ANNUAL REVIEWS

*Annual Review of Biophysics*

# Biomolecular Systems Engineering: Unlocking the Potential of Engineered Allostery via the Lactose Repressor Topology

Thomas M. Groseclose,[1] Ronald E. Rondon,[1] Ashley N. Hersey,[1] Prasaad T. Milner,[1] Dowan Kim,[1] Fumin Zhang,[2] Matthew J. Realff,[1] and Corey J. Wilson[1]

[1]School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, USA; email: corey.wilson@chbe.gatech.edu

[2]School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, USA

Annu. Rev. Biophys. 2021. 50:303–21

First published as a Review in Advance on February 19, 2021

The *Annual Review of Biophysics* is online at biophys.annualreviews.org

https://doi.org/10.1146/annurev-biophys-090820-101708





## Keywords

allostery, engineered transcription factors, synthetic biology, antirepressors

## Abstract

Allosteric function is a critical component of many of the parts used to construct gene networks throughout synthetic biology. In this review, we discuss an emerging field of research and education, biomolecular systems engineering, that expands on the synthetic biology edifice—integrating workflows and strategies from protein engineering, chemical engineering, electrical engineering, and computer science principles. We focus on the role of engineered allosteric communication as it relates to transcriptional gene regulators—i.e., transcription factors and corresponding unit operations. In this review, we (*a*) explore allosteric communication in the lactose repressor LacI topology, (*b*) demonstrate how to leverage this understanding of allostery in the LacI system to engineer non-natural BUFFER and NOT logical operations, (*c*) illustrate how engineering workflows can be used to confer alternate allosteric functions in disparate systems that share the LacI topology, and (*d*) demonstrate how fundamental unit operations can be directed to form combinational logical operations.

## Contents

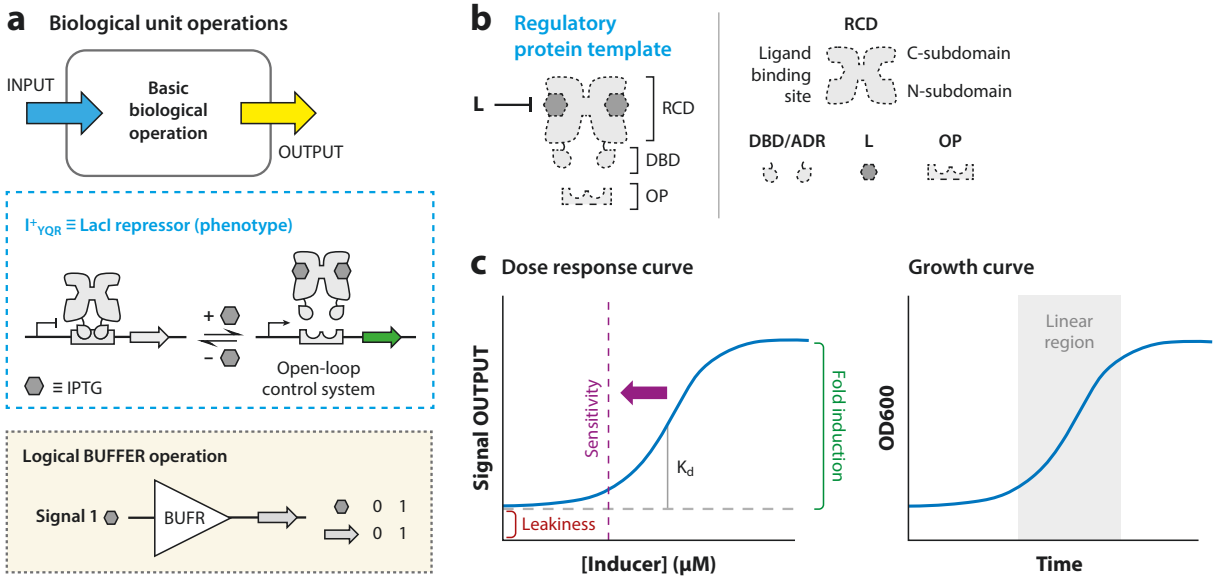




## 1. INTRODUCTION

The ongoing boom in the life sciences has significantly increased our capacity to understand and engineer biological systems with interesting functions and novel capabilities (49, 72). Several fields that have benefitted from—and contributed to—this biological revolution are synthetic biology (5, 28, 66), protein engineering (16, 25, 26, 65), metabolic engineering (20, 31, 32, 45), and systems biology (2, 10, 43). One of the major impediments in synthetic biology and related fields is the lack of reproducibility across laboratories in the performance and production of genetic parts. The general variation in the performance of genetic parts limits the systematic bottom-up development of complex genetic networks. To mitigate variation in performance of nonbiological systems, chemical engineers have developed fields of study such as process design, which enables the systematic organization of many unit operations (i.e., in terms of choice and sequence) to produce a desired physical and/or chemical transformation. Typically, process design starts at the conceptual level (in silico) and ends with fabrication and testing of the system. Likewise, electrical engineering workflows (e.g., circuit design) leverage systematic approaches to facilitate the bottom-up development of systems with highly predictive performance. Unlike systems derived via chemical engineering and electrical engineering workflows, the genetic parts used throughout synthetic biology and metabolic engineering lack systematic development and characterization (i.e., standardized unit operation design and related metrology), which can result in lower performance fidelity of complex systems.

From a chemical engineering standpoint, genetic parts can be viewed as biological unit operations (BUOs) (**Figure 1*a***). In turn, multiple BUOs working in cooperation constitute a complex biological process. Generically, a unit operation represents a fundamental functional object that




## Figure 1

Using the lactose repressor (LacI) topology to illustrate biological unit operations (BUOs). (*a*) A BUO converts an INPUT to a predictable OUTPUT. Wild-type LacI ($I^{+}_{YQR}$), which functions as a Boolean logic BUFFER operation, converts the chemical isopropyl-β-D-thiogalactoside (IPTG) INPUT to a protein OUTPUT, green fluorescent protein (GFP). The schematic shows the putative mechanism of the LacI BUFFER (repressor) operation, with truth table. Ligand (IPTG) binding causes an allosteric (conformational) shift in LacI, resulting in dissociation from the DNA operator, allowing transcription of the downstream gene. (*b*) Building a regulatory protein. A regulatory protein consists of a regulatory core domain (RCD), which exhibits binding specificity for a ligand (L), and a DNA binding domain (DBD), which exhibits binding specificity for a DNA operator (OP). RCDs and DBDs have been shown to be somewhat modular, allowing for pairings to create novel regulatory proteins. DBDs can be altered to confer alternate DNA recognition (ADR). (*c*) Dose response (*left*) and growth curves (*right*). Dose response follows a sigmoidal (Hill) function with respect to inducer ligand concentration, in this case, for a BUFFER operation (e.g., LacI). A basal level of OUTPUT, or leakiness, is characteristic of gene regulatory elements, and a fold induction (change) is the ratio of maximum (saturated) OUTPUT to basal leakiness (or OUTPUT in the 1 state divided by that in the 0 state). Protein sensitivity to ligand can be measured in terms of a dissociation constant (or $K_d$). Note that IPTG binding to the repressor is noncooperative with a Hill constant of 1 (8). The observed dose response (sigmoid) is a consequence of regulation of the OUTPUT and not the effect of ligand binding. A growth curve (*right*) for typical bacteria, with optical density ($OD_{600}$) as a function of time, is shown. Logical operations are assayed within the linear growth region (*gray shaded*). Aspects of this figure are adapted from References 21 and 52, both licensed under CC BY 4.0 (**https://creativecommons.org/licenses/by/4.0/**).

operates with a high degree of fidelity regardless of the process or location of the given operation—provided that the system remains within the performance boundary. In turn, the unit operation's internal complexity can be abstracted away to their process functionality, which chemical engineers can then manipulate free of the cognitive load of their actual implementation. In principle, BUOs can be developed and defined to a similar level of objective performance. Once this is achieved, BUOs can be structured systematically to produce higher-order functionality using strategies, inspired by electrical engineering and computer science principles, with greater fidelity and sophistication (44). Thus, one solution to the current BUO performance fidelity problem is to integrate chemical and electrical engineering workflows. We can leverage and develop certain engineering principles (i.e., circuit design, methodical unit operation development and BUO process design, process simulation, and formal metrology) to guide the development, choice, sequence, and testing of BUOs and complex biomolecular systems derived from the standardized unit operations. Ultimately, BUO structures can be implemented as hardware and software concepts from computing and control systems to enable the production of unprecedented biological programs and related machines. In this review, we present the convergence of these many allied fields (synthetic

biology, protein engineering, and chemical and electrical engineering principles) into what we define as biomolecular systems engineering, specifically through the lens of engineered allostery, which is both one of its key products and one of its key components. In doing so, we showcase how engineered allostery is advancing a discipline that promises to have long-reaching impacts in bioindustry, biotechnology, and human health.

## 2. FOUNDATIONAL REGULATORY UNIT OPERATIONS

Allostery is broadly defined as the signal propagation between at least two functional surfaces upon external stimulus (40, 41). Purportedly, the signal propagation mechanism that influences a given allosteric conformational state involves a defined path along a unique network of residues (34) or via an ensemble of different transition paths (42, 59). Mechanistically, for a two-surface system, (*a*) an effector interacts with one functional surface (e.g., ligand binding site), causing a disturbance to the surface residues; (*b*) this disruption is propagated through a residue network (or ensemble); and (*c*) this interaction typically results in a conformational change in the protein and subsequent activation of the second functional surface (e.g., DNA binding site). Allosteric communication is a hallmark of a wide variety of vital biological functions (1, 11, 12, 39, 42, 46, 48, 54, 71). In this section, we focus on allosteric communication as it relates to regulatory BUOs—i.e., integral unit operations used throughout synthetic biology and biotechnology. Regulatory BUOs can be represented at the transcriptional or translational levels; the former is more commonly employed in the development of synthetic gene networks via the use of transcription factors to activate or repress genes of interest. Broadly speaking, transcription factors are DNA-binding proteins capable of blocking (or recruiting) RNA polymerase activity at the site of genetic promoters, and these functions can be combined in modular ways to engineer synthetic gene networks (37). For the most part, many of the preliminarily engineered bacterial gene circuits were based on a core set of repressors, namely, tetracycline repressor (TetR), lactose repressor (LacI), and bacteriophage λ cI repressor (17, 22–24), which have been extensively studied. In this review, we focus on unit operations composed of a transcription factor and a cognate DNA operator element that regulate the expression of an observable protein [e.g., green fluorescent protein (GFP)]. In the context of control systems engineering, the putative BUO can be represented as an open-loop single-input single-output (SISO) unit operation (**Figure 1*a***).

### 2.1. Structure and Function of a Canonical Regulatory Unit Operation

The LacI regulatory protein has been a workhorse in synthetic biology (6, 17, 56). Given the large number of putative transcription factors represented by the LacI structure and function, we can use information gleaned from this system as a guidepost (or reference) to introduce the role of allosteric communication (and action) into the context of biomolecular systems engineering. Structurally, monomeric LacI is composed of two domains: (*a*) the DNA binding domain (DBD), which interacts with operator DNA, and (*b*) the regulatory core domain (RCD), where the ligand binding pocket is located, at the cleft between the N subdomain and the C subdomain (**Figure 1*b***). Several biophysical (4, 15, 30) and biochemical (60, 69, 70) studies have inferred that the N subdomain is the allosteric component—complemented by the ligand-binding and DNA-binding functional surfaces. A pair of DBDs are required to interact specifically with a DNA element (7, 68). In addition, studies have demonstrated that induction of the LacI protein requires at least two molecules of isopropyl-β-D-thiogalactoside (IPTG), one for each monomer (14). Thus, the functional unit of LacI is regarded as a dimer (30)—although LacI is not classified as a cooperative system (8) (**Figure 1*c***). As it is a model protein, structural and sequence-phenotype information

for LacI is broadly available. X-ray crystallography and nuclear magnetic resonance (NMR) studies have deduced that LacI exists primarily as two steady-state conformations (or ensembles), a DNA-bound conformation and a ligand-bound one (**Figure 1*a***). Upon binding the ligand IPTG, the N subdomain shifts relative to the C subdomain, which remains static, and signal propagation causes a disordering of the DBD (69). The binding of IPTG causes LacI to release operator DNA, allowing (or inducing) the production of an OUTPUT (i.e., gene expression). In logical terms, gene expression goes from the FALSE to the TRUE state upon the addition of an INPUT signal.

## 2.2. Generating a Sequence–Function Map in the LacI System: Phenotyping Allostery

Suckow et al. (60) have performed extensive genetic analysis of the LacI system akin to deep-mutational scanning—albeit over 25 years ago. In their study, partial site saturation across the wild-type LacI primary structure, generating over 4,000 mutants, followed by phenotyping and comparison with the crystal structure, was used to make a sequence–function map to elucidate positions that are relevant for stability and allosteric function. Phenotypes of mutants were binned into three major classes: (*a*) the $I^+$ or wild-type repressor phenotype; (*b*) the $I^S$ or super-repressor phenotype, which cannot dissociate from DNA; and (*c*) the $I^-$ or nonfunctional phenotype, which results in constitutive gene expression, as it seemingly cannot associate with DNA. The $I^S$ phenotype may be caused by a disruption in the native allosteric network, while many instances of the $I^-$ phenotype were likely the result of unfolded (destabilized) proteins. Suckow et al. found that LacI was generally tolerant to mutations, although several positions were particularly sensitive and prone to generating $I^S$ mutants. Analogous functional maps for LacI systems were produced by Meyer et al. (36), with similar results in terms of phenotypic outcomes. Notably, mutations that confer super-repression (i.e., block or disrupt allosteric communication) typically occurred throughout the N subdomain (60).

It is important to note that Suckow et al. (60) and Meyer et al. (36) measured LacI performance (phenotype) by way of the regulation of β-galactosidase as the OUTPUT interface using a blue-white screen. In this assay, when β-galactosidase is produced, the enzyme cleaves X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside), forming a blue pigment that serves as a proxy for LacI-regulated OUTPUT. In the Suckow et al. assay, colonies were grown on indicator plates, then permeabilized in sodium dodecyl sulfate and chloroform solution. Absorbance of the solution was then measured at 420 nm, and phenotypes were subsequently assigned. Comparably, Meyer et al. determined the phenotype of LacI variants by measuring the optical density directly at the colony-forming unit, with and without IPTG. Contemporary assessments of this unit operation are primarily conducted in solution via a microwell plate assay (or via cytometry) by way of a direct observable OUTPUT (e.g., GFP). While the results of the aforementioned assays are in reasonable agreement, differences in phenotypic outcomes do arise (51). Accordingly, differences in transport phenomena must be considered, in addition to catalytic limitation, when choosing an assay for metrology. In general, the GFP screen in solution is regarded as the more accurate system in terms of evaluating true performance metrics, as this assay does not involve a coupled catalytic process and can be facilely measured in solution. Metrology is discussed in greater detail in Section 5.3.

## 2.3. Allosteric Communication In Silico

Targeted molecular dynamics (TMD) simulations using the crystal structures of two LacI conformations [repressed state (Protein Data Bank ID 1EFA) (4) and induced state (Protein Data Bank ID 1TLF) (19)] suggest that the allosteric signal propagates from the inducer binding pocket

through three interconnected routes that span from the N subdomain monomer–monomer interface and the binding pocket on one side to the hinge and DBDs on the other (18). TMD also indicates that many of the putative allosteric positions are synonymous with primary structure positions that are sensitive to perturbation, such that many of these positions also confer superrepression upon perturbation. In other words, a given $I^S$ point mutation can be regarded as a block in allosteric communication. Notably, Flynn et al. could not explicitly study signal propagation in the DBD in either the induced or the repressed state (due to TMD limitations). Namely, in the induced state, the DBD (i.e., amino acid residues 1–60) is intrinsically disordered—thus, any trajectory in this region cannot be calculated because the TMD simulation requires structural information for both the repressed state and the induced state for every position that is evaluated (19). In general, disordered sequences can potentially bind to multiple partners (73). Thus, the absence of this mechanistic information in the TMD study constitutes a significant gap in our understanding of allosteric communication. Accordingly, TMD cannot be used to design or predict new allosteric functions. TMD is one of many in silico strategies used to explore allosteric communication (71); however, to the best of our knowledge, the studies discussed above are the most extensive studies of fundamental allosteric communication in the LacI system. Accordingly, this collection of studies [i.e., of TMD (18) complemented by phenotyping via deep-mutational scanning (60)] introduced the first putative map and action of allosteric communication in the LacI topology, although their results are incomplete.

Clearly, the studies discussed above cannot provide an objective metric for allosteric communication, which will be critical to the de novo design of this brand of signal propagation and function. One potential solution to this problem comes in the form of statistical coupling analysis. Suel and colleagues (29, 33, 50, 61) developed a sequence-based statistical mapping strategy (i.e., a nonpairwise approach) to potentially identify networks of residues that mediate allosteric communication in proteins. These statistical coupling analysis studies revealed that nonallosteric residues (most sites in a given protein) act in an evolutionarily independent manner and are uninfluenced by perturbations (mutation). However, allosteric residues (a small number of positions in a given protein) form coevolving linked networks throughout the structure, producing architectures for mediating long-range communication in proteins. In addition, statistical coupling analysis studies suggest that allosteric residues overlap with residues that are important for protein stability, convoluting the definition and quantification of a given allosteric network. Moreover, statistical coupling analysis and related methodologies are thermodynamic in nature. Therefore, statistical coupling analysis does not provide intrinsic information regarding the underlying mechanism of the interactions between residues, which is the basis for nonheuristic rational design. However, statistical coupling analysis suggests that a hallmark of a given allosteric position is an extreme sensitivity to perturbation. This is evidenced (in part) via the LacI system when deep-mutational scanning is used to test conferred allosteric positions for mutational tolerance (36), although assessment of coevolution (coupling) of these residues has not been evaluated. Considering all of the points discussed above, the de novo design of allosteric communication via contemporary in silico methodologies is seemingly intractable—but new developments (71) continue to show promise. Despite current in silico limitations, other engineering strategies to confer allosteric communication (including functional surfaces) have shown great success; we discuss these in the following sections.

## 3. SINGLE-INPUT SINGLE-OUTPUT UNIT OPERATIONS: BUFFER LOGIC

From the perspective of control systems engineering, the fundamental LacI regulatory operation in synthetic biology can be viewed as an open-loop SISO unit operation (**Figure 1*a***). This SISO

unit operation illustrates the repressor phenotype for the LacI system ($I^{+}_{YQR}$). Assuming that the INPUT signal (IPTG) is introduced in excess ($\gg K_d$ value), and that the unit operation is at steady-state—i.e., within the linear region of the growth phase—as a first approximation, we can represent this SISO as a fundamental Boolean logical operation. Specifically, we can regard this system as a BUFFER logic gate—in which the IPTG signal INPUT results in an OUTPUT (i.e., a protein or, more fundamentally, mRNA) (**Figure 1*a*,*c***). Notably, the linear region is also representative of the continuous growth conditions used in a bioreactor (67). Thus, a given BUFFER logical operation can be executed either in a batch system or under continuous growth conditions. Given that we have a proper metrology to measure and report relative performances, we can treat Boolean biological unit operations in a similar fashion to circuit components on a breadboard—i.e., as electrical engineers would do. An illustration of the expansion to combinational logic is discussed in Section 5. In the interim, the focus is on engineering allosteric communication to create systems of BUFFER unit operations—and, later, systems of the antithetical NOT logical operation. Notably, the designation of other systems that are regulated via inducible promoters (similar to LacI) have also been regarded as BUFFER (as well as NOT) logical gates by other groups (6, 63). The BUFFER gate designation has also been used in synthetic biology to describe biological memory unit operations that employ similar logical controls (57) (a discussion of this work is beyond the scope of this review).

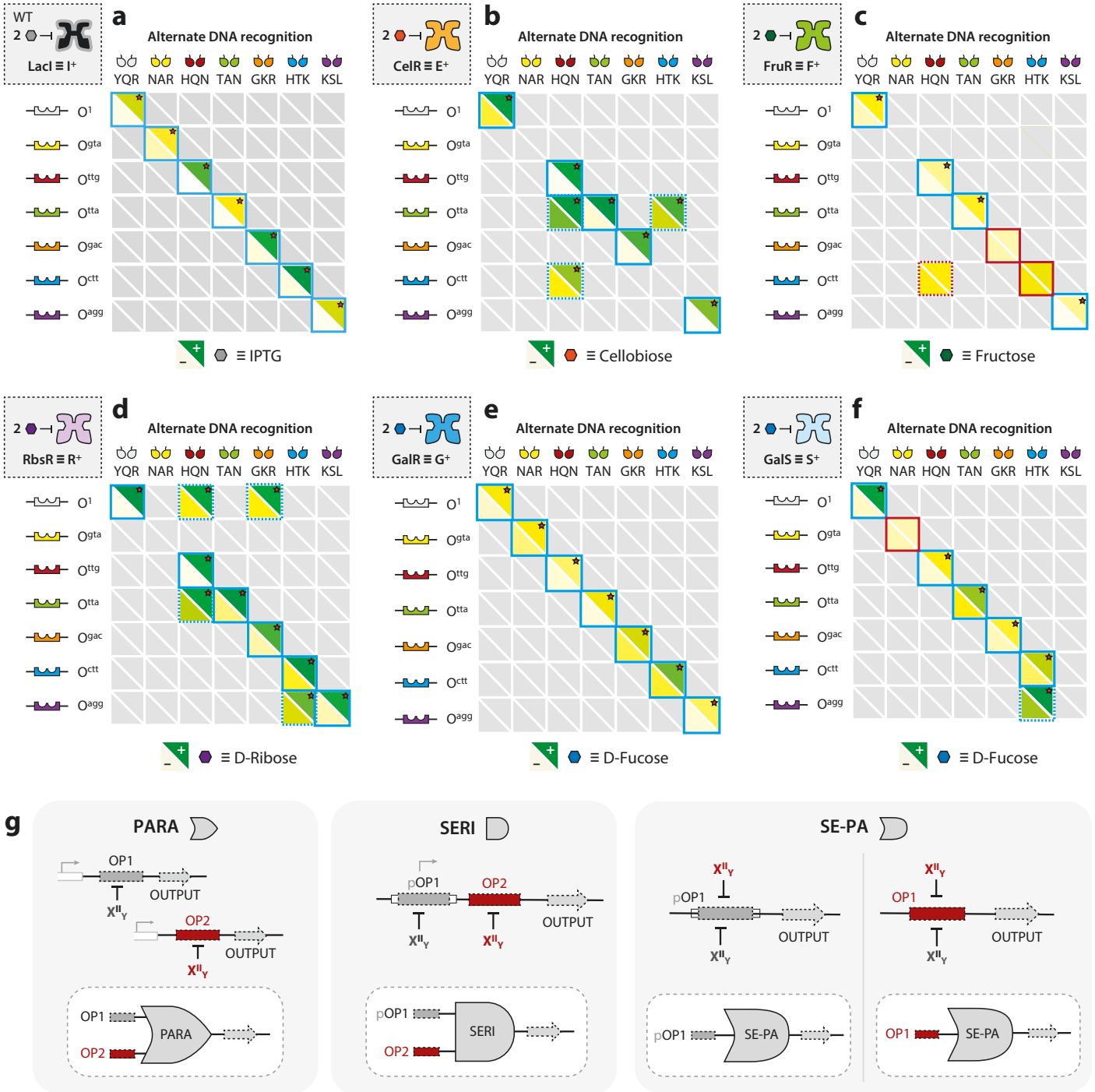


## 3.1. Engineering Alternate DNA Binding Recognition

Several studies have demonstrated that variable DBDs (i.e., allosteric functional surfaces that facilitate modulated interactions with DNA) can be paired with the LacI allosteric RCD unit to confer alternate DNA recognition (ADR) (13, 38, 55, 74) (see **Figure 1*b***). Notably, Milk et al. (38) developed a functional code (although not rational design rules) that can be used to engineer alternate DNA binding modules that are compatible with the native allosteric route (network) present in the wild-type LacI RCD. Alternate DNA binding engineering rules can be gleaned from complementary pairing of residues in the DBD, with a cognate operator DNA element. In Milk et al.'s study, over 8,000 putative alternate DNA-binding functional surface variants were generated, encompassing a fully randomized library of positions Y17, Q18, and R22 (YQR) (38). This library was tested against 64 putative operator variants, with the sequence 5′-A ATT **𝟄𝟄𝟄** AGC GCT **ΨΨΨ** AAT T-3′, where **𝟄** is any nucleotide, and **Ψ** is the complement necessary to achieve full (palindromic) symmetry. After extensive screening (monitoring the inducible regulation of GFP), 332 nonsynonymous transcription factor ($I^{+}_{XXX}$) | operator ($O^{𝟄𝟄𝟄}$) combinations were identified, although they were not explicitly tested for orthogonality. Orthogonality in this case is regarded as exclusive transcription factor binding to a single operator, relative to a finite (or restricted) population of DNA elements.

To demonstrate orthogonality, Rondon & Wilson (53) selected the wild-type LacI regulatory core and paired this domain with eight alternate DNA-binding functional surfaces from the principal design space generated by Milk et al. (38). When the alternate DBDs were combined with the naturally occurring native allosteric route (network) and ligand-binding surface, only six (out of eight) alternate DBDs interacted and functioned with cognate operator DNA. Namely, NAR, HQN, TAN, GKR, HTK, and KSL variants—where the three letter code corresponds to positions 17, 18, and 22 of the primary structure of the DBD—resulted in the repressor phenotype (**Figure 2*a***). Three DBDs (YQR, NAR, and RQR) resulted in noncognate interactions, and one DBD (AWR) failed to interact with any operator DNA. This observation highlights the importance of having rational design rules for intrinsically disordered regions (such as the DBD) to achieve (or not to achieve) specificity and action for a given allosteric transcription factor possessing the given topology.

Chen et al. (9) have also demonstrated that tuning the properties of the DNA elements proximal to the binding objective (operator) can result in fine control over the dynamic range. In their study, a library of promoters were constructed in which a spectrum of dynamic ranges were achieved for several transcription factors. The authors used a modular approach in which promoters were constructed from a moderate design space, and in which the $\sigma^{70}$ binding sites (−10 and −35 hexamers) were varied, in addition to the regions in proximity to these sites. The authors showed that, for a range of moderate $\sigma^{70}$ $K_{eq}$ values, a promoter can be made to exhibit low leakiness and




(Caption appears on following page)

**Figure 2** (*Figure appears on preceding page*)

Repression matrices and metrology for a class of transcription factors. (*a–f*) Repression matrices. Each matrix displays results for a given repressor regulatory core domain (RCD) (for a general repressor, $X^+$, where X represents a protein scaffold, as indicated) equipped with alternate DNA recognition (ADR) domains (across the tops of the matrices) tested against a collection of DNA operators (along the sides of the matrices). DNA binding domain (DBD) names correspond to residues at positions 17, 18, and 22 of the lactose repressor (LacI) DBD, while operator names correspond to nucleotides that differ between distinct operators (i.e., **999**). Cognate interactions are shown in the same color (along the diagonals of the matrices). Each box of the matrix shows the normalized expression level of green fluorescent protein (GFP) in the uninduced (−, *bottom left*) and induced (+, *upper right*) states, with darker green indicating higher expression. Stars indicate statistical significance between the two states. Gray boxes indicate an unresponsive phenotype (high protein expression in both states), likely due to inability to interact with DNA. Blue boxes indicate a repressor ($X^+$) phenotype, whereas red boxes indicate a super-repressor ($X^S$) phenotype. (*g*) Genetic architectures used to construct combinatorial logic. The parallel (PARA) architecture features two distinct operators, OP1 and OP2, upstream of two distinct OUTPUT genes, allowing independent expression of each. The series (SERI) architecture features two distinct operators upstream of a single OUTPUT gene, requiring action from both transcription factors to execute a function. The series-parallel (SE-PA) architecture features a single operator upstream of a single OUTPUT gene, combining aspects of SERI and PARA. In this case, regulatory proteins with common DBDs may simultaneously and independently regulate the operator. This may be executed in the core operator position (intercalated with the promoter, pOP1) (*left*) or in the proximal operator position (*right*), where the operator is between the promoter and OUTPUT gene (OP1). In all cases, each operator binds a distinct, orthogonal transcription factor. Aspects of this figure are adapted from References 21 and 52, both licensed under CC BY 4.0 (**https://creativecommons.org/licenses/by/4.0/**).



high induction (i.e., a large dynamic range). While tuning of $\sigma^{70}$ binding is not directly related to the allosteric properties of a given transcription factor, hexamer engineering should be considered as an additional criterion when engineering allosteric transcription factors used in genetic circuits.

## 3.2. Building Bespoke Transcription Factors with Alternate DNA and Ligand Binding

In contrast to our ability to engineer a DNA-binding functional surface that can pair with the LacI allosteric domain, engineering of the ligand-binding functional surface of LacI has met with limited success. Notably, Taylor et al. (64) used the Rosetta computer-aided protein design suite (in addition to laboratory evolution) to engineer the allosteric transcription factor LacI to respond to new ligands. Briefly, the binding-site residues (that constitute the ligand-binding functional surface) were redesigned to accept disparate small-molecule ligands (i.e., fucose, gentiobiose, lactitol, and sucralose). This study highlighted how altering inducer specificity in these proteins is difficult because substitutions that affect inducer binding may also disrupt allostery. Namely, while alternate ligand binding was achieved, these systems displayed a moderate response to induction. Moreover, these engineered systems displayed cross-reactivity toward the other ligands, including the native ligand IPTG, thus limiting the utility of such systems.

We posit that engineered ligand-binding functional surface residues typically result in moderate to weak function (if any), due in large part to each ligand-binding functional surface requiring a compatible allosteric network. Accordingly, given a ligand-binding functional surface with a compatible allosteric network, alternate DNA binding adaptation can be achieved with significantly greater success. This modular design strategy has recently been demonstrated by several groups via the pairing of disparate RCDs with alternate DNA binding modules to form functional allosteric transcription factors—particularly by way of the broader LacI topology (6, 52, 56). To understand the accomplishments discussed above, we should first appreciate that our understanding of the LacI structure–function relationship has been expanded to study and identify more than 1,000 homologous proteins. This collection of putative allosteric transcription factors are commonly referred to as the LacI/GalR family (35, 58, 62). Each LacI homolog has evolved a unique variation in ligand binding that is complementary to a unique solution to allosteric communication via variations in topology of the RCD, in addition to a DNA-binding functional surface that has

affinity for a specific DNA element. Allosteric transcriptional regulation via modulated DNA binding is the putative communal function of many (if not all) of the members of this protein family. Accordingly, a reasonable supposition that can be inferred from this family of proteins is that the design of an allosterically regulated transcription factor requires simultaneous and reciprocal consideration of all three modules—i.e., ligand binding, DNA binding, and complementary allosteric medium.

To illustrate this modular design (engineering) strategy, Rondon et al. (52) paired nonsynonymous RCDs composed of disparate ligand-binding functional surfaces and compatible allosteric topologies with a set of alternate DNA-binding functional surfaces (that demonstrated functionality with the native LacI regulatory core) to create a large collection of non-natural transcription factors (**Figure 2*b–f***). The design space was composed of 35 putative non-natural transcription factors via five regulatory cores and seven alternate DNA recognition units. Experimental analysis of each non-natural allosteric transcription factor revealed that 27 out of 35 of the putative repressors were functional—interacting with a cognate DNA operator—and were inducible. Six non-natural chimeras interacted with noncognate DNA operators. In this study, two allosteric regulatory cores (GalS and GalR) share a synonymous ligand-binding functional surface and bind the ligand D-fucose. GalR and GalS have primary structures that are 54% identical and can thus be regarded as two systems that utilize variable allosteric networks with a fixed ligand binding site (**Figure 2*e,f***). Accordingly, the differences in primary topology between the GalR and GalS scaffolds present an opportunity to evaluate how variation in the allosteric medium could potentially influence functional outcomes. The GalR and GalS repressors, when paired with the same set of alternate DNA-binding functional surfaces, have different functional outcomes. In this case study, when the functional surfaces (ligand binding and DNA binding) are fixed, but the composition of the allosteric medium varies, the dynamic range between synonymous sets was variable—specifically for GalR and GalS RCDs adapted with YQR, HQN, TAN, GKR, or HTK. GalR paired with NAR resulted in the expected repressor phenotype, whereas GalS paired with the same DBD resulted in the super-repressor phenotype. Similarly, GalR paired with KSL resulted in the expected repressor phenotype; however, when this alternate DBD was paired with the GalS system, the transcription factor became nonfunctional.

## 4. ENGINEERING ALTERNATE ALLOSTERIC COMMUNICATION: NOT LOGIC

Using the LacI system as a model of regulatory response, Poelwijk et al. (47) sought to demonstrate how cells can adapt to environmental variability. Namely, cells bearing the wild-type DNA operator, positioned to facilitate the regulation of the tandem production of a chloramphenicol resistance protein (cmR) and a counter selection marker (sacB, encoding levansucrase, which converts sucrose to levans and is toxic to bacteria) were used to identify antilac variants. *Escherichia coli* containing LacI-regulated cmR and sacB were subjected to environments for which they had varying degrees of fitness depending on the presence of the antibiotic chloramphenicol and a toxin (sucrose), with IPTG controlling expression (via a variant of LacI) of the corresponding selection markers. Optimal fitness in the environment would only be the result of a LacI mutant that exhibits an alternate allosteric response, where DNA binding occurs with ligand binding and no DNA binding occurs in the absence of ligand (i.e., antirepression) (47). Independently, Meyer et al. (36) and Richards et al. (51) conferred antirepression in the LacI scaffold via a two-part workflow described in detail below. This class of engineered antilacs represents a system of NOT gates (antithetical to the BUFFER gates presented in **Figure 2*a***) that can be used to construct combinatorial logic systems, described in Section 5.

## 4.1. Engineering Alternate Allostery: Conferring Antirepression

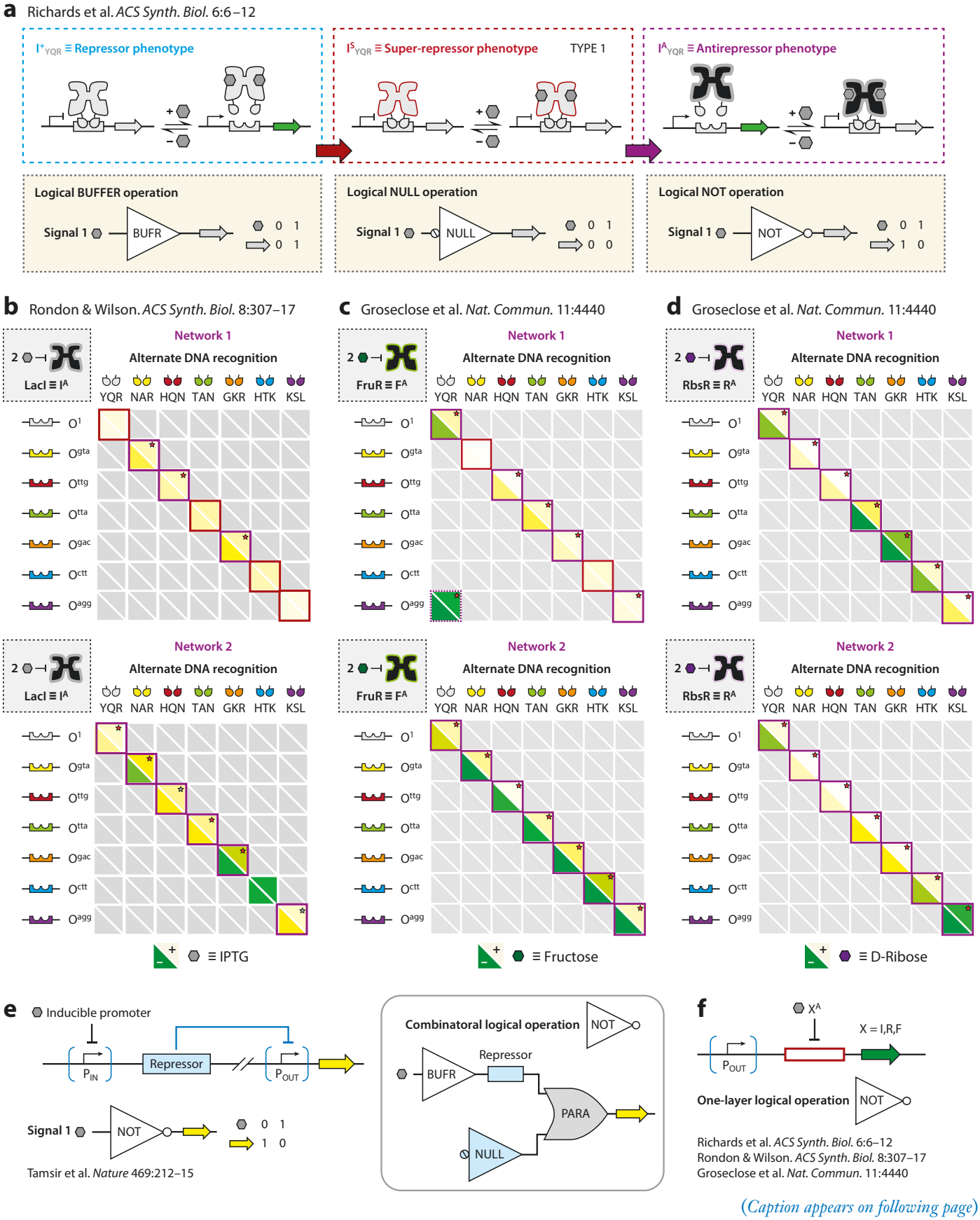
Richards et al. (51) hypothesized that alternate allosteric routes could be conferred in the LacI scaffold by (*a*) first blocking native allosteric communication via a super-repressor ($I^S$) point mutation, (*b*) followed by one or more rounds of laboratory evolution introducing compensatory (functional) mutations (**Figure 3*a***). To test this hypothesis, Richards et al. blocked allosteric communication in the LacI scaffold by introducing four separate super-repressor point mutations (K84A, V95A, V95F, D275F). Using one or more rounds of error-prone polymerase chain reaction, the investigators introduced additional compensatory mutations (complementary to the ligand-binding and DNA-binding functional surfaces) resulting in either (*a*) a conferred antirepressor phenotype ($I^A{}_{YQR}$) (**Figure 3*b***) or (*b*) a rescued repressive phenotype ($I^+{}_{YQR}$). The engineered systems were composed of 11 antirepressors ($I^A{}_{YQR}$) and three alternate repressors ($I^+{}_{YQR}$). All variants retained the initial $I^S$ block, and functional mutations occurred throughout the RCD; however, no mutations were introduced within the DBD. Mutations that conferred alternate allosteric communication had considerable overlap with positions that constitute the trajectories identified by the in silico targeted molecular dynamics study (18). However, at least two alternate solutions have no apparent correlation with putative targeted molecular dynamics trajectories. In other words, the experimental study illustrated that the LacI topology could support multiple (alternate) allosteric networks. Thus, allosteric communication (in the context of this scaffold) can be regarded as plastic, rather than as a fixed (immutable) path. In addition to conferring antirepression, each variant displayed a unique (dose responsive) transfer function and dynamic range (51).

The super-repressor phenotype has a truth table in which the OUTPUT is regarded as zero, regardless of the INPUT state (see **Figure 3*a***). Accordingly, the $I^S$ phenotype can be classified as a noncanonical NULL operation. The $I^S$ phenotype can be conferred via two mechanisms. The first class of $I^S$ variants (type 1) introduces a disruption (block) in allosteric communication. However, type-1 $I^S$ variants retain the ability to bind the ligand (see **Figure 3*a***). The second class of $I^S$ variants (type 2) mitigates allosteric communication by introducing point mutations that disable the ligand-binding functional surface. The ability of type-1 $I^S$ variants to retain ligand binding function (36, 70) facilitates the engineering of alternate allosteric routes in the form of antirepression, as described above, and thus constitutes a more granular design rule. From the vantage point of control systems engineering, the antirepressor phenotype is a single-INPUT single-OUTPUT open-loop system, represented as a NOT logical gate. Importantly, the NOT operations presented in this study have performance conditions that are on par with (but antithetical to) the BUFFER logical operations (**Figures 1*a*** and **3*a***) and can thus be used in collaboration to construct combinational logical functions, as discussed in Section 5.

## 4.2. Engineering Antirepressors with Alternate DNA Binding Function

To test whether DNA-binding functional surface orthogonality is transferable to alternate allosteric routes (i.e., antilac scaffolds), Rondon & Wilson (53) selected 9 out of 14 of the engineered LacI antirepressors from a previous study (51) and paired the engineered regulatory core domains with the six alternate DNA-binding functional surfaces that were shown to be orthogonal in the native repressor ($I^+$) scaffold (**Figures 2*a*** and **3*b***). This resulted in an allosteric design space of 54 putative antirepressors, of which (*a*) 46 functioned as cognate antirepressors; (*b*) 4 systems resulted in the super-repressor phenotype; and (*c*) 3 systems were unresponsive to any operator element, including the cognate operator. In **Figure 3*b***, two sample matrices exemplify how alternate allosteric networks (antirepressor, $I^A$) can influence functional outcomes.

In turn, Groseclose et al. (21) leveraged the workflows used to engineer systems of antilacs outlined in **Figure 3*a,b*** to construct two new classes of antirepressors. The two new systems of




(Caption appears on following page)

**Figure 3** (*Figure appears on preceding page*)

Construction of biotic Boolean NOT operations via antirepressors. (*a*) Evolution of an antirepressor. Beginning with a repressor, in this case, LacI ($I^{+}_{YQR}$), a mutation that blocks the native allosteric network may be introduced to generate a super-repressor ($I^{S}_{YQR}$). Type 1 indicates that the binding pocket is preserved, and only the allosteric network is disrupted. Compensatory mutations, e.g., through laboratory evolution, may be made to the super-repressor to reroute the allosteric network to an antirepressor ($I^{A}_{YQR}$). These phenotypes can also be classified as the operations BUFFER, NULL, and NOT, respectively. Putative mechanisms of the regulatory protein variants interacting with operators (with and without interaction with ligand) are shown in schematics. (*b–d*) Representative antirepression matrices. (*b*) Antirepression matrices for two anti-LacIs with two nonsynonymous allosteric routes [Network 1 (*top*) and Network 2 (*bottom*)]. (*c*) Antirepression matrices for two anti-FruRs, also with two nonsynonymous allosteric routes [Network 1 (*top*) and Network 2 (*bottom*)]. (*d*) Antirepression matrices for two anti-RbsRs, with two nonsynonymous allosteric routes [Network 1 (*top*) and Network 2 (*bottom*)]. Each box of the matrix shows the normalized expression level of green fluorescent protein (GFP) in the uninduced (−, *bottom left*) and induced (+, *upper right*) states, with darker green indicating higher expression. Stars indicate statistical significance between the two states. Gray boxes indicate an unresponsive phenotype (high protein expression in both states), likely due to inability to interact with DNA. Purple boxes indicate an antirepressor ($X^{A}$) phenotype, whereas red boxes indicate a super-repressor ($X^{S}$) phenotype. (*e*) NOT operation constructed by Tamsir et al. (63). An inducible promoter ($P_{IN}$), in this case implicitly regulated by LacI, controls the expression of a repressor, which regulates the expression of an output gene (*yellow*) via an inducible output promoter, $P_{OUT}$. If LacI is induced, and the repressor (functioning as a NULL gate) is not induced, then this results in a combinatorial NOT operation. (*f*) Next-generation NOT gate, reducing the layered NOT operation (shown in panel *e*) from a two-promoter, two–regulatory protein, two-layer system to a unit operation composed of one layer, one promoter, and one regulatory protein. Aspects of this figure are adapted from Reference 21, licensed under CC BY 4.0 (**https://creativecommons.org/licenses/by/4.0/**), and Reference 53, used with permission from the American Chemical Society.



engineered antirepressors were responsive to fructose-1,6-phosphate (anti-FruR), and D-ribose (anti-RbsR), respectively (see **Figure 3*c,d***). In the first stage of the engineering workflow, alternate allosteric communication was conferred in the FruR and RbsR scaffolds using a similar strategy to the one outlined in **Figure 3*a***. In turn, this collection of engineered regulatory cores were adapted with the same set of alternate DBDs used to construct the systems of transcription factors described in **Figures 2*a–f*** and **3*b***. As evidenced in the examples discussed above (i.e., the GalR and GalS systems described in **Figure 2*e,f*** and the antilac systems described in **Figure 3*b***), variations in the given allosteric networks conferred different dynamic ranges in the FruR (**Figure 3*c***) and RbsR (**Figure 3*d***) antirepressor systems.

Collectively, the data for the three antirepressor systems illustrate that (*a*) variation in the allosteric network alone can confer different dynamic ranges (i.e., absolute differences between ON and OFF states); (*b*) alternate allosteric communication in a given topology (with a fixed ligand-binding functional surface) can accommodate a variety of DNA-binding functional surfaces; and (*c*) all alternate allosteric networks are not necessarily compatible with a given DNA-binding functional surface, even if that functional surface has allosteric communication variants that are proximal in primary structure.

## 4.3. Next-Generation NOT Gates: Engineered Antirepressors

Notably, Tamsir et al. (63) devised an alternate scheme that achieved similar NOT logic (**Figure 3*e***). An inducible promoter regulates the expression of a NULL operation (i.e., super-repressor equivalent) that can suppress a cognate promoter via a second layer. Given that a variety of gene regulators can be used to regulate the input promoter, the authors were able to represent a variety of ligand-binding and DNA-binding functional surfaces—although this was achieved via a broad range of topologies and functional mechanisms. Nevertheless, the allosteric antirepressors developed by Rondon & Wilson (53) and Groseclose et al. (21) represent a significant technological advance over the state of the art (**Figure 3*e***), reducing the NOT operation from a two-promoter, two–regulatory protein, two-layer system to a unit operation composed of one layer, one promoter, and one regulatory protein (**Figure 3*f***). Accordingly, this advance over the

state of the art is important in that this collection of next-generation NOT logic operations will facilitate the development of larger and more complex biological programs, as these systems require fewer resources to operate.

## 5. BIOLOGICAL COMBINATIONAL LOGIC CIRCUITS

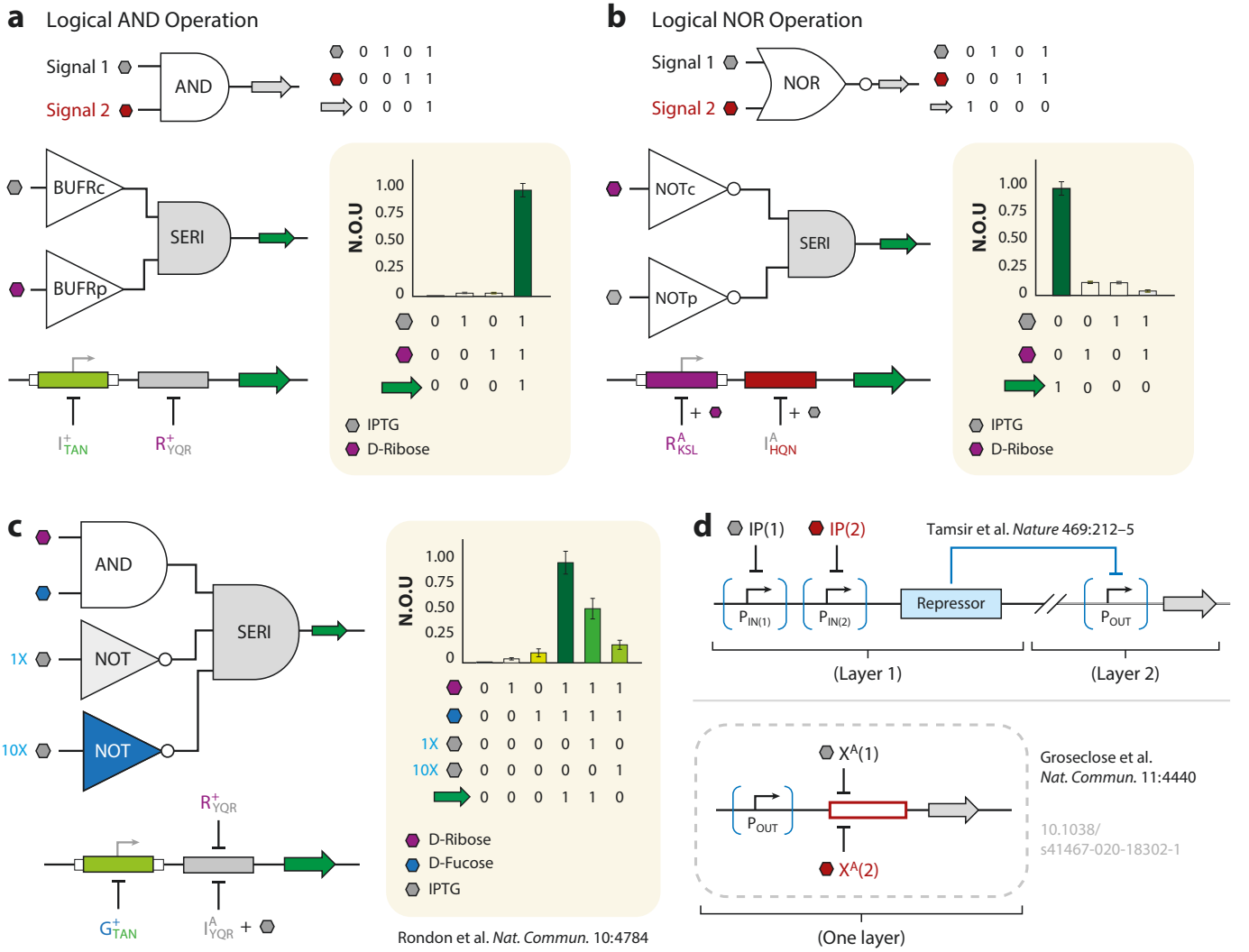
Combinational logic is a concept in which two or more fundamental logical operations can be combined by way of a defined set of rules to create complex (multiple-input) logical operations. Drawing on this electrical engineering concept for inspiration, similar combinational logic circuits can be constructed in biotic systems via the fundamental BUOs BUFFER and NOT described above. Achieving combinational logic in biology required the development of genetic architectures to direct BUFFER (repression) and NOT (antirepression) unit operations (see **Figure 2*g***). In this section, we demonstrate the construction of AND logic and NOR logic, in addition to an advanced (BANDPASS filter) logical operation that utilizes both repressor and antirepressor unit operations.

### 5.1. Building AND Gates from BUFFER Unit Operations

One of the first biological AND gates was constructed by Nielsen et al. (44) and composed of three regulatory proteins and five promoters. In addition, several research groups have constructed less complex two-signal AND gates by way of the combination of two BUFFER unit operations that are directed via the series-parallel architecture (6, 21, 52, 56) or the series genetic architecture (52) described in **Figure 2*g***. In this section, we focus on the latter iterations of AND gates, as these systems offer an advantage to biotic systems by reducing the metabolic burden of the host organism. Briefly, the series genetic architecture contains two DNA operator elements in tandem and can facilitate binding of two disparate repressor (BUFFER) transcription factors, illustrated in **Figure 4*a***. The basic operation of the given biological AND gate is summarized in the truth table, in which the gate can only be in the ON state (i.e., produce GFP) when both signals are present. In turn, biological AND gates constructed by way of a series-parallel genetic architecture only use one DNA operator and require that the two constituent repressors share a common DBD. The key difference between the two varieties of AND gates is that the series-based operation allows for independent tuning of each BUFFER operation, as the constituent repressors do not share a common DBD.

### 5.2. Building Next-Generation NOR Gates

Tamsir et al. (63) constructed one of the first biological NOR gates (see **Figure 4*d***). The NOR logical operation presented in **Figure 4*d*** builds on the NOT gate illustrated in **Figure 3*e***. Tamsir et al. introduced a second input promoter (objectively forming an OR Boolean logic operation) that regulates the production of a NULL output, which functions as an inverter and is directed to a second layer where GFP regulation occurs. In a recent study, Groseclose et al. (21) dramatically simplified the construction of the NOR biological gate using antirepressors, as illustrated in **Figure 4*d***. (Note that this example employs the series-parallel architecture.) To accomplish this, Groseclose et al. leveraged both the series (**Figure 4*b***) and series-parallel (**Figure 4*d***) architectures following a build strategy that was synonymous to the construction of multiple input AND gates—although the NOR combinational logic circuit used disparate NOT unit operations, opposed to BUFFER operations. In summary, the next-generation NOR gate illustrated in **Figure 4*d*** only requires one promoter and two regulatory proteins (i.e., antirepressors), which is two fewer promoters and one fewer regulatory protein relative to the state of the art.




## Figure 4

Achieving combinational logic using systems of engineered allosteric transcription factors. (*a*) A combinatorial AND operation constructed by Rondon et al. (52). Two repressors (BUFFERs), $I^{+}_{TAN}$ and $R^{+}_{YQR}$, regulate the green fluorescent protein (GFP) OUTPUT through operators $O^{tta}$ (core) and $O^{sym}$ (proximal). Only when both ligands, D-ribose and isopropyl-β-D-thiogalactoside (IPTG), are present do both proteins dissociate from DNA, relieving repression and allowing gene expression. (*b*) A combinatorial NOR operation constructed by Groseclose et al. (21). Two antirepressors (NOTs), $R^{A}_{KSL}$ and $I^{A}_{HQN}$, regulate the GFP OUTPUT through operators $O^{agg}$ (core) and $O^{ttg}$ (proximal). Gene expression is abolished to a relative 0 when either ligand is, or both ligands are, present, which causes the respective antirepressor(s) to associate with operator DNA. (*c*) A biological BANDPASS filter constructed from an AND gate with a NOT gate and a second, conditional NOT gate. $G^{+}_{TAN}$ and $R^{+}_{YQR}$ regulate expression of GFP via $O^{tta}$ and $O^{sym}$ operators, resulting in AND function in the presence of fucose and ribose. However, in the presence of 1X IPTG, gene expression is diminished, as $I^{A}_{YQR}$ now is able to anti-induce the $O^{sym}$ operator. In the presence of 10X IPTG, $I^{A}_{YQR}$ anti-induces the operator, and $G^{+}_{YQR}$ is also competitively inhibited by IPTG, further decreasing gene expression. In all instances, plots display normalized output units (NOUs; fluorescence normalized to maximum) for each inducer state. (*d*) Generations of combinatorial NOR operations. (*Top*) The NOR operation constructed by Tamsir et al. (63). Two inducible promoters [IP(1) and IP(2)] are located upstream of a repressor gene, each implicitly regulated by a distinct repressor (BUFFER operation). In the absence of either inducer, both repressors prevent the transcription of the repressor gene—the final output from $P_{OUT}$ is a relative 1. However, in the presence of either (or both) ligands, implicit repressors dissociate from IP(1) and/or IP(2), restoring promoter activity and causing the third repressor (*blue box*) to be transcribed. This repressor functions as an inverter (when it is itself uninduced), diminishing output from $P_{OUT}$. (*Bottom*) The next generation of NOR transcriptional logic, which reduces the operation to a single promoter with two regulatory proteins in one layer. In this case, the presence of either (or both) ligands causes association of the $X^{A}$ antirepressor(s) with the operator (shown in series-parallel in the proximal position), diminishing gene OUTPUT. Aspects of this figure are adapted from References 21 and 52, both licensed under CC BY 4.0 (**https://creativecommons.org/licenses/by/4.0/**).

### 5.3. Understanding Allosteric Communication Through Metrology

There is a dearth of metrology throughout synthetic biology and related fields that employ allosteric systems. This lack of metrology for biological systems has resulted in variable performance in a given genetic part, which reduces our ability to predictably conduct bottom-up design of complex systems with high fidelity. One solution to this problem is to establish a metrology in the context of fundamental BUOs. Toward this end, Nielsen et al. (44) reported the most extensive metrology used to measure fundamental genetic parts. This metrology was inspired by work conducted by Kelly et al. (27), who defined a relative promoter unit (RPU) to report promoter characterization data in compatible units and developed a measurement kit so that researchers might more easily adopt the RPU as a standard unit for reporting promoter activity. Following the workflows developed by Kelly et al., Nielsen et al. converted characterized sensor and gate fluorescence data into the RPU to represent the standard unit of measurement in Cello, enabling predictive circuit design. Cello builds circuits by connecting transcriptional gates, whose common signal carrier is RNA polymerase flux on DNA.

In two additional studies, Rondon et al. (52) and Groseclose et al. (21) established a separate metrology for SISO BUO repressors and antirepressors, respectively. In brief, the metrology consists of three parts: (*a*) defining the conditional units of measurement for a given allosteric transcription factor, (*b*) reproducible realization of units of measurement at steady state, and (*c*) development of a traceability score via the comparison of performance metrics for a given transcription factor to a reference system. The traceability score is given as a set of relative numerical values (i.e., induction units and repression units), linking measurements made for a given transcription factor DNA operator set (non-natural or natural) to the reference standard $I^{+}_{YQR}$ | $O^{1}$. This metrology enables the comparison of the performance metrics of a given transcription factor between laboratories, setting the stage for predicting the performance of complex transcriptional programs prior to their construction. In addition, an established metrology of this variety will facilitate engineering of allosteric communication and, eventually, the de novo design of such systems. Namely, the combination of (*a*) allosteric mapping with (*b*) the ability to engineer alternate allosteric communication, (*c*) objective measurements of performance, and (*d*) transferability to other systems that employ communal allosteric functions will enable the design of bespoke allosteric communication.

## 6. CONCLUDING REMARKS

The development of foundational biological unit operations composed of allosteric transcription factors (i.e., BUFFER and NOT gates) will facilitate the development of a variety of multiple-INPUT single-OUTPUT systems (see **Figure 4*c***). However, the prediction of the performance of a given open-loop multiple-INPUT system will require a deep understanding of single-INPUT unit operations—which will require a deeper understanding of allosteric communication. What is clear from the case studies presented in this review is that the full a priori design of a functional allosteric protein will require the simultaneous design of both functional surfaces, along with the corresponding allosteric topology. Using systematic workflows could potentially simplify the allosteric design problem—i.e., via the development of hierarchical design rules. In addition to facilitating the bottom-up development of multiple-INPUT combinatorial logical systems, viewing BUOs as single-INPUT control systems will enable the intuitive and methodical development of closed-loop biological controllers (3). Archetypical control systems facilitated via the proposed edifice will enable biological engineers to design, build, and test processes that possess the ability to maintain set points, reject disturbances, and be implemented as multiple-input multiple-output controllers.

## DISCLOSURE STATEMENT

The authors are not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

## ACKNOWLEDGMENTS

This work was supported by National Science Foundation grants CBET 1934836, MCB 1921061, CBET 1844289, CBET 1804639, and MCB 1747439 to C.J.W.

## LITERATURE CITED

1. Ahmed MH, Ghatge MS, Safo MK. 2020. Hemoglobin: structure, function and allostery. *Subcell Biochem*. 94:345–82
2. Alon U. 2007. Network motifs: theory and experimental approaches. *Nat. Rev. Genet.* 8:450–61
3. Aoki SK, Lillacci G, Gupta A, Baumschlager A, Schweingruber D, Khammash M. 2019. A universal biomolecular integral feedback controller for robust perfect adaptation. *Nature* 570:533–37
4. Bell CE, Lewis M. 2000. A closer view of the conformation of the Lac repressor bound to operator. *Nat. Struct. Biol.* 7:209–14
5. Cameron DE, Bashor CJ, Collins JJ. 2014. A brief history of synthetic biology. *Nat. Rev. Microbiol.* 12:381–90
6. Chan CTY, Lee JW, Cameron DE, Bashor CJ, Collins JJ. 2016. "Deadman" and "Passcode" microbial kill switches for bacterial containment. *Nat. Chem. Biol.* 12:82–86
7. Chen J, Matthews KS. 1994. Subunit dissociation affects DNA binding in a dimeric *lac* repressor produced by C-terminal deletion. *Biochemistry* 33:8728–35
8. Chen J, Surendran R, Lee JC, Matthews KS. 1994. Construction of a dimeric repressor: dissection of subunit interfaces in *Lac* repressor. *Biochemistry* 33:1234–41
9. Chen Y, Ho JML, Shis DL, Gupta C, Long J, et al. 2018. Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors. *Nat. Commun.* 9:64
10. Choi KR, Jang WD, Yang D, Cho JS, Park D, Lee SY. 2019. Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. *Trends Biotechnol.* 37:817–37
11. Courtney TM, Deiters A. 2019. Enzyme allostery: now controllable by light. *Cell Chem. Biol.* 26:1481–83
12. Coyote-Maestas W, He Y, Myers CL, Schmidt D. 2019. Domain insertion permissibility-guided engineering of allostery in ion channels. *Nat. Commun.* 10:290
13. Daber R, Lewis M. 2009. A novel molecular switch. *J. Mol. Biol.* 391:661–70
14. Daber R, Sharp K, Lewis M. 2009. One is not enough. *J. Mol. Biol.* 392:1133–44
15. Daber R, Stayrook S, Rosenberg A, Lewis M. 2007. Structural analysis of lac repressor bound to allosteric effectors. *J. Mol. Biol.* 370:609–19
16. Dahiyat BI, Mayo SL. 1997. De novo protein design: fully automated sequence selection. *Science* 278:82–87
17. Elowitz MB, Leibler S. 2000. A synthetic oscillatory network of transcriptional regulators. *Nature* 403:335–38
18. Flynn TC, Swint-Kruse L, Kong Y, Booth C, Matthews KS, Ma J. 2003. Allosteric transition pathways in the lactose repressor protein core domains: asymmetric motions in a homodimer. *Protein Sci.* 12:2523–41
19. Friedman AM, Fischmann TO, Steitz TA. 1995. Crystal structure of lac repressor core tetramer and its implications for DNA looping. *Science* 268:1721–27
20. Gordley RM, Bugaj LJ, Lim WA. 2016. Modular engineering of cellular signaling proteins and networks. *Curr. Opin. Struct. Biol.* 39:106–14
21. Groseclose TM, Rondon RE, Herde ZD, Aldrete CA, Wilson CJ. 2020. Engineered systems of inducible anti-repressors for the next generation of biological programming. *Nat. Commun.* 11:4440
22. Guet CC, Elowitz MB, Hsing W, Leibler S. 2002. Combinatorial synthesis of genetic networks. *Science* 296:1466–70

23. Hasty J, Dolnik M, Rottschäfer V, Collins JJ. 2002. Synthetic gene network for entraining and amplifying cellular oscillations. *Phys. Rev. Lett.* 88:148101
24. Hooshangi S, Thiberge S, Weiss R. 2005. Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. *PNAS* 102:3581–86
25. Howell SC, Inampudi KK, Bean DP, Wilson CJ. 2014. Understanding thermal adaptation of enzymes through the multistate rational design and stability prediction of 100 adenylate kinases. *Structure* 22:218–29
26. Huang PS, Boyken SE, Baker D. 2016. The coming of age of *de novo* protein design. *Nature* 537:320–27
27. Kelly JR, Rubin AJ, Davis JH, Ajo-Franklin CM, Cumbers J, et al. 2009. Measuring the activity of BioBrick promoters using an in vivo reference standard. *J. Biol. Eng.* 3:4
28. Khalil AS, Collins JJ. 2010. Synthetic biology: applications come of age. *Nat. Rev. Genet.* 11:367–79
29. Lee J, Natarajan M, Nashine VC, Socolich M, Vo T, et al. 2008. Surface sites for engineering allosteric control in proteins. *Science* 322:438–42
30. Lewis M, Chang G, Horton NC, Kercher MA, Pace HC, et al. 1996. Crystal structure of the lactose operon repressor and its complexes with DNA and inducer. *Science* 271:1247–54
31. Lim WA. 2010. Designing customized cell signalling circuits. *Nat. Rev. Mol. Cell Biol.* 11:393–403
32. Lim WA, June CH. 2017. The principles of engineering immune cells to treat cancer. *Cell* 168:724–40
33. Lockless SW, Ranganathan R. 1999. Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* 286:295–99
34. Ma CW, Xiu ZL, Zeng AP. 2015. Exploring signal transduction in heteromultimeric protein based on energy dissipation model. *J. Biomol. Struct. Dyn.* 33:134–46
35. Meinhardt S, Manley MW Jr., Becker NA, Hessman JA, Maher LJ 3rd, Swint-Kruse L. 2012. Novel insights from hybrid LacI/GalR proteins: family-wide functional attributes and biologically significant variation in transcription repression. *Nucleic Acids Res.* 40:11139–54
36. Meyer S, Ramot R, Kishore Inampudi K, Luo B, Lin C, et al. 2013. Engineering alternate cooperative-communications in the lactose repressor protein scaffold. *Protein Eng. Des. Sel.* 26:433–43
37. Michalodimitrakis K, Isalan M. 2009. Engineering prokaryotic gene circuits. *FEMS Microbiol. Rev.* 33:27–37
38. Milk L, Daber R, Lewis M. 2010. Functional rules for lac repressor–operator associations and implications for protein–DNA interactions. *Protein Sci.* 19:1162–72
39. Miller JP, Melicher MS, Schepartz A. 2014. Positive allostery in metal ion binding by a cooperatively folded β-peptide bundle. *J. Am. Chem. Soc.* 136:14726–29
40. Monod J, Changeux JP, Jacob F. 1963. Allosteric proteins and cellular control systems. *J. Mol. Biol.* 6:306–29
41. Monod J, Wyman J, Changeux JP. 1965. On the nature of allosteric transitions: a plausible model. *J. Mol. Biol.* 12:88–118
42. Motlagh HN, Wrabl JO, Li J, Hilser VJ. 2014. The ensemble nature of allostery. *Nature* 508:331–39
43. Mukhopadhyay A, Redding AM, Rutherford BJ, Keasling JD. 2008. Importance of systems biology in engineering microbes for biofuel production. *Curr. Opin. Biotechnol.* 19:228–34
44. Nielsen AAK, Der BS, Shin J, Vaidyanathan P, Paralanov V, et al. 2016. Genetic circuit design automation. *Science* 352:aac7341
45. Nielsen J, Keasling JD. 2016. Engineering cellular metabolism. *Cell* 164:1185–97
46. Nussinov R, Tsai CJ, Xin F, Radivojac P. 2012. Allosteric post-translational modification codes. *Trends Biochem. Sci.* 37:447–55
47. Poelwijk FJ, de Vos MGJ, Tans SJ. 2011. Tradeoffs and optimality in the evolution of gene regulation. *Cell* 146:462–70
48. Popovych N, Sun S, Ebright RH, Kalodimos CG. 2006. Dynamically driven protein allostery. *Nat. Struct. Mol. Biol.* 13:831–38
49. Purnick PE, Weiss R. 2009. The second wave of synthetic biology: from modules to systems. *Nat. Rev. Mol. Cell Biol.* 10:410–22
50. Raman AS, White KI, Ranganathan R. 2016. Origins of allostery and evolvability in proteins: a case study. *Cell* 166:468–80

51. Richards DH, Meyer S, Wilson CJ. 2017. Fourteen ways to reroute cooperative communications in the lactose repressor: engineering regulatory proteins with alternate repressive functions. *ACS Synth. Biol.* 6:6–12
52. Rondon RE, Groseclose TM, Short AE, Wilson CJ. 2019. Transcriptional programming using engineered systems of transcription factors and genetic architectures. *Nat. Commun.* 10:4784
53. Rondon RE, Wilson CJ. 2019. Engineering a new class of anti-LacI transcription factors with alternate DNA recognition. *ACS Synth. Biol.* 8:307–17
54. Saibil HR, Fenton WA, DK Clare, Horwich AL. 2013. Structure and allostery of the chaperonin GroEL. *J. Mol. Biol.* 425:1476–87
55. Sartorius J, Lehming N, Kisters B, von Wilcken-Bergmann B, Müller-Hill B. 1989. *lac* repressor mutants with double or triple exchanges in the recognition helix bind specifically to *lac* operator variants with multiple exchanges. *EMBO J.* 8:1265–70
56. Shis DL, Hussain F, Meinhardt S, Swint-Kruse L, Bennett MR. 2014. Modular, multi-input transcriptional logic gating with orthogonal LacI/GalR family chimeras. *ACS Synth. Biol.* 3:645–51
57. Siuti P, Yazbek J, Lu TK. 2013. Synthetic circuits integrating logic and memory in living cells. *Nat. Biotechnol.* 31:448–52
58. Sousa FL, Parente DJ, Shis DL, Hessman JA, Chazelle A, et al. 2016. AlloRep: a repository of sequence, structural and mutagenesis data for the LacI/GalR transcription regulators. *J. Mol. Biol.* 428:671–78
59. Stock G, Hamm P. 2018. A non-equilibrium approach to allosteric communication. *Phil. Trans. R. Soc. Lond. B* 373:0187
60. Suckow J, Markiewicz P, Kleina LG, Miller J, Kisters-Woike B, Müller-Hill B. 1996. Genetic studies of the Lac repressor. XV: 4000 single amino acid substitutions and analysis of the resulting phenotypes on the basis of the protein structure. *J. Mol. Biol.* 261:509–23
61. Süel GM, Lockless SW, Wall MA, Ranganathan R. 2003. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat. Struct. Biol.* 10:59–69
62. Swint-Kruse L, Matthews KS. 2009. Allostery in the LacI/GalR family: variations on a theme. *Curr. Opin. Microbiol.* 12:129–37
63. Tamsir A, Tabor JJ, Voigt CA. 2011. Robust multicellular computing using genetically encoded NOR gates and chemical "wires." *Nature* 469:212–15
64. Taylor ND, Garruss AS, Moretti R, Chan S, Arbing MA, et al. 2016. Engineering an allosteric transcription factor to respond to new ligands. *Nat. Methods* 13:177–83
65. Tobin PH, Richards DH, Callender RA, Wilson CJ. 2014. Protein engineering: a new frontier for biological therapeutics. *Curr. Drug Metab.* 15:743–56
66. Voigt CA. 2006. Genetic parts to program bacteria. *Curr. Opin. Biotechnol.* 17:548–57
67. Wehrs M, Tanjore D, Eng T, Lievense J, Pray TR, Mukhopadhyay A. 2019. Engineering robust production microbes for large-scale cultivation. *Trends Microbiol.* 27:524–37
68. Wilson CJ, Das P, Clementi C, Matthews KS, Wittung-Stafshede P. 2005. The experimental folding landscape of monomeric lactose repressor, a large two-domain protein, involves two kinetic intermediates. *PNAS* 102:14563–68
69. Wilson CJ, Zhan H, Swint-Kruse L, Matthews KS. 2007. The lactose repressor system: paradigms for regulation, allosteric behavior and protein folding. *Cell Mol. Life Sci.* 64:3–16
70. Wilson CJ, Zhan H, Swint-Kruse L, Matthews KS. 2007. Ligand interactions with lactose repressor protein and the repressor–operator complex: the effects of ionization and oligomerization on binding. *Biophys. Chem.* 126:94–105
71. Wodak SJ, Paci E, Dokholyan NV, Berezovsky IN, Horovitz A, et al. 2019. Allostery in its many disguises: from theory to applications. *Structure* 27:566–78
72. Woolston BM, Edgar S, Stephanopoulos G. 2013. Metabolic engineering: past and future. *Annu. Rev. Chem. Biomol. Eng.* 4:259–88
73. Wright PE, Dyson HJ. 2015. Intrinsically disordered proteins in cellular signalling and regulation. *Nat. Rev. Mol. Cell Biol.* 16:18–29
74. Zhan J, Ding B, Ma R, Ma X, Su X, et al. 2010. Develop reusable and combinable designs for transcriptional logic gates. *Mol. Syst. Biol.* 6:388

Annual Review of Biophysics
Volume 50, 2021

# Contents

**Indexes**



**Errata**

An online log of corrections to *Annual Review of Biophysics articles* may be found at http://www.annualreviews.org/errata/biophys