Review

Engineering and standardization of posttranscriptional biocircuitry in *Saccharomyces cerevisiae*

John McCarthy

Warwick Integrative Synthetic Biology Centre (WISB) and School of Life Sciences, University of Warwick, Coventry CV4 7AL, UK

*Corresponding author. E-mail: john.mccarthy@warwick.ac.uk

Abstract

This short review considers to what extent posttranscriptional steps of gene expression can provide the basis for novel control mechanisms and procedures in synthetic biology and biotechnology. The term biocircuitry is used here to refer to functionally connected components comprising DNA, RNA or proteins. The review begins with an overview of the diversity of devices being developed and then considers the challenges presented by trying to engineer more scaled-up systems. While the engineering of RNA-based and protein-based circuitry poses new challenges, the resulting ‘toolsets’ of components and novel mechanisms of operation will open up multiple new opportunities for synthetic biology. However, agreed procedures for standardization will need to be placed at the heart of this expanding field if the full potential benefits are to be realized.

Key words: engineering circuitry; RNA; proteins; chemical ligands; yeast

INSIGHT BOX

Synthetic biology is largely based on transcriptional circuitry, in which promoters, terminators and transcription factor sites are utilized to build devices and modules. However, biocircuitry can also be engineered using RNA and/or protein components involved in posttranscriptional steps of gene expression. For example, messenger RNA translation can be modulated by targeting the 5′ untranslated region using RNA-binding proteins, aptamers or aptazymes, whereby degrons or small ligands can feature in the mechanisms used for regulation. Moreover, an even greater range of protein structural and functional properties can also potentially be harnessed for use in engineered regulatory systems. Realization of the full potential of these new directions for synthetic biology will require both further system development and a concerted effort to improve standardization.

INTRODUCTION

In practical terms, synthetic biology primarily refers to the intensified application of state-of-the-art molecular/cell biology techniques (increasingly enabled through automation and machine learning [1]) guided by the application of concepts and strategies that are normally associated with engineering [2, 3]. This distinguishes synthetic biology from discovery bioscience, since it encompasses principles of design and construction, and is thereby obliged to address challenges of predictability and standardization that are more normally associated with engineering disciplines. The construction of synthetic circuitry is being applied to an increasing number of hosts to serve the purposes...
of both fundamental science and biotechnology (e.g. production of recombinant proteins and metabolites, biosensors, etc.).

In the component ‘toolkits’ currently used in synthetic biology, there is a predominance of transcriptional parts, more specifically promoters, transcription factors, transcription factor modulators (metabolites and other ligands) and transcription terminators [4–7]. Formulated somewhat differently, synthetic biology is currently largely about the design, construction and use of engineered DNA [8]. Transcriptional components and modules were inevitably the first choices for synthetic biology because there was a good depth of knowledge of a range of regulatable transcriptional systems in both prokaryotic and eukaryotic organisms. However, there are a number of reasons for exploring options at other points in the gene expression pathway that occur downstream of transcription. Utilizing solely regulatable promoters, for example, can, at least in certain host organisms (including yeast), result in a restricted choice of independently acting systems [9], a less than optimal dynamic range, physiologically (metabolically) disruptive activation/suppression mechanisms [10], and limitations in terms of features such as input–output matching and response times [11]. As a consequence, up-scaling of the size and complexity of logic gate circuitry has generally proved challenging (as has been illustrated by a recent comparison of attempts to build high-complexity circuits [12]).

Here, we examine the scope for development of posttranscriptional biocircuitry, considering the available mechanisms and layers of control. The term biocircuitry is used to describe interactive systems that comprise one or more of the biomolecular types DNA, RNA or protein, thus encompassing a much broader range of components than genetic circuitry. The review also examines the properties of biocircuitry in terms of parameters such as response times (for example, in yeast, translational switching can be initiated in less than 10 min [13]), regulatory dynamic range of regulation, extended scope in terms of diversity of molecular targets, as well as options to create enhanced systems that combine transcriptional and posttranscriptional mechanisms in ways that provide improved functionality. The focus is on Saccharomyces cerevisiae because this is a very important biotech workhorse that is increasingly a host of choice for the engineering of genetic circuitry at different levels. At the same time, it is one of the most tractable and reliable eukaryotic hosts for biocircuitry construction. These characteristics are, as we shall see, important in a number of ways for the wider field. Budding yeast is a good platform for exploring ways to engineer, implement, and ultimately standardize, posttranscriptional toolkits, whereby lessons learned in the process can be expected to inform work using other eukaryotic hosts. Synthetic biology will never achieve (even get close to) the desired-to levels of composability, scalability and robustness [14] without proper standardization (Fig. 1), and we consider the feasibility of defining standards relevant to posttranscriptional constructs. However, at the same time, it is essential to be realistic about the degree to which the noisy analogue circuitry found in living cells can be (re-) engineered to conform to the behavioural characteristics typical of electronic circuitry.

COMPONENTS AND MECHANISMS OF POSTTRANSCRIPTIONAL CONTROL

Recognition of the range of posttranscriptional control mechanisms operating in living organisms initiated with studies performed on Escherichia coli [15–17] but research on eukaryotic posttranscriptional control soon caught up [18–20], and it was not long before yeast became a major focus of work in this area [21]. It has become evident that a number of different types of posttranscriptional component can be incorporated into engineered biocircuitry, either individually or in combinations, depending on the required features of the resulting systems.

The main processes targeted in the posttranscriptional expression pathway are the translation and degradation of mRNA. Achieving specificity largely means utilizing high-affinity control elements in the 5′ or 3′ UTRs, whereby these can be targeted either by RNA-binding proteins (RBPs, usually acting as repressors) or by aptamers, ligands or potentially non-coding RNAs. As we shall see, a highly desirable (and often essential) design feature here is that the activity of such control elements needs to be regulatable. In the context of practical utilization, this review considers the potential for these parts to be incorporated into a usable toolbox for engineering biocircuitry according to principles of predictable construction and functionality. In order to provide a more comprehensive picture of the state of current research on posttranscriptional devices in yeast, this review not only describes posttranscriptional devices that have already been described in the literature, but also considers potential new developments.

RNA-BINDING PROTEINS

One approach to modulating the functionality of target mRNAs is to use RBPs. One potential advantage of pursuing this approach is that there is both an abundance, and a rich variety, of RBPs in living cells [22]. They play important roles in multiple cellular processes, including transcription, splicing, 5′capping, polyadenylation, RNA transport, mRNA translation and degradation. Recent experimental and bioinformatic advances have provided us with an enhanced overview of the multiplicity of structure and function of the RBP families [23, 24]. However, certain characteristics are generally required of an RBP if it is to be a candidate for use as a component of an engineered regulatory device. Most importantly, it needs to be capable of binding with relatively high affinity to a specific RNA structural element, in order to minimize off-target functionality. Moreover, its binding activity needs to be subject to regulation, ideally manifesting an affinity (or intracellular abundance) that can be modulated over an appropriate range with suitable ‘switching’ kinetics. Regulatability of affinity is intrinsic to only a minority of naturally evolved RBPs, and is therefore frequently likely to have to be engineered into the system.

This issue was exemplified early on by the iron regulatory protein, which is naturally found in different versions in various higher eukaryotes. For example, human IRP1 has a very high affinity for the iron-responsive element (IRE) at low iron concentrations, so that the complex between the two strongly inhibits translation via the 5′ UTRs of the ferritin and erythroid 5-aminolevulinic acid synthase mRNAs [25, 26]. The IRP–IRE binding affinity is reduced 50–100-fold at higher iron concentrations and this is the basis for iron-concentration-dependent translational regulation in mammalian cells. However, such iron-dependent regulation is not readily achievable in the yeast cytoplasm, and thus reconstruction of IRP-dependent translational regulation in S. cerevisiae was originally brought about by means of transcriptional regulation of IRP production [27]. A similar approach was taken in exploring the ability of other RNA-binding
proteins (spliceosomal human U1A protein and bacteriophage MS2 coat protein) to act as translational repressors in yeast [28]. This established the principle that any high-affinity RBP can, at least theoretically, be repurposed as a regulatable translational repressor if appropriately targeted to an mRNA.

As we have seen, earlier demonstrations of translational regulation exerted via RNA-binding proteins involved modulation of production of these repressors via regulatable promoters. However, a key objective at this stage of the field’s development is to explore how modulation of the binding of selected RNA-binding proteins can be achieved via posttranscriptional mechanisms, so that de novo engineering of a regulatory device can be rendered independent of transcriptional control (see examples in Fig. 2). One potential strategy is to make the chosen RBP subject to conditional degradation. A number of conditional degron systems have been developed in recent years [29]. Temperature-sensitive degrons were initially employed in yeast predominantly for generating conditional knockouts [30], but they are also of potential interest as fusion components that render RNA-binding protein activity subject to temperature shifts. The target protein is fused at the C-terminal end of a cassette of the form: ubiquitin-Arg-temperature-sensitive dihydrofolate reductase (ts-DHFR).

The use of a temperature shift to induce derepression of a target mRNA will not always be desirable. However, recent research has seen the development of a number of alternatives. In chemically modulated degron systems, small molecules control degron stability. Plant-derived auxin-inducible degron (AID) systems have undergone a number of rounds of development to improve the dynamic range of induction by suppressing auxin-independent degradation of the AID fusion [31, 32]. In an improved auxin response transcription factor (ARF)-AID system, the ARF-interacting domain of ARF called PBI binds to AID to prevent association of the auxin-receptor F-box protein TIR1 with AID, thus preventing auxin-independent degradation by the ubiquitin-mediated proteasomal pathway. In the presence of auxin, on the other hand, TIR1 binding to AID is enabled, thus promoting dissociation of ARF, leading to ubiquitination and proteasome-mediated degradation of the AID-tagged protein. Alternatively, the small molecule associated shutoff (SMASh) system, in which the target protein is tagged (at either terminus) by the HCV NS3 protease together with a degron, has also been tested in yeast. The protease removes itself and the degron from the newly synthesized protein unless an inhibitor such as asunaprevir is added [33]. In contrast, photosensitive degron systems are, in principle, capable of obviating the need for adding small molecule inducers to cell cultures. For example, the light oxygen voltage 2 domain of phototropin from Arabidopsis thaliana has been fused to degrons in such a way as to render access to the degron (and thus activation of protein degradation) subject to (blue light) irradiation [34, 35].

The above consideration of degron systems, while not exhaustive, serves to illustrate that there are options for engineering RBPs that can be regulated via modulation of protein degradation rates. In principle (although not necessarily always in practice), the activity of any type of RNA-binding protein, including those possessing enzyme activity, could be regulated in this manner. For example, future research might reveal whether the cleavage of AGNN-tetraloop-containing RNA elements by the yeast RNase III enzyme Rntp1 [36] can be made subject to degron-mediated modulation (aptamer-dependent control has also been reported for RNase III [37]). Overall, given ongoing progress in identifying novel types of RBP, we can see that there is growing potential for constructing a diverse set of independently acting devices that can be regulated at the protein level. At the same time, it is important to be aware that the degradation properties of any newly engineered fusion between a target protein and a degron are currently not predictable. Thus, variable amounts of optimization work may be necessary in order to achieve satisfactory regulatory characteristics. The design of synthetic RNA-binding domains that can be implemented in such systems may be facilitated by advanced computational methods [38].
CONTROL MEDIATED VIA CIS-ACTING RNA ELEMENTS

An alternative strategy to using a protein repressor to modulate target gene (mRNA) expression is to utilize cis-acting RNA elements that are controlled by other mechanisms. A number of such elements that have evolved naturally in yeast can be co-opted into new types of regulatory circuit. For example, the complex 5′UTR of the yeast transcriptional activator gene GCN4 has been coupled to a recombinant reading frame to enable starvation-induced synthesis of the encoded product [39]. However, it has become evident that a diverse range of systems (with favourably engineered properties) can be more readily built by incorporating exogenous regulatable cis-acting RNA elements into the UTRs of the target mRNA (see, for example, Fig. 2). Considerable progress has been made in the use of cis-acting RNA aptamers as the basis for riboswitch functionality in S. cerevisiae. Following on from the demonstration of aptamer-based ligand-dependent regulation in mammalian cells [40], inducible translational regulation of the yeast cell cycle was achieved by inserting a tetramethylamino-binding aptamer into the 5′UTR of CLB2 [41]. Tetracycline-aptamer-mediated regulation was also found to work effectively in yeast, either by blocking the scanning ribosome when targeted to the 5′UTR [42] or by inhibiting pre-mRNA splicing when targeted to the 5′ splice site of introns inserted into the early part of a reading frame [43]. Further natural/synthetic riboswitches have been characterized in yeast and in other organisms [44].

The most popular approach has been to introduce so-called aptazyme combinations of a self-cleaving ribozyme plus a ligand- (e.g. tetracycline-, theophylline- or neomycin-) binding aptamer into the 3′ or 5′ UTR of a targeted reading frame, rendering mRNA degradation ligand-dependent [45–48]. Published reports have so far indicated that engineered aptazyme systems in yeast can manifest dynamic ranges for ligand-dependent switching of up to 25-fold [47]. Additionally, it has been shown that chemical genetics can be combined with genetic selection techniques to extend the diversity of functional small molecule ligand pairings with synthetic ribozymes in E. coli [49], suggesting that similar approaches could be incorporated into the engineering of aptazyme systems in yeast. Overall, we can expect to observe a growing landscape of synthetic riboswitch-based regulatory systems in this host organism.

The above suggests that it should also be possible to engineer (via a cis-acting aptamer) ligand-dependent regulation into an RBP-encoding mRNA. Regulatory signal inversion would then apply to the production of an active repressor protein (i.e. in terms of the relationship between the ligand or light input and the production of active RBP; Fig. 2). If this device is then coupled to translational inhibition of a downstream target mRNA, the combination of two inhibitory mechanisms would yield, overall, a positive induction module (that can act like a buffered switch; Fig. 3). An alternative architecture, in which RBP production is controlled by a conditional degron instead of a cis-acting aptamer, would manifest the same functionality. Moreover, direct aptamer-dependent regulation of a whole range of target mRNA ORFs is possible. Finally, considering potential future work in broader terms, the design of aptamer-based systems is likely to benefit increasingly from computational methods [50].

CONTROL MEDIATED VIA TRANS-ACTING RNA ELEMENTS?

Intermolecular nucleic acid base-pairing can provide both specificity and high affinity as the basis for precise targeting. Non-coding RNAs, and the proteins that interact with them, perform multiple roles in S. cerevisiae [51]. However, natural microRNAs, which in other organisms guide Argonaute proteins to the 3′UTRs of target mRNAs [52], have not been identified in this organism, and
Figure 3. Posttranscriptional logic gate designs based on components described in this review. All regulatory operations are fulfilled at posttranscriptional steps (here involving mRNA translation). Two illustrative scenarios are shown. (A) Placing aptamer-regulated repressor synthesis in series with repressor-regulated reporter synthesis creates two NOT gates in series, equivalent to a buffered switching module. (B) Targeting two distinct ligand-dependent aptamer switches into one 5′ UTR yields a NOR gate, whereby each input is equivalent to the presence of one ligand molecule. The aptamers can be replaced by aptazymes.

this has so far limited the options for using ncRNAs as parts of engineered circuits compared to plants, animals and viruses [53].

Currently, the most scalable approach to engineering higher order logic gate circuit size and complexity seems to be a largely transcriptional platform involving the use of gRNA-dCas9-Mxi1 complexes in yeast [12]. A major reason for the relatively low signal degradation observed with this approach is likely to be the combination of high specificity and high intermolecular affinities of the system components, coupled to the high transcriptional repression efficacy of dCas9 (whereby Mxi1 is not essential for transcriptional repression in yeast). The effectiveness of this approach demonstrates that there are advantages in using nucleic-acid-interaction-based targeting. At the same time, there are also intrinsic challenges in CRISPR-based strategies [54], and regulatable promoters will need to drive gRNA synthesis at an entry point into the circuitry. Moreover, since this is a transcriptional strategy, we are obliged to ask whether a CRISPR-based system could be targeted to mRNA instead. Recent work with human cell lines has shown that Cas13d endonucleases can be targeted by gRNAs to mRNAs [55], or alternatively that Cas6/Cas13 endonucleases can be targeted to short RNA motifs [56]. We will have to see whether future work generates a substantially posttranscriptional CRISPR-based biocircuitry engineering platform that is applicable to yeast.

PROTEIN-BASED BIOCIRCUITY

This review has so far focused predominantly on how the structure of RNA elements can be engineered in order to create new dynamic functionalities within the cell. RNA-binding proteins can feature in such biocircuitry designs, influencing RNA activities via mechanisms including the modulation of endonuclease cleavage and of translation initiation. However, there are early indications that the full canvas of protein structures and functional capabilities might be exploited to create and implement a substantially larger landscape of protein-based biocircuitry designs [57]. The most exciting aspect here, of course, is the immense diversity of (high-specificity) interactions and functions that nature has demonstrated can be programmed into proteins. At the same time, this structural complexity and flexibility makes precision protein engineering challenging, especially if the aim is to build circuits comprising different interacting protein components. It is interesting to compare the challenges associated with the engineering of (single-stranded) RNA-based biocircuitry, in which the additional degrees of structural and functional complexity relative to (double-stranded) DNA make available distinct mechanisms of action that include inhibitory folded structures, ligand-dependent restructuring, catalysis (cleavage) and sequence-dependent modulation of interactions with ribosomes (and associated factors). Engineering protein circuitry offers an even wider range of capabilities related to molecular binding specificities, conformational adaptability, signalling, transport and catalysis, but generally requires a shift to a very high level of complexity and difficulty in terms of design and implementation.

ANALOGUE VERSUS DIGITAL RESPONSES

Naturally evolved regulatory devices generally manifest analogue responses to changes in the concentrations of their effectors, whether these are proteins, metabolites or chemical ligands. In other words, the degree of activation or suppression responds over a continuous range to progressive changes in
effector molecules. A more digital type of response can be observed if the regulatory device shows ultrasensitivity, often reflected in a high Hill coefficient for binding of the effector molecule [11, 58]. This applies both to RNA and DNA targets, and therefore more digital-type behaviour can, at least to some extent, be engineered into posttranscriptional devices. One means of (partially) achieving this is to modify the effector (for example, a repressor protein) or the binding target (for example, an RNA aptamer) in order to increase the Hill coefficient for binding. On the other hand, it has been suggested that the steps needed to ‘digitize’ biocircuitry behaviour can create unacceptably high costs in terms of operational energy efficiency [59]. The imprecision created by system noise represents a related problem. Noise suppression, for example via measures that minimize cell-to-cell heterogeneity, generally requires additional energy expenditure.

COMPOSABILITY AND SCALABILITY

Precision in the rational engineering of any new system relies on the accurate quantitative characterization of the components that are designed into that system [2, 60]. This is the basis for achieving the composability, scalability and robustness that are still largely missing from the synthetic biology field. Indeed, there has been increasing awareness that the lack of proper standardization is hampering the progress of biological engineering activities. Significant initial steps have been made towards satisfying the requirements for achieving predictable gene expression outcomes from synthetic DNA circuitry in bacterial hosts [61]. Such efforts have been underpinned by a variety of assembly frameworks and strategies, including the International Open Facility Advancing Biotechnology (BIOFAB) and the BioBricks restriction enzyme assembly standard [8, 62], Gibson Assembly, Golden Gate, Modular Overlap-Directed Assembly with Linkers (MODAL), Standard European Vector Architecture (SEVA) and others [64]. However, looking at the bigger picture, standardization for DNA-based components and circuits is not widely applied to the variety of hosts in use at the present time [65], and is very limited for other biomolecular parts and systems involving RNA and proteins. Enabling the truly predictable engineering of biological systems will require the investment of considerable time and energy into standardization at multiple levels, including workflows, system descriptions, measurement techniques, international systems of units and data sharing (Fig. 1 [60, 66–71]).

It is self-evident that the presence of suitable abundance levels of posttranscriptional components, whether these be riboswitch-containing mRNAs or RNA-binding proteins, relies upon transcription. Stable rates of synthesis of these entities can be maintained using constitutive promoters. An important step here is to choose promoters that manifest appropriate values for the transcription rate and transcription stochasticity. A recent study of engineered expression constructs in S. cerevisiae has illustrated how we can quantitate (and optimize) both the average value, and the variation, of the copy number of mRNA molecules per cell [72]. Both smFISH [72] and single-cell RNAseq [73] can be used to determine mRNA-copy-per-cell distributions across cell populations. This type of information is essential for allowing us to balance the optimal dynamic range of posttranscriptional regulation as exercised on each mRNA molecule against the potentially disadvantageous effects of cell-to-cell variation in mRNA copy number. There is a temptation to focus on the maximization of the transcription rate in order to reduce the level of mRNA copy number noise, but this will not necessarily result in a dynamic range of operation of a regulatory mechanism that matches the intended functionality. This can be illustrated by considering the use of a strong promoter to drive transcription of a degron:RNA-binding protein expression construct that will form part of a regulatory circuit. The high transcription rate will help to suppress gene expression noise, but it may also increase the lower limit (base level) of the degron-mediated regulatory range to a level that results in only partial activation of translation of the mRNA targeted by the RNA-binding protein (Fig. 3). These considerations mirror the challenges inherent in input-output tuning for biocircuitry dependent on transcriptional regulation (11), and can be addressed in analogous ways. Computational modelling can help in developing strategies for achieving the right balance in terms of the abundance/activities of all components.

Having selected the constitutive transcription platform used to generate the posttranscriptional components that will build the required biocircuitry (where appropriate, by fine-tuning a synthetic promoter [4, 5, 7]), the next step is to optimize the dynamic range of the regulatory mechanism operating at the mRNA level. Consider an example in which a ligand-dependent aptamer or aptazyme has been engineered into the 5′UTR of an RNA-binding-protein-encoding reading frame (Fig. 3). The dynamic range of the ligand-dependent modulation of the aptamer/aptazyme can be optimized by modifying the aptamer/aptazyme structure [41–44]. The priority will likely be to measure the full extent of the dynamic range of aptamer/aptazyme-mediated regulation in terms of the intracellular RNA-binding protein abundance values. This can be achieved using quantitative mass spectrometry, benchmarked by incorporating peptide standards into the cell extract samples [74, 75]. However, in the interests of pursuing a more readily accessible set of procedures for standardization (see next section), it makes sense to replace the RNA-binding protein with a reporter gene encoding a high-intensity fluorescence protein such as yEGFP. In this way, it becomes feasible to define, at least partially, the profile of any such 5′UTR-mediated posttranscriptional regulatory system in terms of fold-changes in the fluorescence output of yeast cells at different degrees of activation of the aptazyme mechanism.

STANDARDIZATION

There has been a substantial amount of work published on the standardization of design protocols and representation [76–80] as well as (automated) workflows for the assembly of DNA circuitry [81, 82]. In contrast, the development of system characterization protocols and ‘data sheets’ has lagged behind. Returning to the question of the value of standardizing the characterization of system performance using fluorescence reporters, this approach can of course only be upheld if the conditions under which that standardization is performed are strictly defined, and strictly adhered to, by the wider community. Ideally, the starting point should be a specific (genomic) construct in a defined strain, thus avoiding the excessive cell-to-cell heterogeneity in gene expression associated with variations in plasmid copy numbers. An engineered strain should be grown under a standardized set of conditions, and the gene expression output should be measured following standard procedures using defined techniques. The synthetic biology community has yet to agree on a universally applicable set of standards for system characterization. This is unfortunate, since the absence of agreement on standardization is problematic in two ways. First, it holds back progress in our understanding of how to successfully engineer biocircuitry
that follows predictable behavioural patterns. Second, the lack of standardization acts as a barrier to the incorporation of engineering principles that would otherwise facilitate the realization of valuable biotechnological projects. A key challenge is therefore to formulate widely acceptable proposals for international standardization procedures and protocols.

In the context of what has been considered in this review, the initial standardization process could include characterization of the regulatory device (e.g. a 5′UTR-aptazyme-controlled mRNA) using techniques (including those mentioned above) that quantitate the intracellular molecular abundance of mRNA and protein. This could be accompanied by parallel measurements (using a plate reader and a flow cytometer) of the fluorescence output of a version of the genomic construct under investigation that encodes a fluorescence-reporter reading frame. Moreover, it has been demonstrated that the use of independent calibrants and standardized protocols in fluorescence measurements on cells leads to major improvements in precision [83, 84]. Overall, different combinations of the quantitative techniques outlined above could be applied depending on the degrees of accuracy and precision that researchers aspire to (Fig. 4).

A further step towards an ultimate level of standardization would be to involve computational modelling of the relevant synthesis and degradation rates in different system states so as to define the kinetic basis of the measured steady-state abundance values. This would need to include, for example, values for translation initiation events per second (TIPS or TIS\(^{-1}\)) or polypeptides completed per second (PCPS or PCS\(^{-1}\)) for the output protein (generated from the ‘receiver’ mRNA), as the posttranscriptional parallel to measurements of polymerases per second (PoPS (8)). The resulting data can be used to create a standardized quantitative profile and time course of the temporal properties of any activation or suppression process. Such an approach is also of value in the context of incorporating information about system stochasticity and cell-to-cell heterogeneity into the standardization of engineered biosystems.

**STOCHASTICITY, RETROACTIVITY AND ORTHOGONALITY**

Up to now, biosystem data sheets have generally featured population-average values for any given engineered device or module. However, since noise is an intrinsic property of biological systems, there is a strong argument for including at least some stochasticity data in any standard system description. At a minimum, it would be advisable to conduct flow cytometry analysis of the cell-to-cell heterogeneity of gene expression outputs. Such data are important for developing an understanding of the noise intrinsic to the operation of engineered biocircuitry across the cells in a host-organism population. They are also valuable for any researcher wishing to ‘tune’ the stochasticity characteristics of such biocircuitry. It has been observed that the noise associated with a given promoter can be modulated independently of the mean rate of transcription that it drives in the cell [85, 86]. Other work then showed that gene expression noise can also be modulated at the translation level in yeast by modifying the structure of the 5′UTR [72]. Moreover, the incorporation of negative feedback can be utilized to help control, and buffer, gene expression noise [87]. Given the different ways in which the inherent variability of gene expression characteristics can influence system performance and population fitness [88], the inclusion of at least basic stochasticity data in data sheets on engineered biocircuitry would be advantageous.
The dynamic and steady-state characteristics of interactions between devices/modules can also be affected by retroactivity [89–92]. This term refers to the effect of modulatory interactions of a molecule with a downstream device/module on the availability of that molecule to the upstream device/module that generates it. Reference to Figure 3A provides an example as to how certain features of posttranscriptional regulatory mechanisms deserve special consideration in terms of vulnerability to retroactivity effects. The repressor protein generated by the upstream device may need to bind to a high-affinity binding site on multiple copies of the downstream target mRNA. Particular consideration will therefore need to be given to ensuring that an optimal (steady-state) ratio of repressor molecules to mRNA target molecules is maintained in the cell. This should normally be achievable through adjustments of the constitutive transcription rates for the repressor-encoding and reporter-encoding genes and of the repressor affinity for its mRNA binding site. Overall, the consequences of restricted levels of retroactivity in circuitry of this kind designed to perform as steady-state regulatory systems are likely to be of limited significance compared to the impact of retroactivity that underlines important dynamic features of, for example, signalling systems.

Ideally, the synthetic biologist would prefer to isolate biomolecular activities encoded by engineered biocircuitry as far as possible from the cell environment. A high degree of isolation is desired, although not necessarily always achievable, because it minimizes negative effects both on the viability of the host organism and on the expected performance of the biocircuitry. It is also important to recognize that ‘orthogonalization’ may need to be engineered into a system, both to minimize crosstalk between separate domains of engineered genetic circuitry, and to reduce interference between engineered biocircuitry and components that are endogenous to the host. A good starting point, certainly in yeast, is to perform genomic integration of DNA biocircuitry components. This limits the variability of the copy number of the DNA components compared to that observed with extra-chromosomal plasmids, and thus narrows variability in the negative consequences of interactions between host-cell and engineered biocircuitry. Orthogonality issues have inevitably been addressed in most detail, and with most success, in bacterial hosts [93–96]. Progress in yeast has focused primarily on engineered promoters [4, 97], and considerable work remains to be done on the orthogonality of posttranscriptional biocircuitry. However, additional time and effort need to be invested in characterization of posttranscriptional components, and combinations thereof, in order to provide a more accurate picture of the benefits that can be achieved in terms of performance relative to transcriptional circuitry. Moreover, there is scope for biocircuitry built using components at the DNA, RNA and protein levels in the gene expression pathway to be combined together, and this approach has considerable potential that remains to be realized. The use of parallel mechanisms at different levels of the gene expression pathway can be expected to facilitate optimization of key system properties.

At this relatively early stage of development of the posttranscriptional area of biocircuitry engineering, it would be wise to incorporate rigorous standardization as a core element of the research effort. This will greatly accelerate progression towards predictability in design and in design implementation. There is an urgent need for agreement on a set of principles and procedures that will apply to posttranscriptional circuitry, at least in selected host organisms. This could be established as part of a broader set of guidelines that, for example, could extend to include transcriptional regulatory devices. There are strong arguments (extensive established procedures for genetic manipulation, ease of use etc.) for yeast, particularly S. cerevisiae, to play a key role in pioneering efforts in this direction. This could grow in parallel to the developing standardization work being done with E. coli, and we would hopefully see analogous standardization projects developed for other major synthetic biology hosts. It would also be beneficial to acknowledge formally that cell-to-cell heterogeneity is an intrinsic feature of living organisms that influences system performance by incorporating, at least for microbial hosts, some form of stochasticity/variation data in (agreed) standardization procedures.

**FUNDING**

This work was supported by the BioRoBoost consortium [funded by the EU Horizon grant H2020-NMBP-TR-IND-2018-2020/BIOtec-01-2018 (CSA), project ID 820699] and by a Biotechnology and Biological Sciences Research Council (BBSRC) grant (BB/M017982/1) to the Warwick Integrative Synthetic Biology centre (WISB).

**CONFLICT OF INTEREST**

The author declares that he has no conflict of interest.

**REFERENCES**

1. Carbonell P, Radivojevic T, Martin HG. Opportunities at the intersection of synthetic biology, machine learning and automation. ACS Synth Biol 2019;8:1474–7.
2. Purnick PEM, Weiss R. The second wave of synthetic biology: from modules to systems. Nat Rev Mol Cell Biol 2009;10:410–22.
3. Khalil AS, Collins JJ. Synthetic biology: applications come of age. Nat Rev Genet 2010;11:367–79.
4. Khalil AS, Lu TKU, Bashor AJ, et al. A synthetic biology framework for programming eukaryotic transcription functions. Cell 2012;150:647–58.
5. Redden H, Alper HS. The development and characterisation of synthetic minimal yeast promoters. Nat Comm 2015. doi: 10.1038/ncomms8810.
6. Machens F, Balazadeh S, Mueller-Roeber B, et al. Synthetic promoters and transcription factors for heterologous
Integrative Biology

24. Giudice G, Sanchez-Cabo F, Torroja C, et al. ATtRACT—a synthetic biological parts and devices. Nat Biotechnol 2008;26:787–93.

23. Kishore S, Luber S, Zavolan M. Deciphering the role of RNA-binding proteins in yeast. Brief Funct Genomics 2010;9:1492–553.

22. Corley M, Burns MC, Yeo GW. How RNA-binding proteins control gene expression in eukaryotic cells. Trends Biochem Sci 1995;20:191–7.

21. Corley M, Caughman SW, Harford JB, et al. Translational repression by a complex between the iron-responsive element and cellular function. Chem Biol 2013;20:619–26.

20. Mata J, Marguerat S, Bähler J. Post-transcriptional control of Messenger Stability. Mol Syst Biol 2013;9:552–60.

19. Gale M Jr, Tan S-L, Katre MG. Transcriptional control of viral gene expression in eukaryotes. Microbiol Mol Biol Rev 2000;64:239–80.

18. McCarthy JEG, Kollmus H. Cytoplasmic mRNA-protein interactions in eukaryotic gene expression. Trends Biochem Sci 1995;20:356–60.

17. Shimoni Y, Friedlander G, Hetzroni G, et al. Regulation of GCN4 protein-mediated repression in yeast Saccharomyces cerevisiae. J Biol Chem 1998;273:12686–702.

16. McCarthy JEG, Gualerzi C. Translational control of prokaryotic gene expression. Trends Genet 1990;6:78–85.

15. Gold L. Posttranscriptional regulatory mechanisms in Excherichia coli. Annu Rev Biochem 1988;57:199–223.

14. Lucks JB, Qi L, Whitaker WR, et al. Toward scalable parts families for predictable design of biological circuits. Curr Opin Microbiol 2008;11:567–73.

13. Albrecht G, Mösch H-U, Hoffmann B, et al. Monitoring the yeast GCN4 protein-activated response in the yeast Saccharomyces cerevisiae. J Biol Chem 1992;267:12696–702.

12. Gander MW, Vrana JD, Voje WE, et al. Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. ACS Synth Biol 2017;6:547–67.

11. Ang J, Harris E, Hussey BJ, et al. Tuning response curves for synthetic biology. ACS Synth Biol 2013;2:547–67.

10. Ottoz DSM, Rudolf F, Stelling J. Inducible, tightly regulated auxin-inducible degron system preserves native protein levels and enables rapid and specific protein depletion. Genes Dev 2019;33:1–15.

9. Da Silva NA, Srikrishnan S. Introduction and expression of genes with designated base motifs. Nucleic Acids Res 2014;42:e130.

8. Canton B, Labno A, Endy D. Refinement and standardization of synthetic biological parts and devices. Nat Biotechnol 2008;26:787–93.

7. Liu R, Liu L, Li X, et al. Engineering yeast artificial core promoters with designated base motifs. Microb Cell Factories 2020. doi: 10.1186/s12934-020-01305-4.

6. Harford J. Control of Messenger Stability. San Diego, Academic Press, 1993, 239–66.

5. Oliveira CC, Goossen B, Zanchin NIT, et al. Translational repression by the human iron-regulatory factor (IRF) in Saccharomyces cerevisiae. Nucleic Acids Res 1993;21:5316–22.

4. Goossen B, Caughman SW, Harford JB, et al. Translational repression by a complex between the iron-responsive element of ferritin mRNA and its specific cytoplasmic binding protein is position-dependent in vivo. EMBO J 1990;9:4127–33.

3. Kishore S, Luber S, Zavolan M. Deciphering the role of RNA-binding proteins in yeast. Trends Biochem Sci 2010;35:4179–85.

2. Corley M, Burns MC, Yeo GW. How RNA-binding proteins interact with RNA: molecules and mechanisms. Mol Cell 2020;78:10–29. doi: 10.1016/j.molcel.2020.03.011.

1. Shimoni Y, Friedlander G, Hetzroni G, et al. Regulation of gene expression by small non-coding RNAs: a quantitative view. Mol Syst Biol 2007;3:138. doi: 10.1038/msb4100181.
90. Kim KH, Sauro HM. Fan-out in gene regulatory networks. J Biol Eng 2010;4:16.
91. Jayanthi S, Nilgiriwala KS, Del Vecchio D. Retroactivity controls the temporal dynamics of gene transcription. ACS Synth Biol 2013;2:431–41.
92. Pantoja-Hernandez L, Martinez-Garcia JC. Retroactivity in the context of modularly structured biomolecular systems. Front Bioeng Biotechnol 2015;3:85. doi: 10.3389/fbioe.2015.00085.
93. An W, Chin JW. Synthesis of orthogonal transcription-translation networks. Proc Natl Acad Sci U S A 2009;106:8477–82.
94. Rao CV. Expanding the synthetic biology toolbox: engineering orthogonal regulators of gene expression. Curr Opin Biotechnol 2012;23:689–94.
95. Liu CC, Jewett MC, Chin JW, et al. Toward an orthogonal central dogma. Nat Chem Biol 2018;14:103–6.
96. Costello A, Badran AH. Synthetic biological circuits within an orthogonal central dogma. Trends Biotechnol 2021;39:1. doi: 10.1016/j.tibtech.2020.05.013.
97. Blount BA, Weening T, Vasylechko S, et al. Rational diversification of a promoter providing fine-tuned expression and orthogonal regulation for synthetic biology. PLoS One 2012;7:e33279.