range [39,40], but as yet none have all the qualities required for large-scale circuits. A promising compromise is a transcriptional switch based on the *Streptococcus pyogenes* clustered regularly interspaced short palindromic repeat (CRISPR) Cas9 protein, which combines RNA-based programmability with strong binding.

**CRISPR-dCas9 logic gates**

Nuclease-inactive Cas9 (dCas9) retains the ability to tightly bind a target DNA sequence complementary to the spacer of a guide RNA (gRNA). Transcriptional repression by dCas9-mediated CRISPR-interference (CRISPRi) reduces expression by up to 1000-fold [41,42]. The large ON:OFF ratio means CRISPRi can be used for digital-like gene circuits, where layered logic is produced by controlling the expression of downstream gRNAs [43,44]. The large ON:OFF ratio is not sufficient for deeply layered circuits and so CRISPRi must be engineered into an ultrasensitive switch. Gander *et al.* recently fused the Mxi1 chromatin remodeler to dCas9 for improved repression in their yeast gene circuits [4**]. The increased cooperativity due to Mxi1 activity enabled the construction of three-layer logic circuits, plus an impressive seven-layer inverting cascade. dCas9 can also act as a scaffold for transcription activation proteins to switch target promoters ON (CRISPRa) [42,45*,46,47,48*], although generally lower reported ON:OFF ratios combined with a lack of ultrasensitivity.
has so far limited the use of CRISPRa in digital logic circuits.

CRISPRi naturally lends itself to NAND logic through differential expression of the gRNA and protein components as the inputs [43], and split versions of (d)Cas9 have been developed which will enable greater regulatory control and versatility [36,49,50]. Ultimately, dCas9 can effect the decision made by the synthetic computation circuit on the host transcriptome [44,45,47]. Improvements in Cas9 specificity will obviously also be beneficial to CRISPRi circuits [51,52], but orthogonality and modularity in synthetic systems can also be improved by optimisation of the target sequences [53].

### Conclusions

With the resources now available through part mining efforts and use of programmable components (Table 1), combined with rational approaches to refining logic gate characteristics, we anticipate significant increases in the scale and complexity of synthetic biological digital-like logic circuits in the near future. Whilst it is prudent to remember that digital logic is not the best choice for all biological computations (compared to analogue computing strategies it is often more energy-expensive and resource-expensive [54,55]), the ability to program organisms to make robust binary decisions will be fundamental to many applications. The immediate challenges for the field are to improve the digital characteristics of parts to enable deeper layering of circuits, and continue to develop effective computational tools for circuit design.

Because of their programmability it is likely that dCas9 homologs [56] and catalytic mutants of other RNA-guided nucleases [57] will play a central role in the next generation of digital gene circuits, especially where interaction with the host genome is required. Ultrasensitivity in dCas9 or other transcription factors might be improved by the incorporation of sequestration [58], multi-step, or feedback strategies, though the addition of dimerization domains to enable cooperative binding to DNA is an intriguing untested possibility [44].

As both part libraries and the scale of desired synthetic circuits grow, the act of choosing the most appropriate parts will become increasingly automated [24,59]. Nielsen et al. provide an exciting insight into the future of digital bio-programming with their recent work, using computer aided design to produce functional logic circuits with a 75% success rate in the first design-build-test cycle [24]. The study highlights how part insulation is essential for predictive design, and underlines again that improvements to the digital character of gates are required to combat the trend of increasing failure rate with growing circuit depth. As circuit sizes grow, designers will have to consider the functional modularity of parts in order to mitigate the effects of retroactivity [60] and host burden [61]. Host chassis genome minimisation will

### Table 1

<table>
<thead>
<tr>
<th>Method</th>
<th>Repurposing natural gene regulatory modules</th>
<th>Genomic part mining</th>
<th>Protein splitting</th>
<th>CRISPRi</th>
</tr>
</thead>
<tbody>
<tr>
<td>Requirements</td>
<td>Identify parts that have good digital characteristics, and are likely to be orthogonal to the desired host. Characterise function in the synthetic context; test for orthogonality and toxicity.</td>
<td>Choose large part family with a diverse range of specific interactions. Construct a library of parts and reporters; screen for performance, orthogonality, and toxicity.</td>
<td>Divide protein into inactive subunits; testing a library of split versions may be necessary [58]. Fuse to split-tein domains to enable reconstitution of full-length polypeptide.</td>
<td>Design gRNA sequences and cognate operators that are orthogonal to the host and other circuit components. Construct NOT or NOR gates by combining constitutive promoters with operators. Easy to program for large-scale orthogonal circuits.</td>
</tr>
<tr>
<td>Advantages</td>
<td>Can choose parts for specific applications, for example sensing metabolites, or integration of particular signal types. Can generate large sets of orthogonal, composable parts for large-scale circuits. Subsequent tuning can improve function of library variants when applied in a circuit context.</td>
<td>Requires part types that are amenable to high-throughput screening. Construction of a synthetic library can be expensive.</td>
<td>Protein might not split into stable or soluble subunits; subunits may not localise correctly. Crosstalk may occur within parts families that are split into structurally similar subunits.</td>
<td>Not ultrasensitive in native form.</td>
</tr>
<tr>
<td>Weaknesses</td>
<td>May not have broad applicability or be amenable to scale-up for larger circuits.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Examples</td>
<td>[5,12,14*]</td>
<td>[6,15]</td>
<td>[11,17,49,50]</td>
<td>[4,41–44]</td>
</tr>
</tbody>
</table>
hopefully make undesirable interactions between synthetic and host components easier to predict and avoid [62]. Temporal dynamics will also gain importance as circuits become more deeply layered, lengthening the time taken to elicit the ultimate output, and also increasing the likelihood of faults occurring due to signals propagating at different speeds. This will drive the development of new logic gate types with faster switching, perhaps using reversible covalent modification rather than transcription and translation. Digital memory elements, such as bistable switches or gates based on recombinase-mediated DNA flipping [63], can also be employed to improve signal stability and fidelity.

The degree of characterisation that parts are subjected to has so far been fairly ad hoc, on the basis of pragmatic project-specific constraints, but the fabrication of large sets of quality components that can be applied in diverse situations will change this mind-set. Some standardisation is now required for the transition from building at the scale of individual logic gates and simple functions, to the construction of effective, robust systems [64]. Standardisation allows for much of the complexity of biological systems to be ignored at the systems level, abstracting a functional unit to a small set of IN/OUT properties. Canton et al. envisioned datasheets to accompany biological parts [8], which could include switching thresholds, LOW/HIGH output levels, and signal rise time. This abstraction is best suited to highly insulated components, which those involved in decision-making signal processing ideally are. The question of how best to define a standard for biological logic gates is beyond the scope of this review, but will be influenced by the designs of those who develop component libraries, by the properties of the particular part family, and by the requirements and established practices of the community of end-users. Standardisation of notation in the form of the Synthetic Biology Open Language (SBOL) already facilitates the transferability and uptake of new designs [65]; similar case of use and reuse of components will enable synthetic biology to start achieving its potential in real-world applications.

Acknowledgements

This work was supported by UK BBSRC project grants [BB/K016288/1] and [BB/N007212/1]. B.W. acknowledges support from the Scottish Universities Life Sciences Alliance (SULSA).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest


** Cooperative recruitment of chromatin remodelling factors by dCas9-MxiI increased the Hill coefficient of this CRISPRi system to 1.7, enabling construction of the most deeply layered repressive cascade yet. The authors also model the effect of increasing cooperativity on signal degradation, theorising that near-zero degradation can be achieved using switches with Hill coefficients above 2.5.


The authors build a program for computer aided design of biological logic circuits that uses extensive characterisation data of biological parts — in this case of TetR-family-based NOT gates that were modified for improved insulation — to inform the construction of 60 logic circuits. The initial designs produce the correct output for 92% of all input combination states; debugging plasmodium could then be employed to identify problem components.


In this study the Pseudomonas aeruginosa ExsACD partner-swapping regulatory cascade components are repurposed into a tunable bistable system, with a positive-feedback loop generating ultrasensitive switching. Tuning of the hysteretic region of the switch was modelled and verified, showcasing a useful regulatory motif that could be employed to create either digital switches or robust memory elements.


Using modified versions of the CRISPR-Cas9 guide RNA as scaffolds for effector protein recruitment enables simultaneous positive and negative regulation of genes in a heterologous metabolic pathway.


The authors fuse a tripartite VP64-p65-Rta activator to dCas9, enabling fold activation of four orders of magnitude in human cells, and expediting the future use of CRISPRs in digital gene circuits.


Future large-scale CRISPRi gene circuits will need to be orthogonal to the host genome and other circuit components. This work uses computational design of guide RNAs and cognate target promoters to minimise off-target dCas9 binding; synthetic gRNA target sequences are also introduced into native Escherichia coli promoters to create repressible hybrids.
The authors demonstrate a load driver module on the basis of a fast phosphotransfer relay that acts as a functional insulator against retro-active effects of increased load in the form of free transcription factor binding sites. Practical considerations for including load drivers in complex future circuits are discussed.