Genetic and Computational Tools

for Engineering Complex Prokaryotic Systems

Diego Alba Burbano

A dissertation
submitted in partial fulfillment of the
requirements for the degree of

Doctor of Philosophy

University of Washington
2025

Reading Committee:
James M. Carothers, Chair
Neda Bagheri
Jorge Marchand

Program Authorized to Offer Degree
Chemical Engineering

University of Washington

Abstract

Genetic and Computational Tools for Engineering Complex Prokaryotic Systems

Diego Alba Burbano

Chair of the Supervisory Committee:

James M. Carothers

Department of Chemical Engineering

Prokaryotic systems are promising platforms for bioproduction, but the complexity of their gene regulatory and metabolic networks presents significant challenges for engineering efficient biosynthetic pathways. While CRISPR-Cas technologies have revolutionized genome editing in eukaryotes, the development of programmable, multi-layer transcriptional control in bacteria is still emerging. This work outlines our efforts to expand the bacterial CRISPR toolkit and develop predictive models that enable dynamic and rational design of prokaryotic systems. We begin with a review of the diverse CRISPR-Cas systems available for gene editing and regulation in prokaryotes, emphasizing recent innovations that offer precise and flexible control over gene expression. Building on this foundation, we engineered interconnected CRISPRa/i circuits capable of executing dynamic genetic programs in both cell-free and bacterial systems, showing that these circuits can be layered to achieve temporal regulation and logic-based behaviors. To increase circuit complexity, we designed synthetic promoters that can be selectively activated by CRISPRa and are tunable by various inputs, including light, peptides, and small molecules—enabling the creation of large, multi-input transcriptional networks. Finally, we developed automated workflows for constructing large-scale kinetic models using proteomic and metabolomic data, which allow for prediction of system-wide metabolic responses and the identification of promising biosynthetic pathways. Together, these advances provide a versatile framework for programming sophisticated, multi-gene functions in prokaryotes, accelerating the development of next-generation microbial platforms for bioproduction.

Table of Contents

Reading Guide

Chapter 1 is comprised of our review published in *Annual Review of Chemical and Biomolecular Engineering*, "CRISPR Tools for Engineering Prokaryotic Systems: Recent Advances and New Applications", and serves as an introduction to different CRISPR-Cas systems and their applications in prokaryotic genetic engineering, both for gene editing as well as gene regulation.

Chapter 2 is comprised of our manuscript published in *Cell Systems*, "Multi-layer CRISPRa/i Circuits for Dynamic Genetic Programs in Cell-Free and Bacterial Systems", which demonstrates how CRISPR transcriptional control tools can be interconnected to assemble control circuits capable of carrying out dynamic genetic programs in both.

Chapter 3 includes our manuscript published in *Proceedings of the National Academy of Sciences, "*Engineering Activatable Promoters for Scalable and Multi-Input CRISPRa/i Circuits", which describes how prokaryotic promoters can be engineered to enable construction of larger CRISPRa circuits. We also show that CRISPRa/i circuits can be interfaced with a wide variety of input signals, such as light, peptides, and small molecules.

Chapter 4 presents our manuscript, "Automated Assembly of Large Kinetic Metabolic Models of Prokaryotic Systems", which describes workflows for building and parameterizing kinetic models using proteomics and metabolomics data. We introduce computational tools for retrieving enzyme kinetic parameters, parallel model fitting, and demonstrate how the resulting models can predict system-wide effects of metabolic perturbations and support biosynthetic pathway prospecting.

Chapter 1:

CRISPR Tools for Engineering Prokaryotic Systems:

Recent Advances and New Applications

Diego Alba Burbano[+,1,2], Cholpisit Kiattisewee[+,1,2], Ava V. Karanjia[1,2], Ryan A.L. Cardiff[1], Ian D. Faulkner[1,2], Widianti Sugianto[1,2], James M. Carothers*[,1,2]

1: Molecular Engineering & Sciences Institute
and Center for Synthetic Biology
University of Washington
Seattle, WA 98195
United States

2: Department of Chemical Engineering
University of Washington
Seattle, WA 98195
United States

+: These authors contributed equally

*: Corresponding author
jcaroth@uw.edu
206-221-4902

Published on *Annu. Rev. Chem. Biomol. Eng* on April 10th, 2024.
DOI: /10.1146/annurev-chembioeng-100522-114706

Abstract

In the past decades, the broad selection of CRISPR-Cas systems has revolutionized biotechnology by enabling multimodal genetic manipulation in diverse organisms. Rooted in a molecular engineering perspective, in this review we recapitulate the different CRISPR components and how they can be designed for specific genetic engineering applications. We first introduce the repertoire of Cas proteins and tethered effectors used for programing new biological functions through gene editing and gene regulation. We review current guide RNA design strategies and computational tools, and how CRISPR-based genetic circuits can be constructed through the regulated expression of gRNAs. Then, we present recent advances in CRISPR-based biosensing, bioproduction, and biotherapeutics across *in vitro* and *in vivo* prokaryotic systems. Finally, we discuss forthcoming applications in prokaryotic CRISPR technology that will transform synthetic biology principles in the near future.

Keywords: CRISPR, prokaryote, genetic circuits, biosensing, bioproduction, biotherapeutic

Introduction

Synthetic biologists seek to understand and harness biological systems to engineer solutions in a wide variety of fields, ranging from chemical bioproduction to medicine, agriculture, and energy (1, 2). In the past decade, CRISPR-Cas tools have revolutionized biological research by making designing and multiplexing complex genetic functions simpler compared to previous molecular tools. This rapid ability to modify genetic material is streamlining reverse engineering of genotype-phenotype relationships and genetically encoding new desired functions. To expand beyond these applications, molecular tools that facilitate novel genetic manipulation are in great demand (3).

Sequence-specific DNA-binding domains combined with different effector complexes, including nucleases, transcriptional regulators, and base editors, have been successful in many genetic engineering applications (4–6). Protein-based tools consisting of DNA-binding domains derived from zinc-finger nucleases (ZFNs) or transcription activator-like effector proteins (TALEs) offer robust and very precise behaviors (7). However, tuning their effect and designing them to target multiple genes is relatively difficult. In contrast, CRISPR tools facilitate targeting of multiple sites simultaneously since targets are encoded through simple Watson-Crick interactions in short guide RNA (gRNA) sequences. In terms of designability, the discovery of new CRISPR systems, paired with a wide variety of tethered effector domains, have greatly expanded the palette of genetic engineering tools (8–11).

Genetic engineering can be subdivided into 1) gene editing to change the underlying genetic information and 2) gene regulation to change how the information is processed. To date, a wide range CRISPR-Cas techniques have been developed for

gene editing, spanning from DNA double strand break by Cas nucleases to large DNA integrations via CRISPR-assisted transposases (12–15). Base editing is another promising gene editing technique that introduces point mutations into DNA or RNA by using nickase or catalytically-dead Cas proteins (nCas or dCas, respectively) to localize nucleobase editor enzymes to precise target sites (16, 17). Emerging gene editing effectors in combination with improving knowledge in Cas protein engineering has enabled rapid developments of novel CRISPR editing systems (18). Many CRISPR-based tools are now available for up- and down-regulation of gene expression by fusing dCas proteins to transcriptional or translational regulation effector domains. dCas9 was rapidly repurposed for programmable transcriptional repression, termed CRISPR interference (CRISPRi), by physically blocking the transcription initiation or elongation processes (13, 19). Other dCas proteins have also been implemented in the same fashion, including repertoire of dCas9s from different species (20) and dCas12 (21, 22). The transcriptional activation counterpart, termed CRISPR activation (CRISPRa), was promptly developed in eukaryotes based on transcription factors widely applicable in any eukaryotic system, e.g. VP64 (23, 24). In contrast, prokaryotic CRISPRa lagged behind due to limited universal transcriptional activators and stringent design rules for implementation (25, 26). Nevertheless, these genetic engineering tools are now broadly applied in different bacteria (27).

The versatility of CRISPR gene regulation and gene editing tools has fueled many applications across diverse bacteria as well as cell-free systems (Figure 1). In bacteria, CRISPR tools are being used for rapid strain engineering in model organisms such as *E. coli*, mainly for bioproduction applications (28, 29). More recently, CRISPR tools have

also been implemented in non-model microbes such as *C. necator*, *R. sphaeroides*, and *C. autoethanogenum* due to their unique metabolisms and wide range of substrate utilization, including CO_2 (30–33). Beyond metabolic engineering applications, biotherapeutic bacteria harboring CRISPR tools are now capable of recording changes in their environment and deploying novel antimicrobials (34–38). In cell-free systems, purified Cas enzymes and gRNAs are widely used as biosensors for *in vitro* diagnostics of nucleic acids, small molecules, proteins, and even bacteria (39, 40). Moreover, cell-free expression systems (CFES) that also contain the molecular machinery required for gene expression, such as PURE or lysate-based TXTL (41–43), have accelerated the characterization of new CRISPR systems and the development of CRISPR genetic circuits (44–47). In turn, these advancements are paving the way for cell-free bioproduction platforms and biosensors that can multiplex input signals and launch appropriate CRISPR programs in response (42, 48–51).

This article reviews CRISPR-based genetic engineering tools developed to date and their cutting-edge applications across diverse bacteria and cell-free systems. For eukaryotic counterparts, several reviews were recently published (18, 52–54). We first showcase the designability and modularity of CRISPR systems by covering the constituent components, namely the different Cas proteins, tethered effectors, and gRNAs, and how they can be designed for specific genetic engineering applications (Figure 1). We then examine how CRISPR-tools can be integrated into genetic circuits, granting a new layer of programmability and a framework for genetic information processing. We describe how these different CRISPR systems are being used for engineering complex biosensing, bioproduction, and biotherapeutics platforms. Lastly, we

discuss the current challenges that CRISPR technologies encounter and explore the next-generation of CRISPR applications in bacteria.

Figure 1: CRISPR Tools for Genetic Engineering Prokaryotic Systems

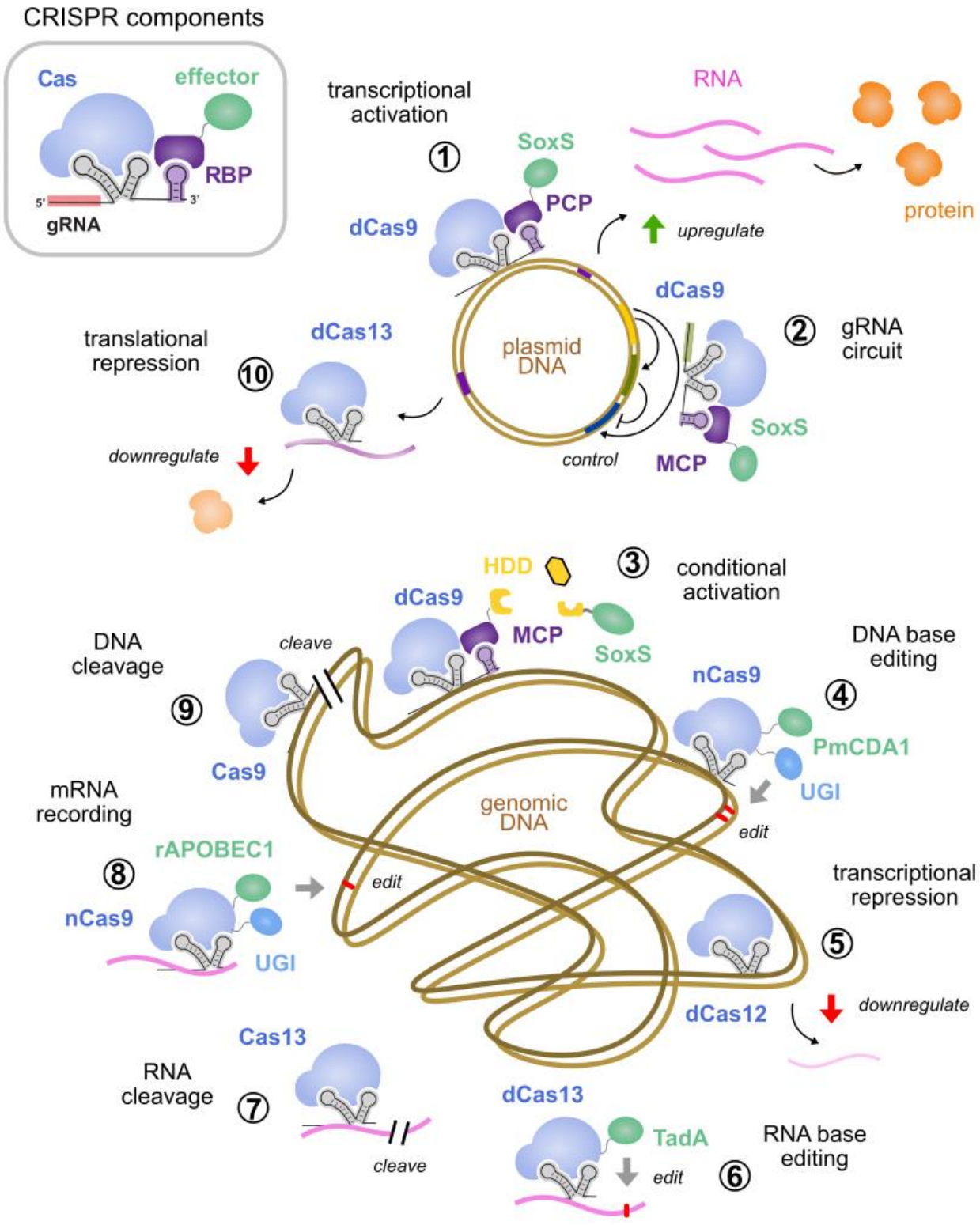


Figure 1: Schematic of different CRISPR tools and their mode of action. The top left panel depicts the main components of a CRISPR tool: Cas protein, guide RNA (gRNA), effector, and RNA binding protein (RBP). For simplicity, different Cas proteins are depicted the same and labeled accordingly. Plasmid and genomic DNA could be used interchangeably. (1) A dCas9-based transcriptional activation tool that recruits the SoxS activator through the RBP PP7 coat protein (PCP) and once localized to a gene of interest and upregulates RNA transcription. (2) A CRISPR circuit operating through the regulated expression of gRNAs that target other gRNAs and controls a gene of interest. In this case, SoxS is recruited through the MS2 coat protein (MCP). (3) A transcriptional activation tool in which recruitment of the effector SoxS is mediated by heterodimerization domains (HDD) responsive to a small-molecule. (4) A nCas9-based tool that uses *Petromyzon marinus* cytidine deaminase 1 (PmCDA1) and uracil DNA glycosylase inhibitor (UGI) for making DNA edits at a location specified by the gRNA. (5) A dCas12-based transcriptional repression tool capable of downregulation transcription of a gene of interest. (6) A dCas13-based tool tether to tRNA adenosine deaminase A (TadA) capable of targeting specific RNAs for base editing. (7) A Cas13-based tool that cleaves target RNA and can trigger collateral cleavage activity of any RNA. (8) A nCas9 tool capable of recording the presence of an mRNA by making DNA edits with the tethered rat apolipoprotein B mRNA editing enzyme catalytic subunit 1 (rAPOBEC1). (9) A Cas9 system capable of making precise DNA double stranded cuts. (10) A dCas13 translational repression system that downregulates protein translation from a specific mRNA.

Engineering CRISPR components to build diverse molecular tools

Since the first demonstrations of genetic engineering with Cas9, a plethora of CRISPR tools based on different Cas proteins have been developed. The fast and wide adoptance of CRISPR tools has been fueled mainly by two inherent properties: i) designability of the CRISPR systems based on different Cas proteins, tethered effectors, and gRNA properties, and ii) multiplexibility of CRISPR programs based on orthogonal gRNAs. In this section, we describe the different CRISPR components, how they can be engineered, and how they can be combined in a modular fashion to suit specific applications.

Figure 2: Design Considerations when Engineering CRISPR Tools

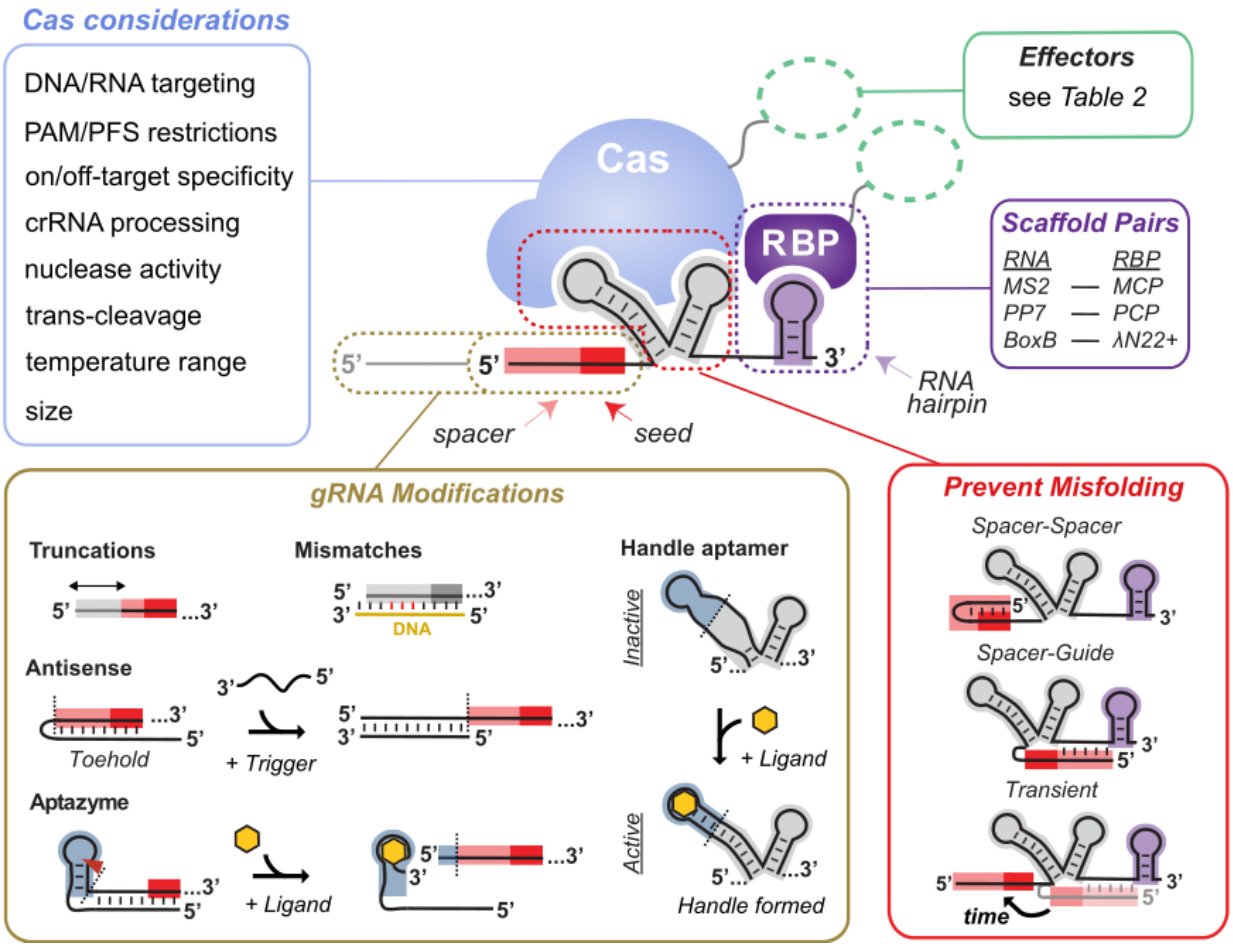


Figure 2: Schematic of the different CRISPR components and how they can be engineered. Given the diversity of Cas proteins (Table 1), several properties, such as DNA or RNA targeting and trans-cleavage activity, must be considered based on the specific application. Effectors for transcriptional regulation or base editing (Table 2) can be covalently tethered to the Cas protein, or recruited through a set of orthogonal scaffold pairs consisting of a RNA binding protein and the cognate RNA hairpin encoded in the gRNA. When designing gRNAs, there are several techniques to tune their activity and computational tools to ensure proper folding (see *gRNA design strategies* section).

The different types of CRISPR-Cas proteins

Natural CRISPR-Cas systems differ greatly in their nucleic acid recognition capabilities and additional functionality such as collateral cleavage and transposase activity (55). The current CRISPR-Cas classification encompasses two distinct classes differentiated by the architecture of the interference module involved in target recognition and nucleic acid cleavage. Class 1 CRISPR systems, subdivided into types I, III, and IV, include ~90% of all CRISPR loci identified in bacterial and archaeal genomes and contain interference modules composed of multiple genes encoding interference subunits. In contrast, class 2 CRISPR systems have condensed interference modules housed within a single multi-domain protein. Importantly, class 2 systems have proven more amenable to engineering efforts due to this compact nature, and therefore will be the focus of this review. Initially represented solely by Cas9 (type II), there are now numerous Cas9 variants as well as Cas12 (type V) and Cas13 (type VI) serving as the basis for the development of versatile and orthogonal molecular tools for biosensing and bioproduction applications.

Type II: the Cas9 workhorse for DNA targeting

The DNA binding activity of Type II CRISPR systems can easily be controlled by an engineered small guide RNA (gRNA), making them particularly well-suited for DNA manipulation applications. In general, type II systems consist of a Cas9 protein and two RNAs, a trans-activating RNA (tracrRNA) and a CRISPR RNA (crRNA). The tracrRNA

serves as the binding scaffold for the Cas9 protein, while the crRNA is the complementary sequence or spacer for the target DNA. In native systems, the tracrRNA and crRNA hybridize; however, many engineered systems fuse the two into a single guide RNA (sgRNA) with similar performance in vitro (12). Cas9 DNA targeting requires two key features: DNA target and gRNA spacer complementarity, and a protospacer adjacent motif (PAM) directly flanking the DNA target, necessary for preventing self-digestion of the gRNA encoding sequence. Mechanistically, Cas9 first scans DNA for potential PAMs and then checks guide-target complementarity. Once guide-target complementarity is validated (56, 57), the nuclease domains of Cas9 generate double-stranded breaks (DSBs) into the target DNA. While mismatched spacer sequences can be tolerated, strict PAM requirements often limit Cas9 targeting (57, 58). For instance, the density of PAM sequences from the most widely used Cas9, SpyCas9 from *Streptococcus pyogenes*, is only ~1 PAM (5'-NGG-3') every 10 bases in the regions upstream of endogenous promoters in the *E. coli* genome (50.5% GC) (25). PAM density for SpyCas9 is potentially higher in organisms whose genome has high GC-content (e.g. *P. putida* 61.5% GC) but will be lower in AT-rich bacteria (e.g. *C. autoethanogenum* 31.1% GC). Therefore, engineering non-model organisms with low GC-content could be challenging with SpyCas9 alone.

To increase the range of PAMs and associated loci for targeting, many groups have mined for new orthologs using bioinformatic tools (59, 60). Other type II Cas9s systems were found to have completely different consensus PAM sequences, a wide range of sizes, and distinct tracrRNAs (Table 1). For example, SmacCas9 from *Streptococcus macacae* has adenine-rich PAM preference (5′-NAAN-3′) which enables

targeting at AT-rich regions (61). ScCas9 from *Streptococcus canis* has less restrictive 5′-NNG-3′ PAM (62). On the other hand, Sth1Cas9 and Sth3Cas9 from *Streptococcus thermophilus* have more stringent PAMs, 5′-NNAGAAW-3′ and 5′-NGGNG-3′, respectively; however, longer PAMs offer significantly lower off-target effects (63, 64). To facilitate prototyping of new Cas proteins, many groups have built computational pipelines to predict functional PAM sequences of various CRISPR-Cas systems. Rybincky and colleagues built Spacer2PAM to predict PAM sequences for 10 microbial organisms (65). Similarly, CASPERpam enables assessment of PAM sequences for all available CRISPR systems in the NCBI database of bacterial genomes (66). For instance, a type II-C CRISPR system from *M. heyeri* investigated by Spacer2PAM showed broader PAM-flexibility (5'-GSNN-3') compared to SpyCas9. Additional details on the size and compatibility of PAMs from different Cas9 orthologs beyond the scope of this review can be found in the recent review by Collias & Beisel 2021 et al. (67).

Table 1: Cas proteins prevalent in Prokaryotic Applications

Type	Cas protein	Size (aa)	PAM/PFS	Target substrates	Refs.
Type II	SpyCas9	~1400	NGG	dsDNA	(12)
	xCas9-NG	~1400	NGN	dsDNA	(68)
	SpRYCas9	~1400	NRN > NYN	dsDNA	(69)

	SmacCas9	~1400	NAA	dsDNA	(61)
	Sth1Cas9, Sth3Cas9	~1200	NNAGAAW, NGGNG	dsDNA	(63, 64)
	ScCas9	~1400	NNG	dsDNA	(62)
	Nme1Cas9, Nme2Cas9, Nme3Cas9	~1000	N4GYTT, N4CC, N4CAAA	dsDNA	(70)
	SauCas9	~1000	NNGRRT	dsDNA	(71)
	CjCas9	~1000	N4RYAC	dsDNA	(72)
	SauriCas9, SacCas9	~1000	NNGG, NAAN	dsDNA	(73)
Type V	Cas12a	~1500	TTV	dsDNA, ssDNA	(22)
	AsCas12a -RR	~1500	TYCV	dsDNA, ssDNA	(74)
	AsCas12a -RVR	~1500	TATV	dsDNA, ssDNA	(74)

	enAsCas12a - HF	~1500	TRTV	dsDNA, ssDNA	(75)
	Cas12b	~1300	TTN	dsDNA, ssDNA	(76)
	Cas12c	~1300	TG or TN	dsDNA, ssDNA	(77, 78)
	Cas12d, or CasY	~1300	TA or TG	dsDNA, ssDNA	(79)
	Cas12e or CasX	~1000	TTCN	dsDNA	(80)
	Cas12f, or Cas14	~600	-	ssDNA	(81, 82)
	Casμ	~500	ATTA	dsDNA	(83)
	Cas12g	~800	-	ssRNA	(84)
	Cas12h	~900	RTR	dsDNA, ssDNA	(85)
	Cas12i	~1,100	TTN	dsDNA, ssDNA	(85)

	Cas12j, or CasΦ	~750	TBN	dsDNA, ssDNA	(86, 87)
	Cas12k	~650	GTN	dsDNA	(88)
	Cas12l	~850	CCY	dsDNA	(89, 90)
	Cas12m	~600	TTN	dsDNA	(78)
	Cas12n	~500	AAN	dsDNA, ssDNA	(91)
Type VI	Cas13a	~1200	3’ H	ssRNA	(92)
	Cas13b	~1100	5’ D, 3’ NNA, NAN	ssRNA	(93)
	Cas13c	~1100	-	ssRNA	(94)
	Cas13d	~900	-	ssRNA	(95)
	Cas13X, Cas13Y	~750	-	ssRNA	(96)

224

As an alternative to mining Cas9 orthologs, protein engineering has been successfully employed to broaden or alter the specificity of PAM recognition domains in common CRISPR nucleases (Table 1). These protein engineering efforts have ranged from chimera construction, random mutagenesis, structure-guided design, and high-throughput PAM determination assays (68, 69, 97–99). Kleinstiver et al. reported modification of SpyCas9 PAM recognition, where they used structural information to perform bacterial directed evolution (97). This approach generated many SpyCas9 variants that shifted the SpyCas9 PAM recognition from 5′-NGG-3′ to 5′-NGA-3′ (VQR variant), 5′-NGAG-3′ (EQR variant), or 5′-NGCG-3′ (VRER variant). Hu et al. used phage-assisted continuous evolution of the a-helical recognition domain to evolve expanded PAM SpyCas9 variants capable of targeting 5′-NG-3′ (xCas9-3.7), 5′-GAA-3′, and 5′-GAT-3′ (68). In pursuit of similar goals, Nishimasu et al. screened structure-guided designs to develop another expanded PAM variant, SpCas9-NG, also targeting 5’-NG-3’ PAMs (69). Comparison of xCas9 and SpCas9-NG demonstrated that SpCas9-NG has superior 5’-NG-3’ PAM recognition. Similar engineering efforts to increase PAM-flexibility were also conducted in other Cas9 variants (100–102). Overall, near PAM-less Cas9 variants whose specificity depends predominantly on target-spacer complementarity have shown to improve target accessibility and efficacy in both eukaryotic (98, 99, 103) and bacterial systems (102, 104, 105).

However, PAM expansion can lead to extended timescales for target recognition and increased off-target effects (57, 58, 106). To address off-target effects, both rational engineering and enzyme selection approaches have been used to develop high-fidelity SpyCas9 variants with reduced off-target activity (107–112). For instance, eSpCas9(1.1),

248 was engineered by disrupting interactions between positively charged residues of Cas9 and negatively charged target DNA backbone (112). Selection-based engineering yielded
250 EvoCas9 and Sniper-Cas9 which showed improved fidelity compared to previous Cas9 variants (107, 111). Introducing mutations that confer high-fidelity traits into a PAM-
252 expanded variant could also mitigate the off-target effect instigated by PAM flexibility (99). However, most of the high-fidelity variants have yet to be characterized across bacteria.

254 In the case of nuclease-deactivated Cas9 (dCas9), SpydCas9 has also been shown to bind non-specifically to NGG PAM sites (106) and to off-target genomic loci with
256 up to six mismatched nucleotides that differ from the spacer (113). dCas9 exhibits toxicity rooting from this promiscuous DNA binding effect when expressed at high level in bacteria
258 (114, 115). Although previous works have demonstrated that Cas9 off-target binding directly scales with PAM flexibility (69, 97, 99), we observed that increased dCas9 PAM-
260 flexibility has no additional growth defect compared to original dCas9, possibly due to reducing CRISPRi-like effect (104). This finding suggests that adverse effects from Cas
262 engineering might behave differently depending on the biological function of the Cas proteins.

264 Overall, PAM expansion has increased our targeting ability with Cas9-based genetic tools, but not without functional tradeoffs. Hence, frameworks for choosing the
266 most effective variant for a given set of targets and specific applications are necessary. For instance, although PAM-flexible dCas9 variants exhibit greater CRISPRa efficiency,
268 CRISPRi gene repression efficiency with the same variants is decreased (104) . To this end, there are now tools such as CasPEDIA, an encyclopedia of CRISPR systems with
270 comprehensive descriptions of enzyme activities and experimental considerations (116).

Systematic characterization of Cas9 variants across bacteria will enhance these databases and streamline the process of choosing the optimal Cas9 variant for a given set of target genes and specific applications.

Type V: a diverse Cas12 family

After the Type II Cas9 CRISPR system, the second most broadly used systems are Type V interference effectors, Cas12 proteins. Cas12 enzymes possess a RuvC-like nuclease domain that sequentially cleaves the strands of the target nucleic acid and, unlike Cas9, generate sticky ends. (21, 22, 117). This makes Cas12 particularly useful for precise *in vitro* DNA assembly. Several Cas12 enzyme variants also contain a subdomain involved in pre-crRNA processing, while others encode a tracrRNA that, along with RNase III, processes the pre-crRNA (21, 118). Recent computational analysis has indicated that type V CRISPR interference effectors evolved from transposon-encoded TnpB nucleases on multiple, independent occasions, resulting in a large pool of type V variants (118, 119). As of 2020, these variants are grouped into 12 subtypes which are highly diverse with respect to the size and architecture of the interference effector proteins as well as molecular mechanisms (120) (Table1).

The corresponding Type V CRISPR proteins can be categorized based on their properties and functions. Cas12a, previously known as Cpf1, has dual-nuclease activity involved in both crRNA biogenesis and target DNA interference (21). Upon recognition of the 5′-TTV-3′ PAM and a 18-23nt seed region complementary to the crRNA sequence, Cas12a induces a staggered double-strand break in the target DNA (22). An intriguing feature of Cas12a is its collateral cleavage activity; after target binding and cleavage, Cas12a turns into an active nuclease that continues to nonspecifically cleave DNA

sequences (121). Although Cas12a was the first discovered and extensively studied type V interference effector, other Cas12 variants offer unique properties better suited for developing them as molecular tools. Cas12b, formerly known as C2c1, requires a tracrRNA, is smaller than Cas12a, and tolerates fewer mismatches (76). Cas12e, also known as CasX, is substantially smaller than Cas12a and exhibits minimal trans-cleavage activity (80). Cas12f and Cas12j, or Cas14 and CasΦ, respectively, are the smallest of the type V interference effectors, having only ~500-700aa (81, 82, 86, 87). Recently, Casμ, another miniature Cas12f interference effector (<490aa), was discovered to have broader PAM preference (83). Cas12g is the only type V RNA-guided ribonuclease that targets a single-stranded RNA substrate (84). Several Cas12 variants are also notable for their distinct compatible PAM sequences. Cas12l, or Casπ, in contrast to the rest of the subtypes, exclusively recognizes C-rich PAMs (5'-CCY-3′) (89, 90). Cas12n recognizes A-rich PAMs (5'-AAN-3′), and consists of only 506aa (91).

Apart from nuclease activity, some Cas12 proteins retain DNA-targeting efficiency but exhibit different biological functions. Cas12c and the recently discovered Cas12m function through binding but not cleavage of the dsDNA target, silencing transcription or replication (77, 78). Cas12k encompasses an inactivated RuvC-like domain also lacking DNAse activity, and associates with a Tn7-like transposase subunits to mediate RNA-guided transposition (76). Even more type V interference effectors are actively being discovered across metagenome-assembled genomes and viruses (122). As with Cas9, there have been many efforts to engineer Cas12 variants in order to improve cleavage accuracy (123), abolished trans-activity (124), increase PAM-flexibility (74, 75, 125), and thermostable mutants for biosensing applications (126).

Type VI: targeting RNA with Cas13

Type VI Cas13 is the most well-studied CRISPR system for directly targeting and manipulating RNA (127). Cas13 is further divided into four subtypes; A (previously known as C2c2), B, C, and D (Table 1). Cas13 systems typically contain two higher eukaryote and prokaryote nucleotide-binding (HEPN) RNase domains. The HEPN domains are able to cleave RNA upon activation through the crRNA base pairing with a matching target RNA (128). Upon activation, Cas13 exhibits collateral activity, similar to Cas12a, which non-specifically cleaves RNAs after cleaving its target. The collateral RNA cleavage activity was proved to prevent bacteriophage infection by shifting the cells into dormant stage (129). Through the HEPN domains, Cas13 is also able to process its own crRNA arrays into mature crRNAs of roughly 60 nt in length (130–132). Natural crRNA spacer length has been reported to vary between 14 and 30 nucleotides (92, 133). Instead of PAM, most Cas13 have a protospacer flanking sequence (PFS) preference which is usually less restrictive compared to the PAM requirements of DNA-targeting Cas proteins (92, 131). Although Cas13a and Cas13b were among the first discovered (127), recent research has shifted attention to Cas13d, which exhibited improved efficacy for RNA silencing and knockdowns (133, 134). Several Cas13 proteins have also been repurposed for various applications beyond RNA cleavage including *in vitro* RNA detection assays (135), RNA editing of mammalian cells and bacteriophage genomes (136–140), and RNA interference (133, 134).

The flexibility in nucleic acid recognition capabilities provided by the different Cas9, Cas12, and Cas13 variants has already fueled many applications in prokaryotic systems, from editing and regulating genes to biosensing nucleic acids and small molecules (39,

141). Nevertheless, the full repertoire of Cas variants developed for eukaryotic applications is yet to be characterized and exploited across different bacteria. Given that genomic GC content varies greatly from 13% to 75% between prokaryotic species (142), the diverse catalog of Cas variants will provide suitable options for implementing CRISPR systems in any microbe with specific traits for each application, including bioproduction, biotherapeutics, and bioremediation.

Expanding CRISPR functionality with tethered effectors

Abolishing Cas nuclease activity generates nickase or catalytically-dead Cas proteins (nCas or dCas, respectively), opening a new paradigm of CRISPR tools for sequence-specific nucleic acid modifications and gene expression regulation. Point mutations in one of the Cas nuclease domains generates a ‘nicking’ variant that only cleaves one strand of DNA, while mutations in both domains form a 'catalytically-dead' variant with no cleavage activity. In prokaryotes, dCas9 was rapidly repurposed for programmable transcriptional repression by physically blocking the transcription initiation or elongation processes. Tethering effector domains to the CRISPR complex enabled sequence-specific localization of effectors to a DNA or RNA of interest, for applications such as transcriptional activation and base editing (6). Effector tethering has been achieved through covalent fusion to the Cas protein, protein-protein interactions (PPI), or recruitment of effectors fused to RNA binding proteins (RBP) that interact non-covalently with modified guide RNAs (scaffold RNA, scRNA). Importantly, orthogonal effector recruitment can be programmed with multiple PPI systems or RBP-RNA hairpin pairs (Figure 2). In contrast to eukaryotes, where CRISPR transcriptional activation (CRISPRa) can readily be achieved with a universal transcriptional activator (143, 144), CRISPRa across prokaryotes remains challenging due to the distinct transcriptional regulation mechanisms present in different bacteria. Here we focus on the development of transcriptional activators for prokaryotic CRISPRa as well as recent advancements in tethering effectors for CRISPR gene editing.

Effectors for CRISPR-based gene regulation

Bikard et al. showed that transcriptional activation could be achieved through recruitment of accessory proteins in prokaryotes similarly to previous works in eukaryotes. They developed a CRISPRa system in *E. coli* by fusing a RNA polymerase (RNAP) omega subunit (RpoZ) to dCas9 (145). One limitation of this system was the need to delete the endogenous *rpoZ* gene. Later work from our group found that bacterial CRISPRa could also function through a bacterial transcription factor, SoxS, using RBP-based recruitment (146). Notably, SoxS-based CRISPRa does not require modification of the genomic background, which streamlines the portability to other bacterial systems. Another approach by Liu et al. achieved CRISPRa through the RBP recruitment of the transcriptional activator PspF (147). Interestingly, we found that PspF and SoxS could be implemented simultaneously through orthogonal RNA hairpin-RBP pairs —MS2/MCP for SoxS and BoxB/λN22 for PspF— and also work complementary to each other. SoxS can activate a broad range of σ^{70} superfamily promoters but is ineffective toward σ^{54}-family promoters (25). On the other hand, PspF system works specifically on the σ^{54}-family promoters due to its unique ATP-dependent activation mechanism but cannot activate σ^{70} family promoters (147). In both cases, strategies to abolish the DNA-binding of transcriptional effectors proved to be crucial since CRISPRa DNA targeting should be governed by dCas9 rather than the transcriptional effectors. In PspF-based CRISPRa, this was achieved by truncating the helix-turn-helix (HTH) DNA-recognition motif of PspF (147). In SoxS-based CRISPRa, we reduced the DNA-binding capacity of heterologous SoxS through multiple mutations, leading to key improvement in CRISPRa performance (25).

Through characterization of different CRISPRa elements and target genes, we have found that CRISPRa performance relies on i) composition of the target promoter, and ii) position of target sequence relative to the transcription start site (Figure 3A) (25, 48). These design factors remain consistent across multiple transcriptional activation systems (25, 147). For instance, the dynamic range of CRISPR-based gene overexpression is dependent on DNA helical phases. Shifting the target site 1-3 bp away from the optimal position leads to a sharp drop in fold-activation (Figure 3B). However, the fold-activation was recovered to a similar level if further shifted to 10-11 bp, resembling that of a full turn of DNA helix. We also found that these two systems exhibited consistent design rules in both *E. coli* and *P. putida*.

Figure 3: Prokaryotic CRISPRa Sequence and Distance Requirements

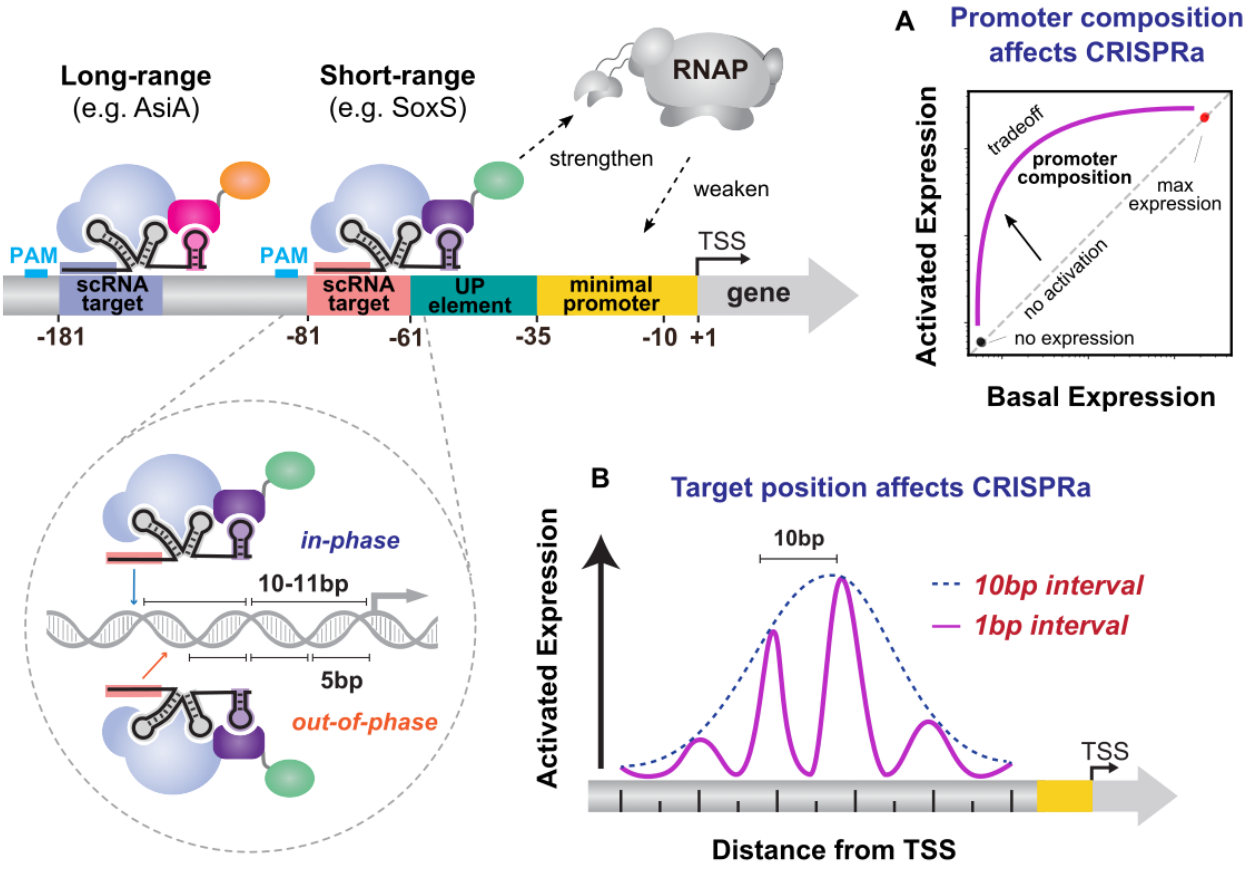


Figure 3: Schematic of a prokaryotic promoter showing stringent requirements for bacterial CRISPRa. RNAP affinity for the minimal promoter and UP-element determines basal expression levels. When CRISPRa is targeted to a promoter with a correct PAM and complementary scRNA target site, RNAP is recruited by the transcriptional activator. A) Promoter composition determines RNAP-promoter and CRISPRa-promoter interactions and hence affects CRISPRa-mediated expression levels. B) CRISPRa-mediated expression is highly dependent on the distance of CRISPRa with respect to the transcription start site (TSS), and exhibits a periodic dependence with peak activities every 10–11 bases, corresponding to a single turn of DNA helix.

Since most bacterial CRISPRa systems are effective in a narrow range of target positions, there have been several attempts to develop CRISPRa systems active at alternative positions (Table 2A). Directed evolution of AsiA, an anti-sigma70 protein from T4 bacteriophage, covalently fused to dCas9 provides long distance CRISPRa beyond previously reported systems (148). Notably, this system was tested in a library of minimal promoters from a metagenomic dataset and found that 248 out of ~8,000 promoters could be upregulated from a long-range CRISPRa. Another strategy to overcome the stringent CRISPRa distance requirements has been to use circular permuted dCas9 (cpdCas9) variants to change the positioning of the tethered effector (149). They achieved increased targeting flexibility by screening a library of cpdCas9s that recruited the N-terminal domain of RpoA (αNTD) of 8 different bacterial species through heterospecific coiled-coiled peptides (SYNZIPs) recruitment (149, 150).

Even though most of the CRISPRa systems in bacteria rely on dCas9 to recruit effector proteins, other Cas systems could also be implemented in a similar manner. Direct fusion of RpoZ or SoxS effectors to dCas12a were implemented for CRISPR activation in *P. polymyxa* and *C. glutamicum*, respectively (151, 152). Type I-E Cascade-mediated bacterial CRISPRa is also recently demonstrated in a CRISPR-free *E. coli* using αNTD as an effector (153) (Table 2A). It should be noted that CRISPRa could be improved from engineering efforts on other parts beyond the effector domain and could also be developed in parallel from effector-specific optimization, e.g. dCas9 engineering for PAM-flexibility and reporter engineering to improve dynamic range (48, 104).

Table 2A: Effectors for prokaryotic CRISPRa, their mechanisms, and design rules.

CRISPRa systems	Mechanism	Optimal sites [a]	Strand	Organisms	Refs.
Effectors tethered directly to dCas					
dCas9-RpoZ	RNAP assembly	-60 to -100	NT	*E. coli,* *B. subtilis,* *L. enzymogenes,* *M. xanthus*	(145, 154–156)
dCas9-RpoA	RNAP assembly	-267 to -415	T	*B. subtilis*	(154)
dCas9-SYNZIP-αNTD	RNAP assembly	-60 to -100	NT	*E. coli*	(149, 150)
dCas9-RpoD	RNAP assembly	-199 to -216	NT	*S. oneidensis*	(157)
dCas9-AsiA [b]	RpoD and RpoB binding	-182 to -252	T	*E. coli,* *S. enterica,* *K. oxytoca*	(148)
dCas12a-SoxS	RpoA and RpoD binding	-50 to -150	T	*P. polymyxa*	(152)
dCas12a-RpoZ	RNAP assembly	-183 to -328	T	*C. glutamicum*	(151)
dCas12a-	Unknown	-90	NT	*B. subtilis*	(158)

RemA					
Cascade-αNTD	RNAP assembly	-100 to -120	T	*E. coli*	(153)
dCas13d-IF	Ribosome recruitment	5'-UTR	-	*E. coli*	(159)
Effectors tethered via guide RNA hairpin					
SoxS [b] (MCP or PCP or MCP + SYNZIP)	RpoA and RpoD binding	-60 to -100	NT	*E. coli, P. putida,* Cell-Free System, Hydrogel	(25, 45, 48, 146, 160, 161)
TetD (MCP)	RpoA and RpoD binding	-60 to -100	NT	*E. coli*	(25, 146)
RpoZ (MCP)	RNAP assembly	-91	NT	*E. coli*	(25, 146)
αNTD (MCP)	RNAP assembly	-60 to -100	T	*E. coli*	(25, 146)
PspF [b] (λN22 or MCP)	RpoN recruitment [c]	-91 to -131[d]	NT	*E. coli, K. oxytoca,* Cell-Free System	(147, 162)

Notes: [a]Relative to transcription start site (TSS), [b]Engineered proteins, [c]RpoN promoter is distinct to other bacterial promoters of RpoD-promoter superfamily, [d]DNA looping is present in the promoter

Effectors for CRISPR-based gene editing

In comparison to traditional CRISPR editing technologies that induce DSB breaks, base editors (BEs) with tethered deaminases only nick one strand of DNA. In bacteria, given the species-dependent variations in editing efficiency (163) and the lethality of

generating DSB in unintended targets (164, 165), BE is often favored over traditional editing through native Cas cleavage and DNA repair mechanisms (8, 9). Current prokaryotic BEs rely on direct tethering of the deaminase to inactivated Cas9 proteins, dCas9 or nCas9, but editing with other Cas variants such as dCas12a has also been demonstrated (Table 2B). After Cas recruitment to the DNA, the deaminase recognizes and edits the target ssDNA in the R-loop structure formed between the gRNA and the target DNA. Each BE system targets a characteristic proximal stretch of ssDNA, or editing window, based on the deaminase and tethering approach (Base editor effectors table). Adenosine base editors (ABE) generate A>G through A-to-I deamination, while cytidine base editors (CBEs) generate C>T substitutions through C-to-U deamination and subsequent DNA replication (Table 2B). Most CBE systems also tether a uracil glycosylase inhibitor (UGI) to prevent base excision repair and reversal of the point mutation (166). Notably, CBEs are more widely used in industrial settings than ABEs, likely due to the higher editing efficiency of CBEs and high GC content of bacterial genomes (167). Moreover, developing RBP-tethered BEs could allow for simultaneous regulation of multiple genes and even opportunities for multiplexed applications with other effector proteins (168). For advanced applications, the potential for multiplexing BE has been shown in multiple non-model organisms (11, 169–172).

Table 2B: Effectors for CRISPR-based gene editing and their editing window.

Editing system	Effector	Cas protein	Organisms	Editing window	Refs.
CBE	rAPOBEC1	Cas9n	*E. coli,* *B. melitensis*	4 to 8	(173)
		Cas9n	*C. glutamicum*	4 to 7	(170)
		Cas9n	*S. typhimurium*	5 to 10	(35)
		enCas9	*P. putida*	3 to 8	(174)
		dCas9	*B. subtilis*	17 to 20	(171)
		dCas12a	*E. coli*	8 to 13, 10 to 12	(175)
		Cas12m/dCas12m	*E. coli*	2 to 5, 13 to 19	(78)
	PmCDA1	dCas9	*E. coli*	1 to 5	(169)
		dCas9	*B. subtilis*	1 to 5	(172)
		dCas9	*Streptomyces*	1 to 5	(176)

			spp.		
		dCas9	*P. polymyxa*	17 to 20	(171)
		dCas9	*Agrobacterium spp.*	16 to 20	(177)
ABE	TadA	dCas9/nCas9	*E. coli*	4 to 8	(178)
		Cas9n	*S. aureus*	4 to 8	(179)
		dxCas9/nxCas9	*P. putida, Pseudomonas spp.*	4 to 8	(180)
EvolvR	Poll3M	nCas9	*E. coli*	~60	(181)
PE	M-MLV2	nCas9	*E. coli*	-	(182)

In addition to deaminases, other effectors have been tethered to expand the CRISPR editing capabilities beyond single-nucleotide substitutions. CRISPR-EvolvR uses an error-prone, nick-translating DNA polymerase (Poll3M) fused to nCas9 to continuously diversify nucleotides within short DNA regions (60 nt) (181). CRISPR prime editing (PE) uses a reverse transcriptase (M-MLV2) fused to nCas9 to introduce deletions, substitutions, and insertions encoded in the 3′-extension of the gRNA (182). Recently, serine integrases have been added to CRISPR-PE for targeted genomic recruitment and integration of desired payloads in eukaryotes (183). However, the low editing efficiency limits the application of these tools: 40%, 20%, 10% for PE-based plasmid deletions, substitutions, and insertions, respectively, and half of that for genomic edits. Editing efficiency is even lower (<2%) when combining modalities or multiplexing two gRNAs.

A major bottleneck for base and prime editing is that gRNA spacer design is more challenging than traditional spacer design as it must also factor the constraints imposed by the editing effector (167). Spacers designed for BE must factor in deaminase type, editing window, target nucleotide position, and desired translation product. For PE, the 3' gRNA extension must be carefully designed in order to enable reverse transcription without affecting DNA targeting. Hence, traditional gRNA spacer design tools often cannot be directly translated into designing spacers for base editing. While a few tools exist to support BE- and PE-gRNA design in eukaryotes (184–187), their development in bacterial systems is fairly limited. Blin and colleagues published CRISPy-web 2.0 for designing BE-gRNAs in non-model organisms such as *Streptomycetes* (188), but has yet to be implemented for other bacteria.

Even when gRNAs can be successfully designed, editing applications are still limited by the toxic effects of different combinations of effectors and Cas proteins (169, 189). For example in *E. coli*, nCas9-PmCDA1 has shown poor transformation efficiency, while dCas9-PmCDA1-UGI-LVA (LVA, degradation tag) demonstrated an improved transformation efficiency and BE editing efficiency (169). Hence, most studies require tedious screening combinations of Cas9 variants, deaminases, and expression cassettes to find a balance between efficiency and toxicity (171, 174, 190). To streamline this process, Shelake and colleagues developed a framework for high-throughput editing efficiency screening in *E. coli*, but only characterized editing efficiencies across a small set of sgRNA targets for GFP reporters (189). Moreover, sequencing-based BE efficiency analysis is currently expensive and time-consuming (191), and *in vivo*, real-time assessment of editing efficiency is only available in mammalian cells (192). Adaptation of these tools for bacterial systems would enable rapid comparison of Cas9 nucleases and deaminases across diverse genetic backgrounds, elucidating the design rules for predictable implementation.

Tethering different effector domains to different Cas proteins has enabled a wide range of functions beyond nucleic acid cleavage. These CRISPR gene regulation and base editing tools have already been implemented individually across different microbes (27, 193–195). Moving forward, simultaneous integration of multiple CRISPR tools in the diverse prokaryotic hosts should be within reach (26). Discovery and engineering of orthogonal hairpins-RBP pairs will enable the multiplexed recruitment of different effectors to specific sites. Moreover, multi-Cas systems for combined DNA and RNA targeting could be achieved with a better understanding of which Cas variants are better suited for

each host as well as expression mechanisms to minimize burden effects (196, 197). Taken together, these advancements will fuel the next generation of prokaryotic CRISPR tools capable of concerted editing and regulation (198).

Tuning CRISPR function through gRNA engineering

RNA-guided Cas proteins make engineering prokaryotic systems more straightforward than previously possible because gRNAs are genetically compact, metabolically inexpensive, and relatively simple to design (199). CRISPR functions can be tuned not only through Cas and effector protein engineering but through gRNA design as well. gRNAs can be engineered through chemical functionalization, sequence modifications, and fusion of other RNA scaffolds. Here we pay special attention to approaches that rely on engineering intrinsic gRNA properties as they are more relevant for prokaryotic applications in which the gRNA(s) are transcribed.

gRNA design strategies

Several gRNA design strategies have been proposed to tune CRISPR activity, mainly through spacer mismatches, truncations, and extensions (26, 200–202) (Figure 2). Mismatches in the Cas9 spacer region have been used to achieve graded levels of gRNA activity (203–205). Generally, mismatches at the PAM-distal end of the spacer are better tolerated than mismatches at the PAM-proximal, or seed region, of the spacer. Mismatches within the seed region exacerbate R-loop destabilization, whereas PAM-distal mismatches reduce the lifetime of CRISPR-DNA complex (206). Although mismatches in the non-seed region can be used to tune gRNA activity, it comes at the cost of gRNA specificity as off-target activity at complementary sites is increased (202).

Truncating the gRNA spacer from the PAM-distal region up to the seed region has also been shown to reduce the level of gRNA activity. A possible concern when using truncated gRNAs is an increase in off-target activity due to lower target complementary constraints. However, in a test set of orthogonal, synthetic, CRISPR-activatable promoters, we found that truncated (11-20 bp) scRNAs did not induce higher expression in their non-cognate, unactivated promoters, suggesting that truncated scRNAs do not generate non-specific activation of heterologous fluorescent proteins (in preparation). Moreover, evidence from eukaryotic gene editing experiments with truncated gRNAs points to the reduced binding energy as a leading factor in increased sensitivity to mismatches (207). These observations have been rationalized in terms of an “excess energy” model, postulating that decreases in binding energy beyond a certain threshold destabilizes binding to both correct and incorrect targets (109, 207–209). In a similar fashion, introducing a proper 5' extension that folds back to cover the spacer region has been shown to increase gRNA specificity as the designed hairpin becomes more stable (210). However, long 5’ extension gRNA spacers could also lead to decreased activity (210, 211).

5’ and 3’ gRNA extensions have been implemented to exert conditional control over CRISPR activity (Figure 2). The general mechanism involves triggering occlusion or reveal of the gRNA spacer or handle in response to a nucleic acid or small molecule. Several gRNA with 5’ toeholds have demonstrated the ability to respond to single DNA or RNA trigger strands (212–214). Aptazymes, or ligand-responsive self-cleaving ribozymes, have been used to cleave a repressive 5′ extension from the gRNA upon the addition of the target ligand (215). Aptamers have also been added into the handle or

560 hairpins to stabilize the gRNA in response to a specific ligand (216). While 3' end aptamers have been demonstrated in sgRNAs (217), 3' extensions could be challenging
562 in a scRNA due to the presence of the effector-recruiting hairpin. Finally, even though most gRNA engineering efforts have been on Cas9 gRNAs, some of the principles should
564 be extendible to gRNAs for Cas12 and Cas13 (210, 218–220).

Computational design of gRNAs

566 Designing gRNAs to target arbitrary DNA or RNA sites requires simply designing the complementary RNA sequence into the spacer. However, accurately predicting gRNA
568 performance in terms of activity and specificity is still challenging (221). One reason is that gRNA activity is highly influenced by target sequence context, both with respect to
570 the sequence surrounding the PAM site and the local target site structure (222). CRISPR complex binding and activity at endogenous target sites can be hindered by other DNA-
572 binding effectors regulating processes such as chromatin accessibility, DNA methylation, and transcriptional regulation (26, 223, 224). Even the inherent resilience of endogenous
574 gene regulatory networks containing feedback or feedforward loops can limit the degree of CRISPR activity. Machine learning models trained on large gene editing or CRISPRi
576 datasets in eukaryotes have successfully been used to extract gRNA design rules (225, 226). However, the applicability of these models for prokaryotic gRNAs has been shown
578 to be rather limited since many of these endogenous target contexts and processes differ greatly across the two domains of life.

580 gRNA activity is also influenced by the secondary structure that the gRNA adopts. The degree of secondary structure in particular, whether internal to the spacer or between

582 the spacer and the rest of the guide, has been observed to be a key determinant of gRNA activity (Figure 2). While some models do account for the energetics of nucleic acid
584 interactions, these have primarily focused on the thermodynamics of spacer-DNA target interactions, ignoring other aspects of guide RNA folding. Ensuring correct gRNA folding
586 is crucial when designing gRNAs with modifications such as extra hairpins or aptamers. To this end, computational RNA folding packages such as ViennaRNA have been used
588 in conjunction with deep learning models to predict on- and off-target activities for bacterial sgRNAs based on RNA folding parameters, melting temperatures, and potential
590 off-target scores (227).

Only until very recently have computational models started to combine all these
592 effects to more accurately predict gRNA performance in prokaryotic systems (228). Developing models that incorporate biophysical aspects of gRNA folding and target
594 sequence context with heuristics from large-scale functional screening is expected to further advance our gRNA design capabilities. Importantly, these models should
596 recapitulate the gRNA performance effects associated with modifications such as spacer truncations and aptamer extensions. Such an advance, if it can minimize or eliminate the
598 current need for experimental verification of new gRNAs, would accelerate the design of CRISPR tools for editing or regulating specific genes to precise levels.

600

602

Building CRISPR-based genetic circuits

Genetic circuits enable natural biological systems to respond to different input stimuli with complex, time-dependent behaviors by regulating the expression of multiple genes. Harnessing these capabilities is fundamental for engineered systems that need to monitor, compute, and respond to internal or external environment changes as demonstrated by the hallmark designs of genetic circuitry (229, 230). However, there are limited examples of engineered genetic circuits capable of multiplexing stimuli and regulating multiple genes. This paucity can be attributed to the limited number of suitable components for implementing scalable circuitry (1, 231–233), and to the difficulty of sequentially combining components into larger, high-level circuits (234–237). Due to the ease of designing and tuning new components that can be interconnected, CRISPR-based tools have rapidly become a suitable framework for building complex genetic circuits. In this section, we focus on how CRISPR tools have been assembled into genetic circuits in bacteria and cell-free expression systems.

CRISPR circuits consist of a series of CRISPR nodes, which are discrete transcriptional or translational units containing target sequences for CRISPRa or CRISPRi (Figure 4A), that are interconnected through gRNAs. Circuit topology is therefore specified by the set of connected nodes and gRNAs and can encode logic gates and network motifs, such as incoherent feedforward loops (Figure 4B), to perform precise functions. Importantly, more complex functions can be achieved by combining several network motifs into a larger circuit.

CRISPR logic gates and small network motifs

The first examples of prokaryotic CRISPR circuits relied on transcriptional repression to create logic gates capable of Boolean operations (NOT, AND, OR, NAND, NOR, etc.) (238). Transcriptional inverters, or NOT gates, have been designed by conditionally expressing sgRNAs from inducible promoters, such that the presence of the inducer leads to repression of a target gene through CRISPRi. Construction of NOR gates has been achieved by combining two NOT gates responsive to different inputs. Moreover, more complex network motifs such as toggle switches that can serve as memory units and incoherent feedforward loops for programming spatio-temporal patterns have been constructed by interconnecting multiple sgRNAs. Notable examples are a recently developed CRISPRi-based synthetic oscillator, termed CRISPRlator (239) and a CRISPRi-based synthetic inverter for antibiotic-free selection plasmid maintenance (240).

Although theoretically any computational operation could be achieved if enough transcriptional repression nodes are interconnected, in practice it is advantageous to combine repression and activation to cover a higher portion of the circuit design space with fewer components (45). The discovery of new transcriptional activators and promoter design rules for effective activation in bacteria has enabled CRISPRa-based circuits (Figure 3) (25, 48). With CRISPRa, construction of AND gates is straightforward through conditional expression of a scRNA and activator, a crRNA and tracrRNA, a toehold-gated gRNA and a trigger; conditional folding of the scRNA; or conditional recruitment of the activator through ligand-induced heterodimerization domains (48, 146, 162) (Figure 4D). OR gates can also be readily assembled by targeting two scRNAs to the same node. Moreover, nodes can be assembled in multi-layer activation cascades to control the

timing and expression of multiple genes. Importantly, circuits combining both CRISPRa and CRISPRi can be constructed through the regulated expression of sg/scRNAs. Implementation of simultaneous CRISPRa and CRISPRi has been demonstrated on both multiple and individual genes (146, 158, 160) , and in multi-layer circuits including activation-repression cascades and incoherent feedforward loops (45)(Figure 4E).

Figure 4: Building CRISPRa/i Circuits for Boolean Logic and Dynamic Control

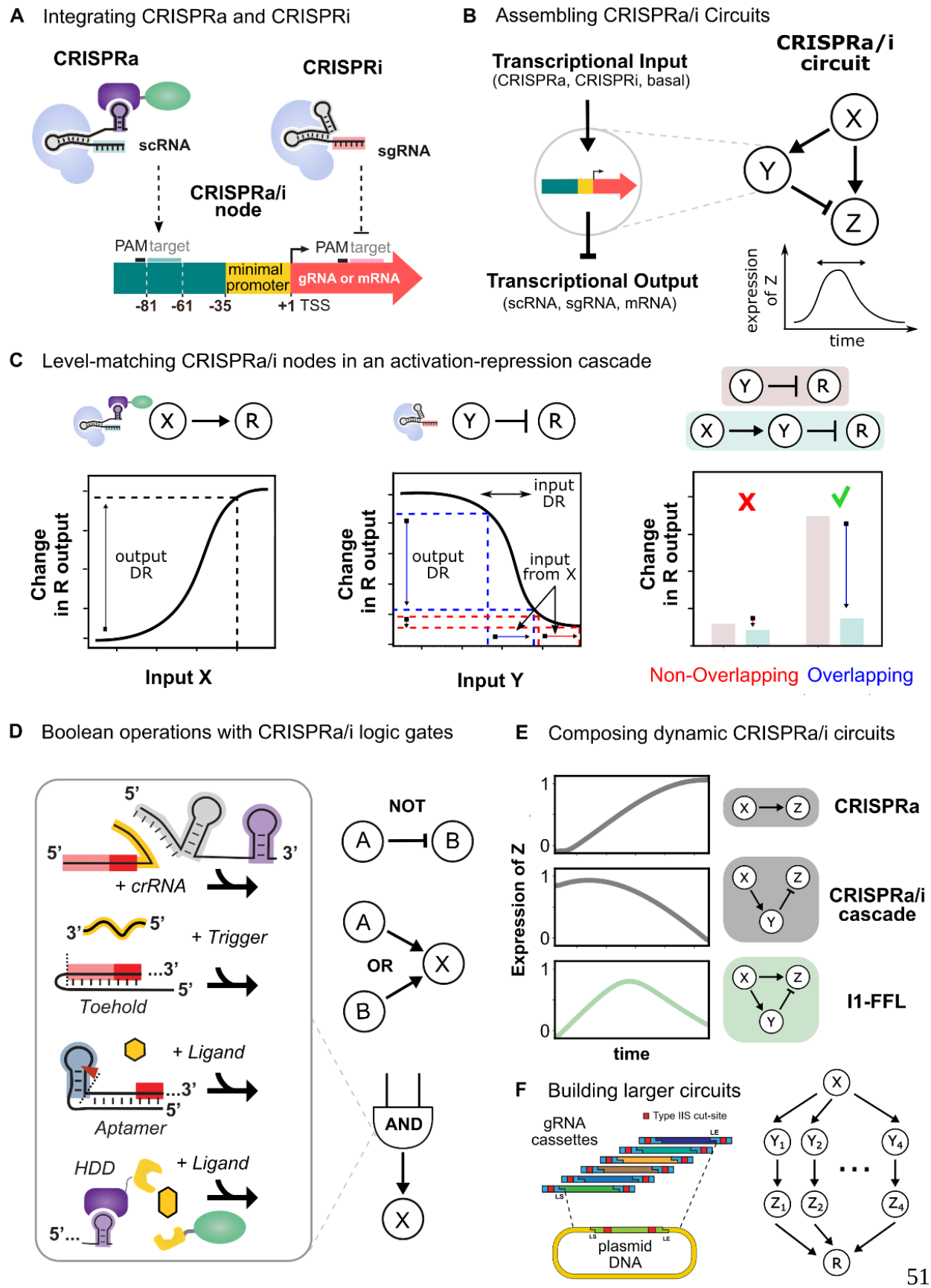


Figure 4: A) Schematic of a CRISPRa/i node, consisting of a transcriptional unit with target site for CRISPR activation and inhibition. B) CRISPR circuits can be assembled by interconnecting CRISPRa/i nodes expressing gRNAs. C) Successful circuit function requires level-matching the output dynamic range of an upstream node with the input dynamic range of a downstream node. On the left and middle panels, dose-response characterization of CRISPRa and CRISPRi, respectively, with overlays of input and output dynamic range (DR). On the right panel, functional characterization of CRISPRi and an activation-repression cascade when the dynamic ranges are matching or mismatched. D) Schematics of different techniques for incorporating Boolean logic into CRISPR circuits via gRNA programming. E) Dynamic gene expression profiles resulting from different CRISPRa/i circuits. The composability of CRISPR circuits enables building larger circuits from small motifs. F) Large CRISPR circuits can be achieved by assembling small gRNA cassettes into multi-guide arrays.

Prospects and challenges of building larger CRISPR circuits

Theoretical and experimental analyses indicate that the CRISPRa/i system is well suited to design complex genetic circuits with many internal nodes (45, 241). Crucial to this task is the ability to level-match the input/output dynamic ranges between sequential nodes. That is, the output transcription levels of upstream nodes encoding gRNAs must be matched to the relevant transcriptional input range of downstream nodes (FIgure 4C). Proper level-matching, and therefore circuit function, can be ensured by tuning gRNA expression through promoter engineering, or gRNA activity through gRNA design.

Several factors must be taken into account when assembling multi-gRNA circuits for CFES or bacteria. In cell-free systems, multiple gRNA can be readily expressed from individual plasmids or linear fragments (45, 242). In living cells, however, the repetitiveness of multiple gRNA parts could be difficult to chemically synthesize and prone to recombination (53). Expression of gRNA arrays from a single transcriptional unit is possible when using self-processing Cas such as Cas12a and Cas13 or incorporation of gRNA processing proteins (Csy4) (130, 243, 244). However, the order of gRNAs in the array also affects their activities, likely due to early termination in gRNA biogenesis (245). Additionally, incorporation of orthogonal CRISPRa/i targets in the gRNA expression cassette is required for gRNAs to regulate the expression of one another (48). Hence, expressing one gRNA per transcriptional unit is more appropriate for programmable gRNA circuits. Several studies have demonstrated the construction of large gRNA circuits using non-redundant parts of promoters, terminators, and gRNA scaffold (115, 246). As the number of regulatory nodes increases, tools to design high-performing, orthogonal promoters and gRNAs will become increasingly important. With a recent workflow for

promoter and gRNA design, circuits with up to three and six layers of activation have been achieved in *E. coli* and *E. coli*-based cell-free lysates, respectively (48) (Figure 4F).

Given that natural CRISPR immunity systems contain ~50 spacers, and some use multiple Cas proteins, more complex CRISPR circuits should be within reach (247). Currently, the main challenge in implementing larger CRISPR circuits is the retroactivity effects observed when expressing multiple gRNAs due to competition for the same pool of Cas proteins (241). For wide circuits regulating many genes in parallel, these effects have been minimized by tuning dCas9 expression levels through feedback control and RBS engineering, enabling simultaneous repression of up to 13 genes (248). For deep circuits with many internal layers, upstream gRNAs have been shown to outcompete downstream gRNAs for dCas9 binding, mainly due to timing of expression. For these circuits, strategies to dynamically regulate upstream gRNA expression, such as reversing CRISPRa complex binding or implementing negative autoregulation motifs, may be necessary. Moreover, we foresee that incorporation of orthogonal CRISPR systems, such as alternative dCas9 variants or even dCas12a and dCas13 systems, could provide additional nodes and layers for CRISPR-based biocomputing.

Forefront applications of CRISPR-based tools

Engineering CRISPR systems for Biosensing

Coupling molecular recognition to the diverse biochemical functions of the CRISPR toolbox has enabled new avenues for biosensing across various biochemical systems, including cell-free systems and bacterial hosts (Figure 5, top). In cell-free systems, the native nucleic acid recognition and cleavage activity of Cas enzymes has fueled *in vitro* detection of diverse biomarkers, from nucleic acids to small molecules (39, 40). Addition of gene expression capabilities to cell-free systems (41–43), has enabled more sophisticated biosensors. In bacteria, CRISPR-based biosensors are being used to record transcriptome changes and screen bioproducing strains. In this section, we recapitulate how different CRISPR systems have been engineered to detect nucleic acids, small molecules, and other environmental cues.

Nucleic acids detection

Nucleic acid detection is a critical tool in many biotechnology applications, from infectious diseases diagnosis to transcriptomic monitoring. The different Cas9, Cas12, and Cas13 systems have all been rapidly adapted for the specific, sensitive, and portable sensing of nucleic acids (39). One approach involves detecting a specific DNA/RNA sequence *in vitro* through gRNA complementarity which trigger cleavage of the surrounding quenched probes through collateral activity (mainly Cas12a and Cas13a) to produce a fluorescent readout (135, 249–251) (Figure 5A). When paired with target amplification processes, such as isothermal recombinase-polymerase amplification (RPA) or loop-mediated isothermal amplification (LAMP), the detection limit of the desired DNA/RNA can be as low as 10 aM (attomolar) and thus useful for SARS-CoV-2 detection

(249, 252–255). This detection approach has also been paired with cell-free expression systems harboring different enzyme-based reporters for higher amplification-free signal, and even packaged into a cell-free wearable biosensor (256, 257). For a more in depth review of CRISPR systems for *in vitro* nucleic acid detection assays, we point the reader to this recent Annual Review (40).

In a similar fashion, engineered Cas genes and gRNAs targeting antibiotic resistance or essential genes can be packaged into bacteriophages or nanoparticles to detect and lyse specific bacteria (258, 259). Cas9 has successfully been used to kill virulent *S. aureus*, *E. coli*, and *E. faecalis* (260–263). Similarly, Cas13 has also been used to detect and control bacteria populations or repurposed to engineer bacteriophage genome (36). Notably, the efficacy of this approach at mitigating infection has been shown to be equal to that of high dose antibiotics in *in vivo* models (261). Given the multiplexibility of CRISPR tools, these systems could readily be engineered for simultaneously targeting different species or multiple sequences in the same bacteria (264), or be interfaced with genetic circuits for logic control.

Engineering tracrRNAs to hybridize with target RNAs as if they were crRNAs creates another approach for detecting nucleic acids (162, 265). This naturally-occurring process was rapidly repurposed for detecting viral RNA *in vitro* as well as endogenous mRNAs in different bacteria. Recently, this approach has been paired with base editing to develop a mRNA recording platform (35). This recording platform has successfully been used for recording infection-induced sRNAs in the intracellular pathogen *Salmonella* and for tracking the spread of antibiotic resistance mobile genetic elements in an *E. coli* population (35). This mRNA detection approach has also been paired with CRISPRa and

connected to a positive feedback loop for signal amplification (162). In a similar fashion, strand displacement of 5' toeholds covering the gRNA spacer or handle have also been used to detect nucleic acids, from endogenous mRNA and sRNA to synthetic ssDNA that can be used for complex computation (212, 218, 219).

Small molecules quantification

Detection of small molecules enables diverse applications, such as monitoring environmental contaminants *in vitro* and quantifying cellular metabolites *in vivo*. Several CRISPR-based small molecule detection systems rely on combining ligand-responsive DNA aptamers or transcription factors with the collateral cleavage activity of Cas12a. In these systems, the presence of the cognate ligand reveals the Cas12a DNA target site that was originally occluded either by the aptamer secondary structure or transcription factor binding (141, 266). Upon interacting with small molecules, the target DNA of Cas12a is exposed and subsequently initiates collateral cleavage of a fluorophore quencher-labeled ssDNA probe. In a similar fashion, ligand-responsive transcription factors have also been used to induce transcription of Cas12a crRNA arrays and Cas13a target transcripts, followed by trans-cleavage of quenched probes (267, 268). For a more in depth review of Cas12 and Cas13 systems for *in vitro* small molecule assays, we point the reader to these recent reviews (269, 270).

Small molecules have also been detected through conditional recruitment of the effector to the CRISPR complex. Many ligand-induced heterodimerization domains have been used in eukaryotes for modulating CRISPRa activity by conditionally recruiting the activator (271). In prokaryotic systems, recent advancements in our understanding of CRISPRa target site requirements and improvements in the dynamic range of activatable

promoters have enabled conditional CRISPRa with abscisic acid (ABA)- and gibberellic acid (GA)-responsive heterodimerization domains (48). These domains were successfully used to build an AND-like logic gate and a CRISPRa cascade in a lysate-based CFES. Implementation of ligand-induced heterodimerization tools in bacteria should also be realizable depending on the effective ligand concentration in the cytosol. Interestingly, the ABA and GA inducible CRISPRa systems possessed similar design rules, suggesting that novel CRISPRa-based biosensors could be developed with other ligand-induced heterodimerization pairs. The repertoire of detectable small molecules could be further expanded through membrane-augmented CFES (272, 273) harboring TCS- and GPCR-coupled CRISPRa/i (274).

Aptamers can be incorporated into sgRNAs in order to stabilize or destabilize the sgRNA structure in response to ligands such as theophylline, 3-methylxanthine, and thiamine (215–217). The observed CRISPRi effect with these sgRNAs was concentration-dependent and ligand-specific, allowing simultaneous repression of two heterologous genes in *E. coli* using different ligands. Similar approaches, especially 5' extension and conditional Cas9 handle folding, can be used to generate ligand-responsive scRNAs for CRISPRa by incorporating existing or novel aptamers (Figure 5B). Combining our expertise in computational RNA design (275) and aptamer engineering (276), we successfully developed ligand-responsive scRNAs acting through a well-characterized theophylline aptamer and a novel human milk oligosaccharide (HMO) aptamer (in preparation). Since HMO was produced intracellularly by a CRISPR program, metabolite sensing could be coupled to the CRISPR circuit to dynamically regulate the metabolic pathway. Further improvement of computational design techniques capable of adhering

to sequence and structure constraints imposed by both aptamer and gRNA should accelerate construction of novel small molecule-responsive CRISPR circuitry.

Environmental cue sensing

For bioproduction, bioremediation, and biotherapeutic applications in which bacteria or CFES are deployed into changing environments, the ability to sense and respond to environmental cues is crucial for carrying out the desired function (277–279). The most general approach for responding to environmental cues relies on signal transduction modules, such as inducible promoters to conditionally express CRISPR components. In fact, in many bacterial genomes, CRISPR systems are already regulated by external stimuli like nutrient availability, stress factors, and quorum sensing (280–282). Successful implementation of engineered inducible systems requires precise tuning of the expression levels to minimize leaky expression while maintaining high enough inducible expression to carry out downstream functions (48, 283). In eukaryotes, many conditional CRISPR expression systems have been used to interface with endogenous and engineered regulatory networks and signaling pathways (284, 285). In bacteria, there are only a few examples of biosensing CRISPR circuits capable of responding to environmental stimuli, such as light and heat (286–289). Some notable environmental conditions capable of gene regulation in bacteria could also be coupled with CRISPR systems, including pH and oxygen levels (277, 290, 291) (Figure 5C). Recent advancements in inducible promoter engineering in bacteria and CFES should accelerate the design of other inducible systems with large output dynamic ranges that can be readily connected with downstream signal processing circuits (48, 292, 293).

As seen in this section, different CRISPR system components have been engineered to modulate CRISPR activity in response to diverse stimuli. Signal transduction modules based on gRNAs, through either aptamer fusions or expression from inducible promoters, are particularly interesting for more complex applications for two reasons. First, biosensing through gRNAs provides a streamlined and scalable framework for integrating multiple signals into one CRISPR circuit (48, 294, 295). And second, gRNA engineering approaches can be used to modulate the response level of each signal transduction module, potentially without engineering the module itself (45, 48). For instance, the leak from an inducible promoter module can be decreased by truncating the gRNA being expressed without modifications at the earlier nodes, such as mutations of the transcription factor protein or altering its expression level. In this fashion, many orthogonal signal transduction modules could be designed to interface with downstream CRISPR circuits capable of signal modulation, pattern formation, and computation. This framework highlights how CRISPR-based biosensors can not only be used for sensing different molecules and producing measurable readouts, but can also be readily engineered to launch any type of CRISPR program.

Figure 5: Forefront Applications of CRISPR Tools

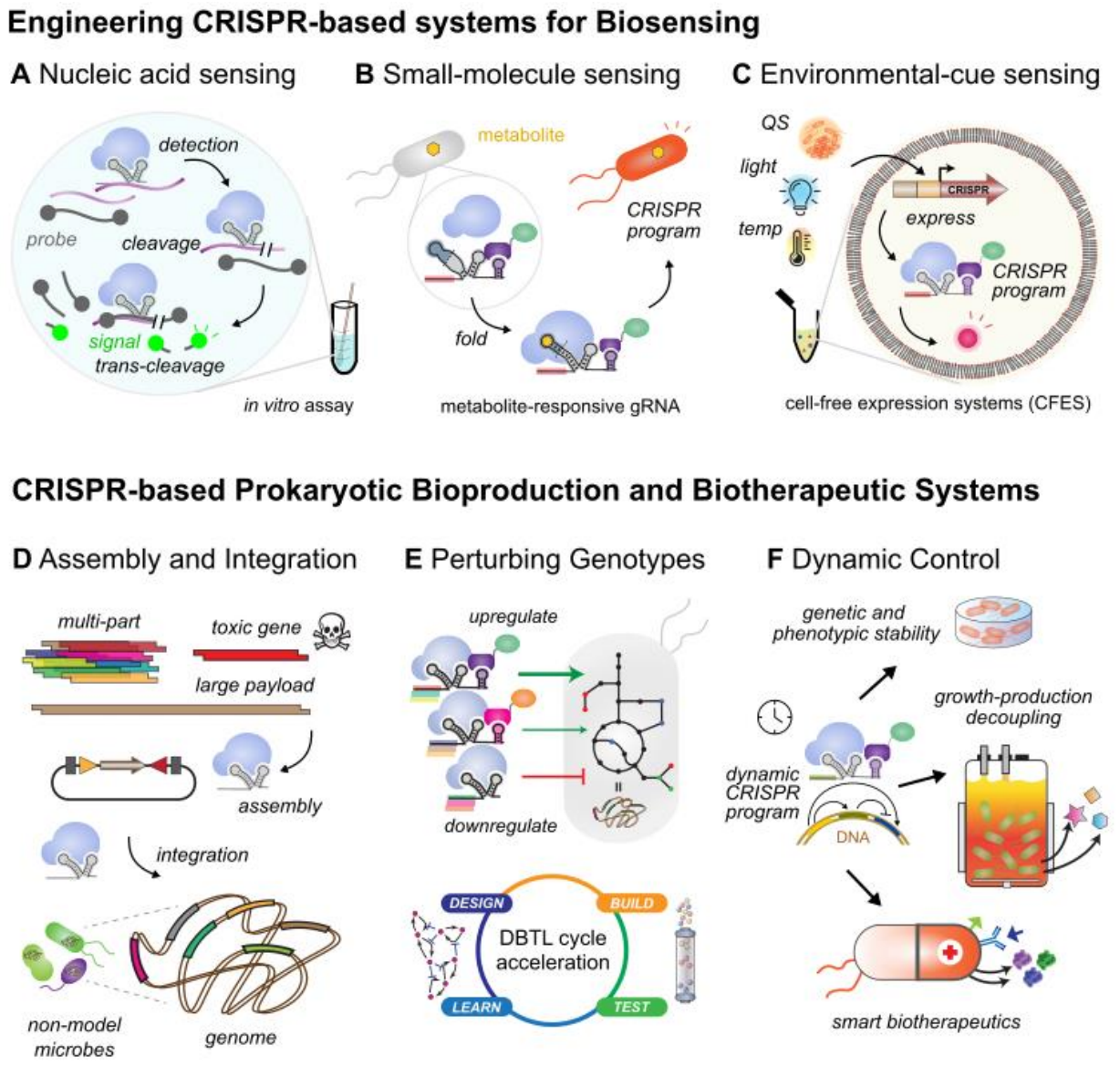


Figure 5: A) CRISPR systems engineered for detecting nucleic acids in *in vitro* assays. The gRNA spacer is designed to be complementary to the nucleic acid sequence, and the trans-cleavage activity of Cas12 or Cas13 is used to cleave the surrounding quenched fluorescent probes, leading to rapid fluorescent readout. B) Example of a small molecule detection *in vivo* scheme in which a metabolite-responsive gRNA triggers CRISPRa of a fluorescent gene. C) Schematic of a cell-free system harboring CRISPR components expressed from inducible promoters. Multiplexed response to different environmental

cues triggers expression of a fluorescent gene through CRISPRa. D) CRISPR tools for accelerating heterologous gene assembly and integration; particularly useful for assembly of large or toxic payload and delivery into non-model microbes lacking genetic tools. E) Schematic of genome-wide CRISPR perturbations to alter prokaryotic metabolism. Coupling gRNA programmability and multiplexibility with high-throughput measurements and computational models could accelerate design-build-test-learn (DBTL) cycles. F) Examples of dynamic CRISPR programs enabling improved genetic and phenotypic stability, dynamic bioproduction, and smart therapeutics capable of sensing specific microenvironments and launching the appropriate therapeutic action.

CRISPR Engineering of Prokaryotic Systems for Bioproduction and Biotherapeutics

CRISPR tools are accelerating the development of bacteria and cell-free expression systems (CFES) for bioproduction and biotherapeutic applications (29, 258, 296–299) (Figure 5, bottom). Engineering microbial chassis for programmable bioproductions and biotherapeutics often involves assembling multiple heterologous genes, fine-tuning their expression levels, and perturbing the endogenous metabolic pathways (300, 301). CRISPR systems streamline the development of microbial chassis by providing versatile toolboxes for manipulating multiple genes simultaneously during DNA assembly and strain engineering (27, 193). When optimizing strains, genome-wide CRISPR perturbations have proven to be a powerful tool for mapping the genotype-phenotype landscape, enabling discovery of new gene functions and selecting desirable bioproduction and biotherapeutic traits (198, 302). In lysate-based CFES, precise control over the enzymatic composition is usually required to maximize product yield (297). CRISPR tools can be used to alter lysate composition through strain engineering and to coordinate multi-enzyme programs through genetic circuits. Additionally, dynamic CRISPR programs enable microbes to adapt and respond to their environment–for instance, inducing lysing only when the biotherapeutic strain has colonized the tumor. When engineering strains for bioproduction, dynamic CRISPR programs can exert metabolic control to maximize flux to the desired product and minimize the accumulation of harmful intermediates. In this section, we focus on how CRISPR tools are enabling engineering complex prokaryotic systems for bioproduction and biotherapeutic applications.

Assembly and integration of engineered pathways

Engineering prokaryotic bioproduction and biotherapeutic platforms often involves manipulating and assembling several heterologous genes. The programmable targeting of CRISPR tools has been repurposed for molecular cloning and introduction of large heterologous pathways into attractive microbial chassis (303–305) (Figure 5D). The engineered near-PAMless SpRY Cas9 variant has been used as a versatile restriction enzyme for molecular cloning due to its programmable targeting of any DNA sequence (306). CRISPR nucleases have also been applied for both integration and counterselection steps and largely accelerated the genome engineering protocol, spanning from gene disruption (knock-out) to gene integration (knock-in) (307–310). For insertion of large genetic payload, several CRISPR systems were developed to enable rapid cloning of biosynthetic gene clusters (BGCs) of up to 145 kb, resulting in the discovery of novel natural products (303, 311–313). In the case of biosynthetic genes that exhibited toxicity in cloning hosts, CRISPRi could be used as an auxiliary system to suppress expression of harmful genes (314). Recent efforts developing CRISPR-assisted recombineering have led to improvements in multiplexed recombineering efficiency (315, 316), streamlining engineering of bioproducing *E. coli* (317–319).

CRISPR tools have also accelerated DNA manipulation in broad ranges of industrially-relevant microorganisms, which were challenging when done via conventional approaches (27, 193, 320). The broad applicability of CRISPR nuclease systems has accelerated genome integration pipelines in attractive microbes, such as *Pseudomonas* and *Acinetobacter* species, known for their broad applications in both chemical bioproduction and therapeutic discovery (307, 309, 310). Highly-efficient chassis-

independent recombinase-assisted genome engineering was employed to deliver CRISPR programs regulating BCGs in unconventional hosts (321). Similarity, recent CRISPR transposase studies have demonstrated programmable insertion of DNA into various bacterial hosts, including the Gram-negative Alphaproteobacteria (*Agrobacterium fabrum*), Betaproteobacteria (*Burkholderia thailandensis*), Gammaproteobacteria (*E. coli*, *P. putida*, and *Klebsiella oxytoca*), and Gram-positive bacteria (*B. subtilis* and *C. glutamicum*) (14, 15, 322, 323).

Perturbing genotypes to explore of phenotypic landscapes

Multi-gene CRISPR perturbations targeting endogenous and heterologous genes for editing or regulation can be used to explore bacterial phenotypes (198, 302). Bacterial phenotypes map ruggedly onto the genetic design space because the underlying metabolic and regulatory networks are complex (324, 325). In principle, combinatorial gene expression programs could be constructed to explore the genetic design space and identify desired phenotypes, such as high bioproduction titers (204, 326, 327). Compared to traditional *cis*-acting approaches, *trans*-acting CRISPR perturbations can dramatically reduce the time needed to engineer and fine-tune multi-gene programs because no additional direct engineering of the heterologous pathway or of the endogenous metabolism is necessary.

CRISPR perturbations include up- or down-regulating gene targets as well as editing, such as to introduce an early stop codon to the open reading frame or abolish enzyme function (167, 328) (Figure 5E). Both Cas9 and Cas12 gene editing have been implemented in metabolically versatile *Rhodobacter* species, leading to enhanced production of bioplastics (32, 329). CRISPR gene editing has also facilitated engineering

of photosynthetic cyanobacteria *Synechococcus* and *Anabaena* (330–332), and obligate anaerobes such as *Clostridia* species (333), demonstrating improved production. Importantly, multiple base-editing events can be instituted in parallel and sequentially, allowing for rapid genome-wide rewiring of the microbial metabolism (11, 31). For example, multiplexed base-editing in *R. sphaeroides* was shown to improve bioproduction of antioxidant coenzyme Q10 (31). Sequential base-editing has also been used to reconstruct *P. putida* phenotypes and introduce auxotrophy towards NADPH (11). Similarly, multiple CRISPRi repression have been implemented in various non-model organisms suitable for CO2-utilizing organisms for chemical productions, including cyanobacteria and acetogen (30, 334–336). For CRISPRa, various systems have been applied for bioproductions in *E. coli*, *P. putida*, *P. polymyxa*, *S. venezuelae*, and *M. xanthus*, enabling bioproduction of various compounds ranging from commodity chemicals—ethanol, butanediol, and mevalonate, to therapeutics—biopterins, jadomycin, epothilone (146, 150, 152, 160, 337). Notably, through CRISPRa of three heterologous genes, we have demonstrated production of biopterin in *E. coli* and *P. putida* (160, 161), and valuable oligosaccharide products in *E. coli* (in preparation). Crucial to this task was the ability to rapidly explore the bioproduction landscape by generating a combinatorial 64-member library of truncated scRNAs to upregulate expression of the three heterologous genes to precise levels (in preparation). Multiple studies have also demonstrated simultaneous CRISPRa and CRISPRi perturbations to achieve phenotypes inaccessible at the single gene perturbation level, across different prokaryotic systems, ranging from *E. coli*, *S. enterica*, *B. subtilis*, and *P. putida* (45, 146, 158, 160, 338).

Multiple groups have implemented genome-scale CRISPR programs for improved bioproduction. Coexpression of 20 sgRNAs to simultaneously downregulate six genes (*ackA, iclR, poxB, pta, sdhC, sdhD*) redirected *E. coli* metabolism to increase succinic acid production by 150-fold (246). Similar multiplexed CRISPRi programs have also led to increased production of isopentenol and free fatty acids (339, 340). Simultaneously CRISPR base editing of up to ten genes has been used to optimize lycopene production in *C. glutamicum*, accelerate glycerol utilization in *B. subtilis,* and increase PCA production in *P. putida* (11, 341). With high programmability and ease of design via short RNA sequences, CRISPR systems are permissive toward automated robotic platforms which could revolutionize the deployment of Design-Build-Test-Learn (DBTL) cycle (342). To achieve deeper understanding of novel engineered strains, genome-scale CRISPR programs could be reinforced by microbial single-cell RNA sequencing (343). This approach could enable massively-parallel investigations of individual CRISPR programs in a large gRNA library and feed comprehensive information back to the computational platform. In combination with advancements in massive DNA library synthesis, high-throughput sequencing, and metabolite-responsive biosensor, genome-wide screening for gene targets that improve productivity could be achieved in a single-cell level (344–346).

Similarly, CRISPR tools can be used to improve bioproduction in *in vitro* systems. Cell-free systems, ranging from purified enzymes (347–349) to lysate-based expression systems, have proven to be viable bioproduction platforms for a wide range of products, such as small molecules, vaccines, and even recombinant phages (297, 350–352). For more in-depth analyses on various cell-free bioproduction systems, we refer to other

reviews (353–355). Since lysate-based CFES maintain some of the cellular enzymes and metabolites (356, 357), the background reactions could divert flux away from heterologous pathways and decrease product yield. Proteomics and metabolomics lysate analysis have elucidated how the strain genotype, culture conditions, and lysate preparation, impact the final lysate composition and in turn production capacity of proteins and small molecules (356–359). Hence, recent efforts have engineered strains by knocking out, knocking down, or up-regulating target genes to produce a more favorable proteome for improved protein expression and redirecting metabolic flux in the resulting lysate (297, 360, 361). Taking advantage of the full suite of CRISPR tools would enable rapid strain engineering and screening of different lysate compositions for improved CFES bioproduction. Paired with novel lysates from non-model organisms with more desirable endogenous metabolism for bioproduction (362, 363), CRISPR-tools offer a promising route for engineering more efficient CFES bioproduction platforms.

CRISPR perturbations could also be used to identify genes with desirable traits for biotherapeutic applications. Biotherapeutic microbes, or live biotherapeutic products (LBP) as defined by the FDA, are becoming an exciting chassis for diagnostic and therapeutic applications (364). Recombinant LBPs are currently being explored to treat infectious diseases, inflammatory bowel diseases, and cancer (365). LBPs harboring CRISPR tools are already capable of recording changes in their environment and deploying novel antimicrobials (34, 35, 37, 366). One advantage of LBP is that they can colonize solid tumors and activate an immune response. However, bacteria are rapidly targeted and depleted by the immune system. Hence, CRISPRa/i could be used to

discover genes that affect tumor colonization and improve resistance to the host immune system.

CRISPR tools are undoubtedly accelerating our ability to explore bacterial phenotypes. CRISPRa/i at the transcriptional and translational level in particular, have proved to be powerful techniques for engineering metabolic pathways without altering the underlying genetics and regulation. While CRISPRa/i perturbations could be hardcoded into the final strain, there are advantages in having genetic programs to dynamically control phenotype, for instance to induce lysing only once a biotherapeutic strain is localized to the tumor, or to decouple cell growth and product synthesis in a bioproducing strain.

Exerting dynamic control over multiple gene

CRISPR tools can be used to dynamically sense multiple stimuli and respond to them through the coordinated expression of multiple genes (Figure 5F). In biotherapeutic applications, LBPs must sense characteristic features of their target microenvironment, such as low O_2 and high lactate levels when targeting tumors. CRISPR circuits could be readily designed to sense these microenvironment signals, process them, and launch the appropriate therapeutic action. Additionally, engineered LBPs should possess biocontainment circuits that enable both selective removal from the host and prevent proliferation in the wild (367, 368). Recently, Rottinghaus and colleagues demonstrated a CRISPR-based killswitch in *E. coli* Nissle strain for *in vivo* use (288). The system induces cell death through multiplexed CRISPRi targeting of essential genes in response to a chemical inducer as well as temperature.

In engineered biproduction systems, dynamic CRISPR control of metabolic pathways is a highly effective approach to maximize flux and minimize the accumulation of harmful intermediates, optimizing overall product yield (369). Recently, implementation of a CRISPRi dynamic feedback control increased the production of N-acetylglucosamine from 81.7 g/L to 131.6 g/L in *Bacillus subtilis* (370). The control circuit was designed to downregulate expression of competing pathways in response to the key intermediate GlcN6P. Crucial to the function of these circuits is the availability of pathway metabolite biosensors. Negative feedback control has also been used to reduce metabolic burden. sgRNAs expressed from native heat stress–related promoters allow the dynamic downregulation of heterologous genes, improving growth and achieving greater total protein production over time (197).

In addition to flux and burden control, decoupling cell growth and product synthesis is another approach to maximize overall production (371) . Towards this goal, quorum sensing (QS) has been used to monitor cell density in order to prioritize biomass accumulation early and product synthesis later. A QS-based CRISPRi circuit has been used in rapamycin producing *Streptomyces* for dynamic downregulation of genes in three competing pathways, TCA cycle, FA synthesis, and AAA synthesis. Notably, compared to static engineering strategies that result in growth arrest and low rapamycin titers, QS-based CRISPRi resulted in increased rapamycin titers to the highest reported level of 1836 mg/l (372). Similar strategies have been used to improve naringenin and salicylic acid production in engineered *E. coli* through QS-based CRISPRi of essential genes (373). Importantly, QS-based CRISPR circuits have also been shown to improve

bioproduction robustness across industrially-relevant culturing conditions (374), and enable dynamic programming of *E. coli–P. putida* consortium composition (375).

A parallel approach to decouple cell growth and product synthesis involves inducing the metabolic program during stationary phase. Expression of sgRNAs from stationary phase promoters enabled growth-to-production switching and high-titer shikimic acid and glutaric acid production (376). As with QS-based CRISPR circuits, induction of sgRNAs during stationary phase can also improve robustness and scalability of *E. coli* cultures producing alanine, citramalate and xylitol (377, 378). In the cyanobacteria *Synechocystis,* inducible CRISPRi of citrate synthase *gltA* at low culture densities increased CO_2 partitioning to n-butanol, but sacrificed volumetric productivity due to low cell growth (379). Similar effects have been observed in *P. putida* strains harboring knockouts or CRISPRi programs to boost acetyl-CoA levels (380, 381). In order to maintain high carbon partitioning into the desired product and increase overall CO_2 fixation rate, a control strategy to dynamically cycle between growth and production phases has been proposed . In lactate-secreting *Synechocystis,* intermittent CRISPRi growth arrest significantly increased cumulative lactate titers, and allowed stable lactate production for one month (382).

An emerging approach to maintain stable bioproduction genotypes and phenotypes during prolonged culture times involves encapsulating bioproduction microbes in biocompatible materials, forming engineered living materials (ELMs) (383). We have recently fabricated *E. coli* ELMs carrying dynamic CRISPR programs for long-term bioproduction of biopterin (161). These multi-gene CRISPR programs could be used to switch bioproduction on and off in several cycles for at least 19 days, indicating that

on-demand programming of ELMs could be signaled with CRISPR circuitry. With implementation of CRISPR circuit in the interface of material and living microbes, more complex programming, such as spatially-compartmentalized microbial consortia of multiple bioproduction modules (384–386), could be constructed.

In cell-free expression systems, dynamic CRISPRa/i circuits have also enabled precise temporal control of gene expression through delays, pulses, and filters (45, 46, 48). These CRISPRa/i circuits could provide an efficient mechanism for implementing dynamic multigene control programs for bioproduction applications. For instance, circuits capable of delaying gene expression could be useful when enzymes consume protein expression resources or produce toxic intermediates. In membrane-augmented CFES, dynamic CRISPRa/i circuits could be used to investigate and optimize how membrane enzyme expression impacts membrane insertion and function (272). In lysate engineering, if the desired proteome is detrimental to cell growth, dynamic CRISPRa/i programs could be used to switch into the desired proteome once high biomass levels have been achieved.

Diverse CRISPR tools are widely being adopted to accelerate the engineering of prokaryotic systems. Further coupling new CRISPR tools for rapid strain and lysate engineering with artificial intelligence and machine learning approaches (AI/ML) to recommend promising candidates (387–389), it could be possible to assess fewer genetic programs and achieve optimal strain in less time and with cheaper cost. Together with advancements in genome-scale modeling and computational gRNA design, these approaches will greatly accelerate the development of sophisticated bioproduction and biotherapeutic platforms.

Future prospects

CRISPR tools have proven invaluable for engineering prokaryotic systems across many fields, from biosensing to bioproduction and biotherapeutics. The ability of CRISPR tools to precisely modify genetic material and encoding new desired functions is unparalleled among traditional genetic engineering tools. The vast repertoire of different Cas proteins, tethered effector domains, and gRNA design strategies grants flexibility in engineering tools with different capabilities for specific tasks. While CRISPR transcriptional regulation and base editing are already well established, novel tools are being developed to engineer prokaryotic systems through other molecular processes, such as DNA looping, DNA replication, RNA translation, and protein degradation (159, 390–393).

Recently, dCas9-based CRISPR tools have been repurposed for modulating DNA replication. Plasmid copy-number manipulation in *E. coli* has been achieved by controlling centrations of DNA replication machinery (392, 393). Targeting the genomic DNA replication parts also provide programmable control of growth rate of individual bacteria in a co-culture of *E. coli* and *P. putida* (375). Implementation of similar strategies with broad-host-range CRISPRa tools could enable further replication control in non-model bacteria with a more limited set of different copy-numbers plasmids. We envision such tools could be integrated with dynamic CRISPR circuits to enable growth arrest for allocating metabolic resources towards bioproduction, or population control in a microbial consortia.

CRISPR-based translational regulation in prokaryotes has also been recently developed (159) and its potential is largely still unexplored. With the RNA-targeting

property of Cas13, translational control over a specific gene in a multi-gene operon could be possible. In *E. coli*, 68% of all genes are contained in operons, with some containing as many as 14 genes (394). Precise regulation of this significant number of genes with current dCas9/12-based tools is limited because transcription of all genes in the operon is affected. Hence, combining dCas13- and dCas9/12-based gene regulation would greatly expand our ability to precisely regulate genes within bacterial genomes.

In a similar fashion, Cas13-guided RNA editing has been demonstrated in mammalian cells by tethering the deaminase domain to dCas13 (138) but is yet to be implemented in prokaryotic systems. Given that some effectors of the DNA base-editors target RNA natively in bacteria (395), programmable prokaryotic Cas13-based RNA-editing systems should be within reach. Similar to CRISPR DNA-targeting tools, realizing the potential of CRISPR RNA-targeting in prokaryotes will require learning the rules to effectively and predictably repress, activate, and edit genes at the mRNA level.

The flexibility granted by CRISPR tools to localize different effector(s) to many target(s) has enabled the transition from a disarray of one-off molecular tools to a framework under which multiple tools can be orchestrated to engineer more complex systems (98, 396). Moving forward, simultaneous integration of multiple CRISPR tools in the diverse prokaryotic hosts should be immediately feasible considering that the distinct gRNA structures of Cas variants allow each CRISPR system to function orthogonally (26). The diverse catalog of Cas variants will provide suitable options for implementing CRISPR systems in any microbe with specific traits for each application. Given the number of orthogonal hairpin-RBP pairs, different effectors can be programmatically recruited to specific gRNAs, and therefore to specific nucleic acid targets. Harnessing the

designability of orthogonal CRISPR systems, further complex circuitry such as conditional gene activation (397), and synergistic gene activation (in preparation). Overall, harnessing the diversity of natural and engineered CRISPR-Cas systems will provide streamlined avenues for developing prokaryotic systems as the next generation of biosensors, dynamic bioproduction platforms, and smart microbial biotherapeutics.

Disclosure Statements

D.A.B., C.K., R.A.L.C., I.D.F., and J.M.C. are co-inventors on intellectual property filings that include CRISPR-related technologies. J.M.C is a member of the Wayfinder Biosciences scientific advisory board.

Acknowledgement

We thank members of the Carothers Research Group and members of the Zalatan Research Group for advice, materials, and comments on the manuscript. The authors' work in this area was supported in part by US NSF Award MCB 2032794 (J.M.C.), NSF Award CBET 1844152 (J.M.C.), NSF Award EF-1935087 (J.M.C.), US Department of Energy (DOE) BETO DE-EE0008927 (J.M.C.), DOE BER DE-SC0023091 (J.M.C.) and DOE ARPA-E DE-AR00002387-1567 (J.M.C). This material is based upon work supported by the NSF Graduate Research Fellowship Program under Grant No. DGE-2140004 (D.A.B., A.V.K.). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the NSF.

References

1. Nielsen J, Keasling JD. 2016. Engineering Cellular Metabolism. *Cell*. 164(6):1185–97

2. Voigt CA. 2020. Synthetic biology 2020–2030: six commercially-available products that are changing our world. *Nat. Commun.* 11(1):6379

3. Cameron DE, Bashor CJ, Collins JJ. 2014. A brief history of synthetic biology. *Nat. Rev. Microbiol.* 12(5):381–90

4. Flores-Bautista E, Hernandez-Guerrero R, Huerta-Saquero A, Tenorio-Salgado S, Rivera-Gomez N, et al. 2020. Deciphering the functional diversity of DNA-binding transcription factors in Bacteria and Archaea organisms. *PLOS ONE*. 15(8):e0237135

5. Porto EM, Komor AC, Slaymaker IM, Yeo GW. 2020. Base editing: advances and therapeutic opportunities. *Nat. Rev. Drug Discov.* 19(12):839–59

6. Wang H, Russa ML, Qi LS. 2016. CRISPR/Cas9 in Genome Editing and Beyond. *Annu. Rev. Biochem.* 85(1):227–64

7. Gaj T, Gersbach CA, Barbas CF III. 2013. ZFN, TALEN, and CRISPR/Cas-based methods for genome engineering. *Trends Biotechnol.* 31(7):397–405

8. Arroyo-Olarte RD, Bravo Rodríguez R, Morales-Ríos E. 2021. Genome Editing in Bacteria: CRISPR-Cas and Beyond. *Microorganisms*. 9(4):844

9. Vento JM, Crook N, Beisel CL. 2019. Barriers to genome editing with CRISPR in

bacteria. *J. Ind. Microbiol. Biotechnol.* 46(9–10):1327–41

10. Vigouroux A, Bikard D. 2020. CRISPR Tools To Control Gene Expression in Bacteria. *Microbiol. Mol. Biol. Rev.* 84(2):e00077-19

11. Volke DC, Martino RA, Kozaeva E, Smania AM, Nikel PI. 2022. Modular (de)construction of complex bacterial phenotypes by CRISPR/nCas9-assisted, multiplex cytidine base-editing. *Nat. Commun.* 13(1):3026

12. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. 2012. A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science*. 337(6096):816–21

13. Qi LS, Larson MH, Gilbert LA, Doudna JA, Weissman JS, et al. 2013. Repurposing CRISPR as an RNA-guided platform for sequence-specific control of gene expression. *Cell*. 152(5):1173–83

14. Tou CJ, Orr B, Kleinstiver BP. 2023. Precise cut-and-paste DNA insertion using engineered type V-K CRISPR-associated transposases. *Nat. Biotechnol.*

15. Vo PLH, Ronda C, Klompe SE, Chen EE, Acree C, et al. 2021. CRISPR RNA-guided integrases for high-efficiency, multiplexed bacterial genome engineering. *Nat. Biotechnol.* 39(4):480–89

16. Komor AC, Kim YB, Packer MS, Zuris JA, Liu DR. 2016. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature*. 533(7603):420–24

17. Nishida K, Arazoe T, Yachie N, Banno S, Kakimoto M, et al. 2016. Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science*. 353(6305):aaf8729

18. Anzalone AV, Koblan LW, Liu DR. 2020. Genome editing with CRISPR–Cas nucleases, base editors, transposases and prime editors. *Nat. Biotechnol.* 38(7):824–44

19. Gilbert LA, Horlbeck MA, Adamson B, Villalta JE, Chen Y, et al. 2014. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. *Cell*. 159(3):647–61

20. Rock JM, Hopkins FF, Chavez A, Diallo M, Chase MR, et al. 2017. Programmable transcriptional repression in mycobacteria using an orthogonal CRISPR interference platform. *Nat. Microbiol.* 2(4):16274

21. Fonfara I, Richter H, Bratovič M, Le Rhun A, Charpentier E. 2016. The CRISPR-associated DNA-cleaving enzyme Cpf1 also processes precursor CRISPR RNA. *Nature*. 532(7600):517–21

22. Zetsche B, Gootenberg JS, Abudayyeh OO, Slaymaker IM, Makarova KS, et al. 2015. Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*. 163(3):759–71

23. Casas-Mollano JA, Zinselmeier MH, Sychla A, Smanski MJ. 2023. Efficient gene activation in plants by the MoonTag programmable transcriptional activator. *Nucleic Acids Res.*, p. gkad458

24. Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, et al. 2015. Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. *Cell*. 160(1–2):339–50

25. Fontana J, Dong C, Kiattisewee C, Chavali VP, Tickman BI, et al. 2020. Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. *Nat. Commun.* 11(1):1618

26. Fontana J, Sparkman-Yager D, Zalatan JG, Carothers JM. 2020. Challenges and opportunities with CRISPR activation in bacteria for data-driven metabolic engineering. *Curr. Opin. Biotechnol.* 64:190–98

27. Call SN, Andrews LB. 2022. CRISPR-Based Approaches for Gene Regulation in Non-Model Bacteria. *Front. Genome Ed.* 4:

28. Cho S, Shin J, Cho B-K. 2018. Applications of CRISPR/Cas System to Bacterial Metabolic Engineering. *Int. J. Mol. Sci.* 19(4):

29. Shi S, Qi N, Nielsen J. 2022. Microbial production of chemicals driven by CRISPR-Cas systems. *Curr. Opin. Biotechnol.* 73:34–42

30. Fackler N, Heffernan J, Juminaga A, Doser D, Nagaraju S, et al. 2021. Transcriptional control of Clostridium autoethanogenum using CRISPRi. *Synth. Biol.* 6(1):ysab008

31. Luo Y, Ge M, Wang B, Sun C, Wang J, et al. 2020. CRISPR/Cas9-deaminase enables robust base editing in Rhodobacter sphaeroides 2.4.1. *Microb. Cell*

Factories. 19(1):93

32. Mougiakos I, Orsi E, Ghiffary MR, Post W, de Maria A, et al. 2019. Efficient Cas9-based genome editing of Rhodobacter sphaeroides for metabolic engineering. *Microb. Cell Factories*. 18(1):204

33. Xiong B, Li Z, Liu L, Zhao D, Zhang X, Bi C. 2018. Genome editing of Ralstonia eutropha using an electroporation-based CRISPR-Cas9 technique. *Biotechnol. Biofuels*. 11(1):172

34. Bhattarai-Kline S, Lear SK, Fishman CB, Lopez SC, Lockshin ER, et al. 2022. Recording gene expression order in DNA by CRISPR addition of retron barcodes. *Nature*. 608(7921):217–25

35. Jiao C, Reckstadt C, König F, Homberger C, Yu J, et al. 2023. RNA recording in single bacterial cells using reprogrammed tracrRNAs. *Nat. Biotechnol.* 41(8):1107–16

36. Kiga K, Tan X-E, Ibarra-Chávez R, Watanabe S, Aiba Y, et al. 2020. Development of CRISPR-Cas13a-based antimicrobials capable of sequence-specific killing of target bacteria. *Nat. Commun.* 11(1):2934

37. Rodrigues M, McBride SW, Hullahalli K, Palmer KL, Duerkop BA. 2019. Conjugative Delivery of CRISPR-Cas9 for the Selective Depletion of Antibiotic-Resistant Enterococci

38. Schmidt R, Steinhart Z, Layeghi M, Freimer JW, Bueno R, et al. 2022. CRISPR

activation and interference screens decode stimulation responses in primary human T cells

39. Kaminski MM, Abudayyeh OO, Gootenberg JS, Zhang F, Collins JJ. 2021. CRISPR-based diagnostics. *Nat. Biomed. Eng.* 5(7):643–56

40. Tian T, Zhou X. 2023. CRISPR-Based Biosensing Strategies: Technical Development and Application Prospects. *Annu. Rev. Anal. Chem.* 16(1):311–32

41. Garamella J, Marshall R, Rustad M, Noireaux V. 2016. The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. *ACS Synth. Biol.* 5(4):344–55

42. Garenne D, Thompson S, Brisson A, Khakimzhan A, Noireaux V. 2021. The all-E. coliTXTL toolbox 3.0: new capabilities of a cell-free synthetic biology platform. *Synth. Biol.* 6(1):ysab017

43. Lavickova B, Maerkl SJ. 2019. A Simple, Robust, and Low-Cost Method To Produce the PURE Cell-Free System. *ACS Synth. Biol.* 8(2):455–62

44. Marshall R, Maxwell CS, Collins SP, Jacobsen T, Luo ML, et al. 2018. Rapid and Scalable Characterization of CRISPR Technologies Using an E. coli Cell-Free Transcription-Translation System. *Mol. Cell*. 69(1):146-157.e3

45. Tickman BI, Burbano DA, Chavali VP, Kiattisewee C, Fontana J, et al. 2021. Multi-layer CRISPRa/i circuits for dynamic genetic programs in cell-free and bacterial systems. *Cell Syst.*, p. S2405471221004191

46. Westbrook A, Tang X, Marshall R, Maxwell CS, Chappell J, et al. 2019. Distinct timescales of RNA regulators enable the construction of a genetic pulse generator. *Biotechnol. Bioeng.* 116(5):1139–51

47. Wimmer F, Mougiakos I, Englert F, Beisel CL. 2022. Rapid cell-free characterization of multi-subunit CRISPR effectors and transposons. *Mol. Cell*. 82(6):1210-1224.e6

48. Alba Burbano D, Cardiff RAL, Tickman BI, Kiattisewee C, Maranas CJ, et al. 2023. Engineering activatable promoters for scalable and multi-input CRISPRa/i circuits. *Proc. Natl. Acad. Sci.* 120(30):e2220358120

49. Cui Y, Chen X, Wang Z, Lu Y. Cell-Free PURE System: Evolution and Achievements. *BioDesign Res.* 2022:9847014

50. Karim AS, Liew F (Eric), Garg S, Vögeli B, Rasor BJ, et al. 2020. Modular cell-free expression plasmids to accelerate biological design in cells. *Synth. Biol.* 5(1):ysaa019

51. Silverman AD, Karim AS, Jewett MC. 2020. Cell-free gene expression: an expanded repertoire of applications. *Nat. Rev. Genet.* 21(3):151–70

52. Brandt K, Barrangou R. 2019. Applications of CRISPR Technologies Across the Food Supply Chain. *Annu. Rev. Food Sci. Technol.* 10(1):133–50

53. McCarty NS, Graham AE, Studená L, Ledesma-Amaro R. 2020. Multiplexed CRISPR technologies for gene editing and transcriptional regulation. *Nat.*

1294 *Commun.* 11(1):1281

54. Simeonov DR, Marson A. 2019. CRISPR-Based Tools in Immunity. *Annu. Rev.*
1296 *Immunol.* 37(1):571–97

55. Hille F, Richter H, Wong SP, Bratovič M, Ressel S, Charpentier E. 2018. The
1298 Biology of CRISPR-Cas: Backward and Forward. *Cell*. 172(6):1239–59

56. Anders C, Niewoehner O, Duerst A, Jinek M. 2014. Structural basis of PAM-
1300 dependent target DNA recognition by the Cas9 endonuclease. *Nature*.
513(7519):569–73

1302 57. Sternberg SH, Redding S, Jinek M, Greene EC, Doudna JA. 2014. DNA
interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*.
1304 507(7490):62–67

58. Anderson EM, Haupt A, Schiel JA, Chou E, Machado HB, et al. 2015. Systematic
1306 analysis of CRISPR-Cas9 mismatch tolerance reveals low levels of off-target
activity. *J. Biotechnol.* 211:56–65

1308 59. Chylinski K, Le Rhun A, Charpentier E. 2013. The tracrRNA and Cas9 families of
type II CRISPR-Cas immunity systems. *RNA Biol.* 10(5):726–37

1310 60. Fonfara I, Le Rhun A, Chylinski K, Makarova KS, Lécrivain A-L, et al. 2014.
Phylogeny of Cas9 determines functional exchangeability of dual-RNA and Cas9
1312 among orthologous type II CRISPR-Cas systems. *Nucleic Acids Res.* 42(4):2577–
90

61. Chatterjee P, Lee J, Nip L, Koseki SRT, Tysinger E, et al. 2020. A Cas9 with PAM recognition for adenine dinucleotides. *Nat. Commun.* 11(1):2474

62. Chatterjee P, Jakimo N, Jacobson JM. 2018. Minimal PAM specificity of a highly similar SpCas9 ortholog. *Sci. Adv.* 4(10):eaau0766

63. Horvath P, Romero DA, Coûté-Monvoisin A-C, Richards M, Deveau H, et al. 2008. Diversity, Activity, and Evolution of CRISPR Loci in *Streptococcus thermophilus*

64. Müller M, Lee CM, Gasiunas G, Davis TH, Cradick TJ, et al. 2016. Streptococcus thermophilus CRISPR-Cas9 Systems Enable Specific Editing of the Human Genome. *Mol. Ther.* 24(3):636–44

65. Rybnicky GA, Fackler NA, Karim AS, Köpke M, Jewett MC. 2022. Spacer2PAM: A computational framework to guide experimental determination of functional CRISPR-Cas system PAM sequences. *Nucleic Acids Res.* 50(6):3523–34

66. Mendoza BJ, Trinh CT. 2018. In Silico Processing of the Complete CRISPR-Cas Spacer Space for Identification of PAM Sequences. *Biotechnol. J.* 13(9):1700595

67. Collias D, Beisel CL. 2021. CRISPR technologies and the search for the PAM-free nuclease. *Nat. Commun.* 12(1):555

68. Hu JH, Miller SM, Geurts MH, Tang W, Chen L, et al. 2018. Evolved Cas9 variants with broad PAM compatibility and high DNA specificity. *Nature*. 556(7699):57–63

69. Nishimasu H, Shi X, Ishiguro S, Gao L, Hirano S, et al. 2018. Engineered CRISPR-Cas9 nuclease with expanded targeting space. *Science*. 361(6408):1259–62

70. Edraki A, Mir A, Ibraheim R, Gainetdinov I, Yoon Y, et al. 2019. A Compact, High-Accuracy Cas9 with a Dinucleotide PAM for In Vivo Genome Editing. *Mol. Cell*. 73(4):714-726.e4

71. Ran FA, Cong L, Yan WX, Scott DA, Gootenberg JS, et al. 2015. In vivo genome editing using Staphylococcus aureus Cas9. *Nature*. 520(7546):186–91

72. Schmidheini L, Mathis N, Marquart KF, Rothgangl T, Kissling L, et al. 2023. Continuous directed evolution of a compact CjCas9 variant with broad PAM compatibility. *Nat. Chem. Biol.*

73. Hu Z, Wang S, Zhang C, Gao N, Li M, et al. 2020. A compact Cas9 ortholog from Staphylococcus Auricularis (SauriCas9) expands the DNA targeting scope. *PLOS Biol.* 18(3):e3000686

74. Gao L, Cox DBT, Yan WX, Manteiga JC, Schneider MW, et al. 2017. Engineered Cpf1 variants with altered PAM specificities. *Nat. Biotechnol.* 35(8):789–92

75. Kleinstiver BP, Sousa AA, Walton RT, Tak YE, Hsu JY, et al. 2019. Engineered CRISPR–Cas12a variants with increased activities and improved targeting ranges for gene, epigenetic and base editing. *Nat. Biotechnol.* 37(3):276–82

76. Strecker J, Jones S, Koopal B, Schmid-Burgk J, Zetsche B, et al. 2019.

Engineering of CRISPR-Cas12b for human genome editing. *Nat. Commun.* 10(1):212

77. Huang CJ, Adler BA, Doudna JA. 2022. A naturally DNase-free CRISPR-Cas12c enzyme silences gene expression. *Mol. Cell*. 82(11):2148-2160.e4

78. Wu WY, Mohanraju P, Liao C, Adiego-Pérez B, Creutzburg SCA, et al. 2022. The miniature CRISPR-Cas12m effector binds DNA to block transcription. *Mol. Cell*. 82(23):4487-4502.e7

79. Burstein D, Harrington LB, Strutt SC, Probst AJ, Anantharaman K, et al. 2017. New CRISPR–Cas systems from uncultivated microbes. *Nature*. 542(7640):237–41

80. Liu J-J, Orlova N, Oakes BL, Ma E, Spinner HB, et al. 2019. CasX enzymes comprise a distinct family of RNA-guided genome editors. *Nature*. 566(7743):218–23

81. Karvelis T, Bigelyte G, Young JK, Hou Z, Zedaveinyte R, et al. 2020. PAM recognition by miniature CRISPR–Cas12f nucleases triggers programmable double-stranded DNA target cleavage. *Nucleic Acids Res.* 48(9):5016–23

82. Xu X, Chemparathy A, Zeng L, Kempton HR, Shang S, et al. 2021. Engineered miniature CRISPR-Cas system for mammalian genome regulation and editing. *Mol. Cell*. 81(20):4333-4345.e4

83. Sharrar A, Tacca LA de, Collingwood T, Meacham Z, Rabuka D, et al. 2023.

1374 μCas, a novel class of miniature type-V Cas12f nucleases with diverse PAM

84. Li Z, Zhang H, Xiao R, Han R, Chang L. 2021. Cryo-EM structure of the RNA-
1376 guided ribonuclease Cas12g. *Nat. Chem. Biol.* 17(4):387–93

85. Yan WX, Hunnewell P, Alfonse LE, Carte JM, Keston-Smith E, et al. 2019.
1378 Functionally diverse type V CRISPR-Cas systems. *Science*. 363(6422):88–91

86. Pausch P, Al-Shayeb B, Bisom-Rapp E, Tsuchida CA, Li Z, et al. 2020. CRISPR-
1380 CasΦ from huge phages is a hypercompact genome editor

87. Wang Y, Qi T, Liu J, Yang Y, Wang Z, et al. 2023. A highly specific CRISPR-
1382 Cas12j nuclease enables allele-specific genome editing

88. Strecker J, Ladha A, Gardner Z, Schmid-Burgk JL, Makarova KS, et al. 2019.
1384 RNA-guided DNA insertion with CRISPR-associated transposases

89. Sun A, Li C-P, Chen Z, Zhang S, Li D-Y, et al. 2023. The compact Casπ (Cas12l)
1386 'bracelet' provides a unique structural platform for DNA manipulation. *Cell Res.* 33(3):229–44

1388 90. Urbaitis T, Gasiunas G, Young JK, Hou Z, Paulraj S, et al. 2022. A new family of CRISPR-type V nucleases with C-rich PAM recognition. *EMBO Rep.*
1390 23(12):e55481

91. Chen W, Ma J, Wu Z, Wang Z, Zhang H, et al. 2023. Cas12n nucleases, early
1392 evolutionary intermediates of type V CRISPR, comprise a distinct family of miniature genome editors. *Mol. Cell*. 83(15):2768-2780.e6

92. Abudayyeh OO, Gootenberg JS, Essletzbichler P, Han S, Joung J, et al. 2017. RNA targeting with CRISPR–Cas13. *Nature*. 550(7675):280–84

93. Smargon AA, Cox DBT, Pyzocha NK, Zheng K, Slaymaker IM, et al. 2017. Cas13b Is a Type VI-B CRISPR-Associated RNA-Guided RNase Differentially Regulated by Accessory Proteins Csx27 and Csx28. *Mol. Cell*. 65(4):618-630.e7

94. Pyzocha NK, Chen S. 2018. Diverse Class 2 CRISPR-Cas Effector Proteins for Genome Engineering Applications. *ACS Chem. Biol.* 13(2):347–56

95. Yan WX, Chong S, Zhang H, Makarova KS, Koonin EV, et al. 2018. Cas13d Is a Compact RNA-Targeting Type VI CRISPR Effector Positively Modulated by a WYL-Domain-Containing Accessory Protein. *Mol. Cell*. 70(2):327-339.e5

96. Xu C, Zhou Y, Xiao Q, He B, Geng G, et al. 2021. Programmable RNA editing with compact CRISPR–Cas13 systems from uncultivated microbes. *Nat. Methods*. 18(5):499–506

97. Kleinstiver BP, Prew MS, Tsai SQ, Topkar VV, Nguyen NT, et al. 2015. Engineered CRISPR-Cas9 nucleases with altered PAM specificities. *Nature*. 523(7561):481–85

98. Legut M, Daniloski Z, Xue X, McKenzie D, Guo X, et al. 2020. High-Throughput Screens of PAM-Flexible Cas9 Variants for Gene Knockout and Transcriptional Modulation. *Cell Rep.* 30(9):2859-2868.e5

99. Walton RT, Christie KA, Whittaker MN, Kleinstiver BP. 2020. Unconstrained

genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science*, p. eaba8853

100. Ma D, Xu Z, Zhang Z, Chen X, Zeng X, et al. 2019. Engineer chimeric Cas9 to expand PAM recognition based on evolutionary information. *Nat. Commun.* 10(1):560

101. Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, et al. 2014. Crystal Structure of Cas9 in Complex with Guide RNA and Target DNA. *Cell*. 156(5):935–49

102. Teng Y, Wang J, Jiang T, Zou Y, Yan Y. 2023. Engineering a Streptococcus Cas9 Ortholog with an RxQ PAM-Binding Motif for PAM-Free Gene Control in Bacteria. *ACS Synth. Biol.* 12(9):2764–72

103. Kim N, Kim HK, Lee S, Seo JH, Choi JW, et al. 2020. Prediction of the sequence-specific cleavage activity of Cas9 variants. *Nat. Biotechnol.* 38(11):1328–36

104. Kiattisewee C, Karanjia AV, Legut M, Daniloski Z, Koplik SE, et al. 2022. Expanding the Scope of Bacterial CRISPR Activation with PAM-Flexible dCas9 Variants. *ACS Synth. Biol.* 11(12):4103–12

105. Klanschnig M, Cserjan-Puschmann M, Striedner G, Grabherr R. 2022. CRISPRactivation-SMS, a message for PAM sequence independent gene up-regulation in Escherichia coli. *Nucleic Acids Res.*, p. gkac804

106. Jones DL, Leroy P, Unoson C, Fange D, Ćurić V, et al. 2017. Kinetics of dCas9

target search in *Escherichia coli*

107. Casini A, Olivieri M, Petris G, Montagna C, Reginato G, et al. 2018. A highly specific SpCas9 variant is identified by in vivo screening in yeast. *Nat. Biotechnol.* 36(3):265–71

108. Chen JS, Dagdas YS, Kleinstiver BP, Welch MM, Sousa AA, et al. 2017. Enhanced proofreading governs CRISPR–Cas9 targeting accuracy. *Nature*. 550(7676):407–10

109. Kleinstiver BP, Pattanayak V, Prew MS, Tsai SQ, Nguyen NT, et al. 2016. High-fidelity CRISPR–Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*. 529(7587):490–95

110. Kulcsár PI, Tálas A, Huszár K, Ligeti Z, Tóth E, et al. 2017. Crossing enhanced and high fidelity SpCas9 nucleases to optimize specificity and cleavage. *Genome Biol.* 18(1):190

111. Lee JK, Jeong E, Lee J, Jung M, Shin E, et al. 2018. Directed evolution of CRISPR-Cas9 to increase its specificity. *Nat. Commun.* 9(1):3048

112. Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX, Zhang F. 2016. Rationally engineered Cas9 nucleases with improved specificity

113. Kim D, Bae S, Park J, Kim E, Kim S, et al. 2015. Digenome-seq: genome-wide profiling of CRISPR-Cas9 off-target effects in human cells. *Nat. Methods*. 12(3):237–43

114. Cho S, Choe D, Lee E, Kim SC, Palsson B, Cho B-K. 2018. High-Level dCas9 Expression Induces Abnormal Cell Morphology in Escherichia coli. *ACS Synth. Biol.* 7(4):1085–94

115. Zhang S, Voigt CA. 2018. Engineered dCas9 with reduced toxicity in bacteria: implications for genetic circuit design. *Nucleic Acids Res.* 46(20):11115–25

116. Adler BA, Trinidad MI, Bellieny-Rabelo D, Zhang E, Karp HM, et al. 2023. CasPEDIA Database: A Functional Classification System for Class 2 CRISPR-Cas Enzymes

117. Swarts DC, van der Oost J, Jinek M. 2017. Structural Basis for Guide RNA Processing and Seed-Dependent DNA Targeting by CRISPR-Cas12a. *Mol. Cell*. 66(2):221-233.e4

118. Shmakov S, Abudayyeh OO, Makarova KS, Wolf YI, Gootenberg JS, et al. 2015. Discovery and Functional Characterization of Diverse Class 2 CRISPR-Cas Systems. *Mol. Cell*. 60(3):385–97

119. Koonin EV, Makarova KS, Zhang F. 2017. Diversity, classification and evolution of CRISPR-Cas systems. *Environ. Microbiol. CRISPRcas9*. 37:67–78

120. Makarova KS, Wolf YI, Iranzo J, Shmakov SA, Alkhnbashi OS, et al. 2020. Evolutionary classification of CRISPR–Cas systems: a burst of class 2 and derived variants. *Nat. Rev. Microbiol.* 18(2):67–83

121. Chen JS, Ma E, Harrington LB, Da Costa M, Tian X, et al. 2018. CRISPR-Cas12a

target binding unleashes indiscriminate single-stranded DNase activity

122. Al-Shayeb B, Skopintsev P, Soczek KM, Stahl EC, Li Z, et al. 2022. Diverse virus-encoded CRISPR-Cas systems include streamlined genome editors. *Cell*. 185(24):4574-4586.e16

123. Wörle E, Jakob L, Schmidbauer A, Zinner G, Grohmann D. 2021. Decoupling the bridge helix of Cas12a results in a reduced trimming activity, increased mismatch sensitivity and impaired conformational transitions. *Nucleic Acids Res.* 49(9):5278–93

124. Ma E, Chen K, Shi H, Stahl EC, Adler B, et al. 2022. Improved genome editing by an engineered CRISPR-Cas12a. *Nucleic Acids Res.* 50(22):12689–701

125. Tóth E, Varga É, Kulcsár PI, Kocsis-Jutka V, Krausz SL, et al. 2020. Improved LbCas12a variants with altered PAM specificities further broaden the genome targeting range of Cas12a nucleases. *Nucleic Acids Res.* 48(7):3722–33

126. Nguyen LT, Rananaware SR, Yang LG, Macaluso NC, Ocana-Ortiz JE, et al. 2023. Engineering highly thermostable Cas12b via de novo structural analyses for one-pot detection of nucleic acids. *Cell Rep. Med.* 4(5):101037

127. O'Connell MR. 2019. Molecular Mechanisms of RNA Targeting by Cas13-containing Type VI CRISPR–Cas Systems. *CRISPR Basic Biol. Its Technol. Appl.* 431(1):66–87

128. Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, et al.

2016. C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. *Science*. 353(6299):aaf5573

129. Meeske AJ, Nakandakari-Higa S, Marraffini LA. 2019. Cas13-induced cellular dormancy prevents the rise of CRISPR-resistant bacteriophage. *Nature*. 570(7760):241–45

130. Bandaru S, Tsuji MH, Shimizu Y, Usami K, Lee S, et al. 2020. Structure-based design of gRNA for Cas13. *Sci. Rep.* 10(1):11610

131. East-Seletsky A, O'Connell MR, Burstein D, Knott GJ, Doudna JA. 2017. RNA Targeting by Functionally Orthogonal Type VI-A CRISPR-Cas Enzymes. *Mol. Cell*. 66(3):373-383.e3

132. East-Seletsky A, O'Connell MR, Knight SC, Burstein D, Cate JHD, et al. 2016. Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. *Nature*. 538(7624):270–73

133. Konermann S, Lotfy P, Brideau NJ, Oki J, Shokhirev MN, Hsu PD. 2018. Transcriptome Engineering with RNA-Targeting Type VI-D CRISPR Effectors. *Cell*. 173(3):665-676.e14

134. Charles EJ, Kim SE, Knott GJ, Smock D, Doudna J, Savage DF. 2021. Engineering improved Cas13 effectors for targeted post-transcriptional regulation of gene expression

135. Kellner MJ, Koob JG, Gootenberg JS, Abudayyeh OO, Zhang F. 2019.

SHERLOCK: nucleic acid detection with CRISPR nucleases. *Nat. Protoc.* 14(10):2986–3012

136. Adler BA, Hessler T, Cress BF, Lahiri A, Mutalik VK, et al. 2022. Broad-spectrum CRISPR-Cas13a enables efficient phage genome editing. *Nat. Microbiol.* 7(12):1967–79

137. Borrajo J, Javanmardi K, Griffin J, Martin SJS, Yao D, et al. 2023. Programmable multi-kilobase RNA editing using CRISPR-mediated trans-splicing

138. Cox DBT, Gootenberg JS, Abudayyeh OO, Franklin B, Kellner MJ, et al. 2017. RNA editing with CRISPR-Cas13. *Science*. 358(6366):1019–27

139. Guan J, Oromí-Bosch A, Mendoza SD, Karambelkar S, Berry JD, Bondy-Denomy J. 2022. Bacteriophage genome engineering with CRISPR–Cas13a. *Nat. Microbiol.* 7(12):1956–66

140. Qu L, Yi Z, Zhu S, Wang C, Cao Z, et al. 2019. Programmable RNA editing by recruiting endogenous ADAR using engineered RNAs. *Nat. Biotechnol.* 37(9):1059–69

141. Liang M, Li Z, Wang W, Liu J, Liu L, et al. 2019. A CRISPR-Cas12a-derived biosensing platform for the highly sensitive detection of diverse small molecules. *Nat. Commun.* 10(1):3672

142. Hildebrand F, Meyer A, Eyre-Walker A. 2010. Evidence of Selection upon Genomic GC-Content in Bacteria. *PLOS Genet.* 6(9):e1001107

143. Maeder ML, Linder SJ, Cascio VM, Fu Y, Ho QH, Joung JK. 2013. CRISPR RNA–guided activation of endogenous human genes. *Nat. Methods*. 10(10):977–79

144. Perez-Pinera P, Kocak DD, Vockley CM, Adler AF, Kabadi AM, et al. 2013. RNA-guided gene activation by CRISPR-Cas9–based transcription factors. *Nat. Methods*. 10(10):973–76

145. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA. 2013. Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. *Nucleic Acids Res*. 41(15):7429–37

146. Dong C, Fontana J, Patel A, Carothers JM, Zalatan JG. 2018. Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. *Nat Commun*. 9(1):2489

147. Liu Y, Wan X, Wang B. 2019. Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. *Nat. Commun.* 10(1):3693

148. Ho H-I, Fang JR, Cheung J, Wang HH. 2020. Programmable CRISPR-Cas transcriptional activation in bacteria. *Mol. Syst. Biol.* 16(7):e9427

149. Villegas Kcam MC, Tsong AJ, Chappell J. 2021. Rational engineering of a modular bacterial CRISPR–Cas activation platform with expanded target range. *Nucleic Acids Res.* 49(8):4793–4802

150. Ameruoso A, Villegas Kcam MC, Cohen KP, Chappell J. 2022. Activating natural product synthesis using CRISPR interference and activation systems in

Streptomyces

151. Liu W, Tang D, Wang H, Lian J, Huang L, Xu Z. 2019. Combined genome editing and transcriptional repression for metabolic pathway engineering in Corynebacterium glutamicum using a catalytically active Cas12a. *Appl. Microbiol. Biotechnol.* 103(21):8911–22

152. Schilling C, Koffas MAG, Sieber V, Schmid J. 2020. Novel Prokaryotic CRISPR-Cas12a-Based Tool for Programmable Transcriptional Activation and Repression. *ACS Synth. Biol.* 9(12):3353–63

153. Villegas Kcam MC, Tsong AJ, Chappell J. 2022. Uncovering the Distinct Properties of a Bacterial Type I-E CRISPR Activation System. *ACS Synth. Biol.* 11(2):1000–1003

154. Lu Z, Yang S, Yuan X, Shi Y, Ouyang L, et al. 2019. CRISPR-assisted multi-dimensional regulation for fine-tuning gene expression in Bacillus subtilis. *Nucleic Acids Res.* 47(7):e40–e40

155. Peng R, Wang Y, Feng W, Yue X, Chen J, et al. 2018. CRISPR/dCas9-mediated transcriptional improvement of the biosynthetic gene cluster for the epothilone production in Myxococcus xanthus. *Microb. Cell Factories*. 17(1):15

156. Yu L, Su W, Fey PD, Liu F, Du L. 2018. Yield Improvement of the Anti-MRSA Antibiotics WAP-8294A by CRISPR/dCas9 Combined with Refactoring Self-Protection Genes in Lysobacter enzymogenes OH11. *ACS Synth. Biol.* 7(1):258–66

157. Chen Y, Niu X, Cheng M, Wang L, Sun P, et al. 2022. CRISPR/dCas9-RpoD-Mediated Simultaneous Transcriptional Activation and Repression in Shewanella oneidensis MR-1. *ACS Synth. Biol.* 11(6):2184–92

158. Wu Y, Liu Y, Lv X, Li J, Du G, Liu L. 2020. CAMERS-B: CRISPR/Cpf1 assisted multiple-genes editing and regulation system for Bacillus subtilis. *Biotechnol Bioeng*. 117(6):1817–25

159. Otoupal PB, Cress BF, Doudna JA, Schoeniger JS. 2022. CRISPR-RNAa: targeted activation of translation using dCas13 fusions to translation initiation factors. *Nucleic Acids Res.* 50(15):8986–98

160. Kiattisewee C, Dong C, Fontana J, Sugianto W, Peralta-Yahya P, et al. 2021. Portable bacterial CRISPR transcriptional activation enables metabolic engineering in Pseudomonas putida. *Metab. Eng.* 66:283–95

161. Sugianto W, Altin-Yavuzarslan G, Tickman BI, Kiattisewee C, Yuan S-F, et al. 2023. Gene expression dynamics in input-responsive engineered living materials programmed for bioproduction. *Mater. Today Bio*. 20:100677

162. Liu Y, Pinto F, Wan X, Yang Z, Peng S, et al. 2022. Reprogrammed tracrRNAs enable repurposing of RNAs as crRNAs and sequence-specific RNA biosensors. *Nat. Commun.* 13(1):1937

163. Javaid N, Choi S. 2021. CRISPR/Cas System and Factors Affecting Its Precision and Efficiency. *Front. Cell Dev. Biol.* 9:

164. Jiang W, Bikard D, Cox D, Zhang F, Marraffini LA. 2013. CRISPR-assisted editing of bacterial genomes. *Nat. Biotechnol.* 31(3):233–39

165. Vercoe RB, Chang JT, Dy RL, Taylor C, Gristwood T, et al. 2013. Cytotoxic Chromosomal Targeting by CRISPR/Cas Systems Can Reshape Bacterial Genomes and Expel or Remodel Pathogenicity Islands. *PLOS Genet.* 9(4):e1003454

166. Wang L, Xue W, Yan L, Li X, Wei J, et al. 2017. Enhanced base editing by co-expression of free uracil DNA glycosylase inhibitor. *Cell Res.* 27(10):1289–92

167. Wang Y, Liu Y, Zheng P, Sun J, Wang M. 2021. Microbial Base Editing: A Powerful Emerging Technology for Microbial Genome Engineering. *Trends Biotechnol.* 39(2):165–80

168. Zhang S, Feng S, Jiang W, Huang X, Chen J. 2019. Construction and optimization of a base editor based on the MS2 system. *Anim. Models Exp. Med.* 2(3):185–90

169. Banno S, Nishida K, Arazoe T, Mitsunobu H, Kondo A. 2018. Deaminase-mediated multiplex genome editing in Escherichia coli. *Nat. Microbiol.* 3(4):423–29

170. Heo YB, Hwang G-H, Kang SW, Bae S, Woo HM. 2022. High-Fidelity Cytosine Base Editing in a GC-Rich Corynebacterium glutamicum with Reduced DNA Off-Target Editing Effects

171. Kim MS, Kim H-R, Jeong D-E, Choi S-K. 2021. Cytosine Base Editor-Mediated

Multiplex Genome Editing to Accelerate Discovery of Novel Antibiotics in Bacillus subtilis and Paenibacillus polymyxa. *Front. Microbiol.* 12:691839

172. Yu S, Price MA, Wang Y, Liu Y, Guo Y, et al. 2020. CRISPR-dCas9 Mediated Cytosine Deaminase Base Editing in Bacillus subtilis. *ACS Synth. Biol.* 9(7):1781–89

173. Zheng K, Wang Y, Li N, Jiang F-F, Wu C-X, et al. 2018. Highly efficient base editing in bacteria using a Cas9-cytidine deaminase fusion. *Commun. Biol.* 1:32

174. Sun J, Lu L-B, Liang T-X, Yang L-R, Wu J-P. 2020. CRISPR-Assisted Multiplex Base Editing System in Pseudomonas putida KT2440. *Front. Bioeng. Biotechnol.* 8:905

175. Li X, Wang Y, Liu Y, Yang B, Wang X, et al. 2018. Base editing with a Cpf1-cytidine deaminase fusion. *Nat. Biotechnol.* 36(4):324–27

176. Zhao Y, Tian J, Zheng G, Chen J, Sun C, et al. 2020. Multiplex genome editing using a dCas9-cytidine deaminase fusion in Streptomyces. *Sci. China Life Sci.* 63(7):1053–62

177. Rodrigues SD, Karimi M, Impens L, Van Lerberge E, Coussens G, et al. 2021. Efficient CRISPR-mediated base editing in Agrobacterium spp. *Proc. Natl. Acad. Sci. U. S. A.* 118(2):e2013338118

178. Xin X, Li J, Zhao D, Li S, Xie Q, et al. 2019. Double-Check Base Editing for Efficient A to G Conversions. *ACS Synth. Biol.* 8(12):2629–34

179. Zhang Y, Zhang H, Wang Z, Wu Z, Wang Y, et al. 2020. Programmable adenine deamination in bacteria using a Cas9–adenine-deaminase fusion. *Chem. Sci.* 11(6):1657–64

180. Abdullah, Wang P, Han T, Liu W, Ren W, et al. 2022. Adenine Base Editing System for Pseudomonas and Prediction Workflow for Protein Dysfunction via ABE. *ACS Synth. Biol.* 11(4):1650–57

181. Halperin SO, Tou CJ, Wong EB, Modavi C, Schaffer DV, Dueber JE. 2018. CRISPR-guided DNA polymerases enable diversification of all nucleotides in a tunable window. *Nature*. 560(7717):248–52

182. Tong Y, Jørgensen TS, Whitford CM, Weber T, Lee SY. 2021. A versatile genetic engineering toolkit for E. coli based on CRISPR-prime editing. *Nat. Commun.* 12(1):5206

183. Yarnall MTN, Ioannidi EI, Schmitt-Ulms C, Krajeski RN, Lim J, et al. 2023. Drag-and-drop genome insertion of large sequences without double-strand DNA cleavage using CRISPR-directed integrases. *Nat. Biotechnol.* 41(4):500–512

184. Dandage R, Després PC, Yachie N, Landry CR. 2019. beditor: A Computational Workflow for Designing Libraries of Guide RNAs for CRISPR-Mediated Base Editing. *Genetics*. 212(2):377–85

185. Hsu JY, Grünewald J, Szalay R, Shih J, Anzalone AV, et al. 2021. PrimeDesign software for rapid and simplified design of prime editing guide RNAs. *Nat. Commun.* 12(1):1034

186. Hwang G-H, Park J, Lim K, Kim S, Yu J, et al. 2018. Web-based design and analysis tools for CRISPR base editing. *BMC Bioinformatics*. 19(1):542

187. Standage-Beier K, Tekel SJ, Brafman DA, Wang X. 2021. Prime Editing Guide RNA Design Automation Using PINE-CONE. *ACS Synth. Biol.* 10(2):422–27

188. Blin K, Shaw S, Tong Y, Weber T. 2020. Designing sgRNAs for CRISPR-BEST base editing applications with CRISPy-web 2.0. *Synth. Syst. Biotechnol.* 5(2):99–102

189. Shelake RM, Pramanik D, Kim J-Y. 2022. In Vivo Rapid Investigation of CRISPR-Based Base Editing Components in Escherichia coli (IRI-CCE): A Platform for Evaluating Base Editing Tools and Their Components. *Int. J. Mol. Sci.* 23(3):1145

190. Chen Y, Hysolli E, Chen A, Casper S, Liu S, et al. 2022. Multiplex base editing to convert TAG into TAA codons in the human genome. *Nat. Commun.* 13(1):4482

191. Kluesner MG, Nedveck DA, Lahr WS, Garbe JR, Abrahante JE, et al. 2018. EditR: A Method to Quantify Base Editing from Sanger Sequencing. *CRISPR J.* 1(3):239–50

192. Rieffer AE, Chen Y, Salamango DJ, Moraes SN, Harris RS. 2023. APOBEC Reporter Systems for Evaluating diNucleotide Editing Levels. *CRISPR J.*

193. Lu L, Shen X, Sun X, Yan Y, Wang J, Yuan Q. 2022. CRISPR-based metabolic engineering in non-model microorganisms. *Curr. Opin. Biotechnol.* 75:102698

194. Naduthodi MIS, Barbosa MJ, van der Oost J. 2018. Progress of CRISPR-Cas

Based Genome Editing in Photosynthetic Microbes. *Biotechnol. J.* 13(9):1700591

195. Volke DC, Orsi E, Nikel PI. 2023. Emergent CRISPR–Cas-based technologies for engineering non-model bacteria. *Curr. Opin. Microbiol.* 75:102353

196. Bellato M, Frusteri Chiacchiera A, Salibi E, Casanova M, De Marchi D, et al. 2022. CRISPR Interference Modules as Low-Burden Logic Inverters in Synthetic Circuits. *Front. Bioeng. Biotechnol.* 9:

197. Ceroni F, Boo A, Furini S, Gorochowski TE, Borkowski O, et al. 2018. Burden-driven feedback control of gene expression. *Nat. Methods*. 15(5):387–93

198. Lian J, HamediRad M, Hu S, Zhao H. 2017. Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system. *Nat. Commun.* 8(1):1688

199. Shaytan AK, Novikov RV, Vinnikov RS, Gribkova AK, Glukhov GS. 2022. From DNA-protein interactions to the genetic circuit design using CRISPR-dCas systems. *Front. Mol. Biosci.* 9:

200. Filippova J, Matveeva A, Zhuravlev E, Stepanov G. 2019. Guide RNA modification as a way to improve CRISPR/Cas9-based genome-editing systems. *Biochimie*. 167:49–60

201. Moon SB, Kim DY, Ko J-H, Kim J-S, Kim Y-S. 2019. Improving CRISPR Genome Editing by Engineering Guide RNAs. *Trends Biotechnol.* 37(8):870–81

202. Todor H, Silvis MR, Osadnik H, Gross CA. 2021. Bacterial CRISPR screens for gene function. *Curr. Opin. Microbiol.* 59:102–9

203. Hawkins JS, Silvis MR, Koo B-M, Peters JM, Osadnik H, et al. 2020. Mismatch-CRISPRi Reveals the Co-varying Expression-Fitness Relationships of Essential Genes in Escherichia coli and Bacillus subtilis. *Cell Syst.* 11(5):523-535.e9

204. Jost M, Santos DA, Saunders RA, Horlbeck MA, Hawkins JS, et al. 2020. Titrating gene expression using libraries of systematically attenuated CRISPR guide RNAs. *Nat. Biotechnol.* 38(3):355–64

205. Van Hove B, De Wannemaeker L, Missiaen I, Maertens J, De Mey M. 2023. Taming CRISPRi: Dynamic range tuning through guide RNA diversion. *New Biotechnol.* 77:50–57

206. Singh D, Sternberg SH, Fei J, Doudna JA, Ha T. 2016. Real-time observation of DNA recognition and rejection by the RNA-guided endonuclease Cas9. *Nat. Commun.* 7(1):12778

207. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. 2014. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. *Nat. Biotechnol.* 32(3):279–84

208. Tsai SQ, Zheng Z, Nguyen NT, Liebers M, Topkar VV, et al. 2015. GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases. *Nat. Biotechnol.* 33(2):187–97

209. Tsai SQ, Joung JK. 2016. Defining and improving the genome-wide specificities of CRISPR–Cas9 nucleases. *Nat. Rev. Genet.* 17(5):300–312

210. Kocak DD, Josephs EA, Bhandarkar V, Adkar SS, Kwon JB, Gersbach CA. 2019.

Increasing the specificity of CRISPR systems with engineered RNA secondary structures. *Nat. Biotechnol.* 37(6):657–66

211. Mullally G, van Aelst K, Naqvi MM, Diffin FM, Karvelis T, et al. 2020. 5′ modifications to CRISPR–Cas9 gRNA can change the dynamics and size of R-loops and inhibit DNA cleavage. *Nucleic Acids Res.* 48(12):6811–23

212. Li Y, Teng X, Zhang K, Deng R, Li J. 2019. RNA Strand Displacement Responsive CRISPR/Cas9 System for mRNA Sensing. *Anal. Chem.* 91(6):3989–96

213. Lin J, Liu Y, Lai P, Ye H, Xu L. 2020. Conditional guide RNA through two intermediate hairpins for programmable CRISPR/Cas9 function: building regulatory connections between endogenous RNA expressions. *Nucleic Acids Res.* 48(20):11773–84

214. Siu K-H, Chen W. 2019. Riboregulated toehold-gated gRNA for programmable CRISPR–Cas9 function. *Nat. Chem. Biol.* 15(3):217–20

215. Tang W, Hu JH, Liu DR. 2017. Aptazyme-embedded guide RNAs enable ligand-responsive genome editing and transcriptional activation. *Nat. Commun.* 8(1):15939

216. Kundert K, Lucas JE, Watters KE, Fellmann C, Ng AH, et al. 2019. Controlling CRISPR-Cas9 with ligand-activated and ligand-deactivated sgRNAs. *Nat. Commun.* 10(1):2127

217. Liu Y, Zhan Y, Chen Z, He A, Li J, et al. 2016. Directing cellular information flow via CRISPR signal conductors. *Nat. Methods*. 13(11):938–44

218. Collins SP, Rostain W, Liao C, Beisel CL. 2021. Sequence-independent RNA sensing and DNA targeting by a split domain CRISPR–Cas12a gRNA switch. *Nucleic Acids Res.* 49(5):2985–99

219. Oesinghaus L, Simmel FC. 2019. Switching the activity of Cas12a using guide RNA strand displacement circuits. *Nat. Commun.* 10(1):2092

220. Wang Y, Wang Y, Pan D, Yu H, Zhang Y, et al. 2022. Guide RNA engineering enables efficient CRISPR editing with a miniature Syntrophomonas palmitatica Cas12f1 nuclease. *Cell Rep.* 40(13):

221. Liu G, Zhang Y, Zhang T. 2020. Computational approaches for effective CRISPR guide RNA design and evaluation. *Comput. Struct. Biotechnol. J.* 18:35–44

222. Corsi GI, Qu K, Alkan F, Pan X, Luo Y, Gorodkin J. 2022. CRISPR/Cas9 gRNA activity depends on free energy changes and on the target PAM context. *Nat. Commun.* 13(1):3006

223. Jensen KT, Fløe L, Petersen TS, Huang J, Xu F, et al. 2017. Chromatin accessibility and guide sequence secondary structure affect CRISPR-Cas9 gene editing efficiency. *FEBS Lett.* 591(13):1892–1901

224. Weiss T, Crisp PA, Rai KM, Song M, Springer NM, Zhang F. 2022. Epigenetic features drastically impact CRISPR–Cas9 efficacy in plants. *Plant Physiol.*

190(2):1153–64

225. Doench JG, Fusi N, Sullender M, Hegde M, Vaimberg EW, et al. 2016. Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* 34(2):184–91

226. Moreno-Mateos MA, Vejnar CE, Beaudoin J-D, Fernandez JP, Mis EK, et al. 2015. CRISPRscan: designing highly efficient sgRNAs for CRISPR-Cas9 targeting in vivo. *Nat. Methods*. 12(10):982–88

227. Wang L, Zhang J. 2019. Prediction of sgRNA on-target activity in bacteria by deep learning. *BMC Bioinformatics*. 20(1):517

228. Yu Y, Gawlitt S, Sousa LB de A e, Merdivan E, Piraud M, et al. 2023. Improved prediction of bacterial CRISPRi guide efficiency from depletion screens through mixed-effect modeling and data integration

229. Elowitz MB, Leibler S. 2000. A synthetic oscillatory network of transcriptional regulators. *Nature*. 403(6767):335–38

230. Gardner TS, Cantor CR, Collins JJ. 2000. Construction of a genetic toggle switch in Escherichia coli. *Nature*. 403(6767):339–42

231. English MA, Gayet RV, Collins JJ. 2021. Designing Biological Circuits: Synthetic Biology Within the Operon Model and Beyond. *Annu. Rev. Biochem.* 90(1):221–44

232. Hasty J, McMillen D, Collins JJ. 2002. Engineered gene circuits. *Nature*.

420(6912):224–30

233. Jeong D, Klocke M, Agarwal S, Kim J, Choi S, et al. 2019. Cell-Free Synthetic Biology Platform for Engineering Synthetic Biological Circuits and Systems. *Methods Protoc.* 2(2):

234. Brophy JAN, Voigt CA. 2014. Principles of genetic circuit design. *Nat. Methods*. 11(5):508–20

235. Gander MW, Vrana JD, Voje WE, Carothers JM, Klavins E. 2017. Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. *Nat. Commun.* 8(1):15459

236. Lucks JB, Qi L, Whitaker WR, Arkin AP. 2008. Toward scalable parts families for predictable design of biological circuits. *Curr. Opin. Microbiol.* 11(6):567–73

237. Qian Y, Huang H-H, Jiménez JI, Del Vecchio D. 2017. Resource Competition Shapes the Response of Genetic Circuits. *ACS Synth. Biol.* 6(7):1263–72

238. Jin M, Garreau de Loubresse N, Kim Y, Kim J, Yin P. 2019. Programmable CRISPR-Cas Repression, Activation, and Computation with Sequence-Independent Targets and Triggers. *ACS Synth. Biol.* 8(7):1583–89

239. Santos-Moreno J, Tasiudi E, Stelling J, Schaerli Y. 2020. Multistable and dynamic CRISPRi-based synthetic circuits. *Nat. Commun.* 11(1):2746

240. Kim SK, Kim H, Woo SG, Kim TH, Rha E, et al. 2022. CRISPRi-based programmable logic inverter cascade for antibiotic-free selection and maintenance of multiple plasmids. *Nucleic Acids Res.* 50(22):13155–71

241. Clamons S, Murray R. 2019. Modeling predicts that CRISPR-based activators, unlike CRISPR-based repressors, scale well with increasing gRNA competition and dCas9 bottlenecking. *bioRxiv*, p. 719278

242. Khakimzhan A, Garenne D, Tickman B, Fontana J, Carothers J, Noireaux V. 2021. Complex dependence of CRISPR-Cas9 binding strength on guide RNA spacer lengths. *Phys. Biol.* 18(5):056003

243. Banerjee D, Eng T, Lau AK, Sasaki Y, Wang B, et al. 2020. Genome-scale metabolic rewiring improves titers rates and yields of the non-native product indigoidine at scale. *Nat. Commun.* 11(1):5385

244. McCarty NS, Shaw WM, Ellis T, Ledesma-Amaro R. 2019. Rapid Assembly of gRNA Arrays via Modular Cloning in Yeast. *ACS Synth. Biol.* 8(4):906–10

245. Liao C, Ttofali F, Slotkowski RA, Denny SR, Cecil TD, et al. 2019. Modular one-pot assembly of CRISPR arrays enables library generation and reveals factors influencing crRNA biogenesis. *Nat. Commun.* 10(1):2948

246. Reis AC, Halper SM, Vezeau GE, Cetnar DP, Hossain A, et al. 2019. Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays. *Nat. Biotechnol.* 37(11):1294–1301

247. Martynov A, Severinov K, Ispolatov I. 2017. Optimal number of spacers in CRISPR arrays. *PLOS Comput. Biol.* 13(12):e1005891

248. Huang H-H, Bellato M, Qian Y, Cárdenas P, Pasotti L, et al. 2021. dCas9

regulator to neutralize competition in CRISPRi circuits. *Nat. Commun.* 12(1):1692

249. Broughton JP, Deng X, Yu G, Fasching CL, Servellita V, et al. 2020. CRISPR–Cas12-based detection of SARS-CoV-2. *Nat. Biotechnol.* 38(7):870–74

250. Myhrvold C, Freije CA, Gootenberg JS, Abudayyeh OO, Metsky HC, et al. 2018. Field-deployable viral diagnostics using CRISPR-Cas13

251. Qiu M, Zhou X-M, Liu L. 2022. Improved Strategies for CRISPR-Cas12-based Nucleic Acids Detection. *J. Anal. Test.* 6(1):44–52

252. Gootenberg JS, Abudayyeh OO, Kellner MJ, Joung J, Collins JJ, Zhang F. 2018. Multiplexed and portable nucleic acid detection platform with Cas13, Cas12a, and Csm6

253. Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, et al. 2017. Nucleic acid detection with CRISPR-Cas13a/C2c2

254. Li S-Y, Cheng Q-X, Wang J-M, Li X-Y, Zhang Z-L, et al. 2018. CRISPR-Cas12a-assisted nucleic acid detection. *Cell Discov.* 4(1):20

255. Patchsung M, Jantarug K, Pattama A, Aphicho K, Suraritdechachai S, et al. 2020. Clinical validation of a Cas13-based assay for the detection of SARS-CoV-2 RNA. *Nat. Biomed. Eng.* 4(12):1140–49

256. Nguyen PQ, Soenksen LR, Donghia NM, Angenent-Mari NM, de Puig H, et al. 2021. Wearable materials with embedded synthetic biology sensors for biomolecule detection. *Nat. Biotechnol.* 39(11):1366–74

257. Zhang C, Zhang P, Ren H, Jia P, Ji J, et al. 2022. Synthetic biology-powered biosensors based on CRISPR/Cas mediated cascade signal amplification for precise RNA detection. *Chem. Eng. J.* 446:136864

258. Mayorga-Ramos A, Zúñiga-Miranda J, Carrera-Pacheco SE, Barba-Ostria C, Guamán LP. 2023. CRISPR-Cas-Based Antimicrobials: Design, Challenges, and Bacterial Mechanisms of Resistance. *ACS Infect. Dis.* 9(7):1283–1302

259. Uribe RV, Rathmer C, Jahn LJ, Ellabaan MMH, Li SS, Sommer MOA. 2021. Bacterial resistance to CRISPR-Cas antimicrobials. *Sci. Rep.* 11(1):17267

260. Bikard D, Euler CW, Jiang W, Nussenzweig PM, Goldberg GW, et al. 2014. Exploiting CRISPR-Cas nucleases to produce sequence-specific antimicrobials. *Nat. Biotechnol.* 32(11):1146–50

261. Cobb LH, Park J, Swanson EA, Beard MC, McCabe EM, et al. 2019. CRISPR-Cas9 modified bacteriophage for treatment of Staphylococcus aureus induced osteomyelitis and soft tissue infection. *PLOS ONE*. 14(11):e0220421

262. Neil K, Allard N, Roy P, Grenier F, Menendez A, et al. 2021. High-efficiency delivery of CRISPR-Cas9 by engineered probiotics enables precise microbiome editing. *Mol. Syst. Biol.* 17(10):e10335

263. Ram G, Ross HF, Novick RP, Rodriguez-Pagan I, Jiang D. 2018. Conversion of staphylococcal pathogenicity islands to CRISPR-carrying antibacterial agents that cure infections in mice. *Nat. Biotechnol.* 36(10):971–76

264. Wan X, Li Q, Olsen RH, Meng H, Zhang Z, et al. 2022. Engineering a CRISPR interference system targeting AcrAB-TolC efflux pump to prevent multidrug resistance development in Escherichia coli. *J. Antimicrob. Chemother.* 77(8):2158–66

265. Jiao C, Sharma S, Dugar G, Peeck NL, Bischler T, et al. 2021. Noncanonical crRNAs derived from host transcripts enable multiplexable RNA detection by Cas9

266. Dai Y, Somoza RA, Wang L, Welter JF, Li Y, et al. 2019. Exploring the Trans-Cleavage Activity of CRISPR-Cas12a (cpf1) for the Development of a Universal Electrochemical Biosensor. *Angew. Chem. Int. Ed.* 58(48):17399–405

267. Iwasaki RS, Batey RT. 2020. SPRINT: a Cas13a-based platform for detection of small molecules. *Nucleic Acids Res.* 48(17):e101–e101

268. Mahas A, Wang Q, Marsic T, Mahfouz MM. 2022. Development of Cas12a-Based Cell-Free Small-Molecule Biosensors via Allosteric Regulation of CRISPR Array Expression. *Anal. Chem.* 94(11):4617–26

269. Chen K, Shen Z, Wang G, Gu W, Zhao S, et al. 2022. Research progress of CRISPR-based biosensors and bioassays for molecular diagnosis. *Front. Bioeng. Biotechnol.* 10:

270. Kadam US, Cho Y, Park TY, Hong JC. 2023. Aptamer-based CRISPR-Cas powered diagnostics of diverse biomarkers and small molecule targets. *Appl. Biol. Chem.* 66(1):13

271. Cunningham-Bryant D, Sun J, Fernandez B, Zalatan JG. 2019. CRISPR–Cas-Mediated Chemical Control of Transcriptional Dynamics in Yeast. *ChemBioChem*. 20(12):1519–23

272. Kruyer NS, Sugianto W, Tickman BI, Alba Burbano D, Noireaux V, et al. 2021. Membrane Augmented Cell-Free Systems: A New Frontier in Biotechnology. *ACS Synth. Biol.* 10(4):670–81

273. Sachse R, Dondapati SK, Fenz SF, Schmidt T, Kubick S. 2014. Membrane protein synthesis in cell-free systems: From bio-mimetic systems to bio-membranes. *FEBS Lett.* 588(17):2774–81

274. Kipniss NH, Dingal PCDP, Abbott TR, Gao Y, Wang H, et al. 2017. Engineering cell sensing and responses using a GPCR-coupled CRISPR-Cas system. *Nat. Commun.* 8(1):2212

275. Burke CR, Sparkman-Yager D, Carothers JM. 2017. Multi--state design of kinetically--controlled RNA aptamer ribosensors. *bioRxiv*

276. Hwang C, Carothers JM. 2016. Label-free selection of RNA aptamers for metabolic engineering. *Vitro Sel. Evol.* 106:37–41

277. Chien T, Harimoto T, Kepecs B, Gray K, Coker C, et al. 2022. Enhancing the tropism of bacteria via genetically programmed biosensors. *Nat. Biomed. Eng.* 6(1):94–104

278. Jung JK, Alam KK, Verosloff MS, Capdevila DA, Desmau M, et al. 2020. Cell-free

biosensors for rapid detection of water contaminants. *Nat. Biotechnol.* 38(12):1451–59

279. Karbelkar AA, Reynolds EE, Ahlmark R, Furst AL. 2021. A Microbial Electrochemical Technology to Detect and Degrade Organophosphate Pesticides. *ACS Cent. Sci.* 7(10):1718–27

280. Patterson AG, Jackson SA, Taylor C, Evans GB, Salmond GPC, et al. 2016. Quorum Sensing Controls Adaptive Immunity through the Regulation of Multiple CRISPR-Cas Systems. *Mol. Cell*. 64(6):1102–8

281. Patterson AG, Chang JT, Taylor C, Fineran PC. 2015. Regulation of the Type I-F CRISPR-Cas system by CRP-cAMP and GalM controls spacer acquisition and interference. *Nucleic Acids Res.* 43(12):6038–48

282. Perez-Rodriguez R, Haitjema C, Huang Q, Nam KH, Bernardis S, et al. 2011. Envelope stress is a trigger of CRISPR RNA-mediated DNA silencing in Escherichia coli. *Mol. Microbiol.* 79(3):584–99

283. Nielsen AA, Voigt CA. 2014. Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. *Mol. Syst. Biol.* 10(11):763

284. Baeumler TA, Ahmed AA, Fulga TA. 2017. Engineering Synthetic Signaling Pathways with Programmable dCas9-Based Chimeric Receptors. *Cell Rep.* 20(11):2639–53

285. Shao J, Wang M, Yu G, Zhu S, Yu Y, et al. 2018. Synthetic far-red light-mediated

CRISPR-dCas9 device for inducing functional neuronal differentiation

286. Jiang F, Liu J-J, Osuna BA, Xu M, Berry JD, et al. 2019. Temperature-Responsive Competitive Inhibition of CRISPR-Cas9. *Mol. Cell*. 73(3):601-610.e5

287. Richter F, Fonfara I, Bouazza B, Schumacher CH, Bratovič M, et al. 2016. Engineering of temperature- and light-switchable Cas9 variants. *Nucleic Acids Res.* 44(20):gkw930

288. Rottinghaus AG, Ferreiro A, Fishbein SRS, Dantas G, Moon TS. 2022. Genetically stable CRISPR-based kill switches for engineered microbes. *Nat. Commun.* 13(1):672

289. Wu P, Chen Y, Liu M, Xiao G, Yuan J. 2021. Engineering an Optogenetic CRISPRi Platform for Improved Chemical Production. *ACS Synth. Biol.* 10(1):125–31

290. Moser F, Espah Borujeni A, Ghodasara AN, Cameron E, Park Y, Voigt CA. 2018. Dynamic control of endogenous metabolism with combinatorial logic circuits. *Mol. Syst. Biol.* 14(11):e8605

291. Stirling F, Naydich A, Bramante J, Barocio R, Certo M, et al. 2020. Synthetic Cassettes for pH-Mediated Sensing, Counting, and Containment. *Cell Rep.* 30(9):3139-3148.e4

292. 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(1):64

293. Liu X, Gupta STP, Bhimsaria D, Reed JL, Rodríguez-Martínez JA, et al. 2019. De novo design of programmable inducible promoters. *Nucleic Acids Res.* 47(19):10452–63

294. Fontana J, Dong C, Ham JY, Zalatan JG, Carothers JM. 2018. Regulated Expression of sgRNAs Tunes CRISPRi in E. coli. *Biotechnol J*. 13(9):e1800069

295. Shaw WM, Studená L, Roy K, Hapeta P, McCarty NS, et al. 2022. Inducible expression of large gRNA arrays for multiplexed CRISPRai applications. *Nat. Commun.* 13(1):4984

296. Ramachandran G, Bikard D. 2019. Editing the microbiome the CRISPR way. *Philos. Trans. R. Soc. B Biol. Sci.* 374(1772):20180103

297. Rasor BJ, Yi X, Brown H, Alper HS, Jewett MC. 2021. An integrated in vivo/in vitro framework to enhance cell-free biosynthesis with metabolically rewired yeast extracts. *Nat. Commun.* 12(1):5139

298. Roberts A, Barrangou R. 2020. Applications of CRISPR-Cas systems in lactic acid bacteria. *FEMS Microbiol. Rev.* 44(5):523–37

299. Teng Y, Jiang T, Yan Y. 2023. The expanded CRISPR toolbox for constructing microbial cell factories. *Trends Biotechnol.*

300. Fatma Z, Schultz JC, Zhao H. 2020. Recent advances in domesticating non-model microorganisms. *Biotechnol. Prog.* 36(5):e3008

301. Yan Q, Fong SS. 2017. Challenges and Advances for Genetic Engineering of Non-model Bacteria and Uses in Consolidated Bioprocessing. *Front. Microbiol.* 8(2060):

302. Lian J, Schultz C, Cao M, HamediRad M, Zhao H. 2019. Multi-functional genome-wide CRISPR system for high throughput genotype–phenotype mapping. *Nat. Commun.* 10(1):5794

303. Jiang W, Zhao X, Gabrieli T, Lou C, Ebenstein Y, Zhu TF. 2015. Cas9-Assisted Targeting of CHromosome segments CATCH enables one-step targeted cloning of large gene clusters. *Nat. Commun.* 6(1):8101

304. Kang H-S, Charlop-Powers Z, Brady SF. 2016. Multiplexed CRISPR/Cas9- and TAR-Mediated Promoter Engineering of Natural Product Biosynthetic Gene Clusters in Yeast. *ACS Synth. Biol.* 5(9):1002–10

305. Li S-Y, Zhao G-P, Wang J. 2016. C-Brick: A New Standard for Assembly of Biological Parts Using Cpf1. *ACS Synth. Biol.* 5(12):1383–88

306. Christie KA, Guo JA, Silverstein RA, Doll RM, Mabuchi M, et al. 2023. Precise DNA cleavage using CRISPR-SpRYgests. *Nat. Biotechnol.* 41(3):409–16

307. Biggs BW, Bedore SR, Arvay E, Huang S, Subramanian H, et al. 2020. Development of a genetic toolset for the highly engineerable and metabolically versatile Acinetobacter baylyi ADP1. *Nucleic Acids Res.* 48(9):5169–82

308. Collias D, Vialetto E, Yu J, Co K, Almási É d. H, et al. 2023. Systematically

attenuating DNA targeting enabled CRISPR-driven editing in bacteria. *Nat. Commun.* 14(1):680

309. Cook TB, Rand JM, Nurani W, Courtney DK, Liu SA, Pfleger BF. 2018. Genetic tools for reliable gene expression and recombineering in Pseudomonas putida. *J. Ind. Microbiol. Biotechnol.* 45(7):517–27

310. Wirth NT, Kozaeva E, Nikel PI. 2019. Accelerated genome engineering of Pseudomonas putida by I-SceI-mediated recombination and CRISPR-Cas9 counterselection. *Microb Biotechnol*. n/a(n/a):

311. Enghiad B, Huang C, Guo F, Jiang G, Wang B, et al. 2021. Cas12a-assisted precise targeted cloning using in vivo Cre-lox recombination. *Nat. Commun.* 12(1):1171

312. Liang M, Liu L, Xu F, Zeng X, Wang R, et al. 2022. Activating cryptic biosynthetic gene cluster through a CRISPR–Cas12a-mediated direct cloning approach. *Nucleic Acids Res.* 50(6):3581–92

313. Shola DTN, Yang C, Kewaldar V-S, Kar P, Bustos V. 2020. New Additions to the CRISPR Toolbox: CRISPR-CLONInG and CRISPR-CLIP for Donor Construction in Genome Editing. *CRISPR J.* 3(2):109–22

314. Wen X, Zhang Y, Cheng H, An J, Guo Y, et al. 2021. A CRISPR/dCas9-assisted system to clone toxic genes in Escherichia coli. *Biochim. Biophys. Acta BBA - Gen. Subj.* 1865(11):129994

315. Garst AD, Bassalo MC, Pines G, Lynch SA, Halweg-Edwards AL, et al. 2017. Genome-wide mapping of mutations at single-nucleotide resolution for protein, metabolic and genome engineering. *Nat. Biotechnol.* 35(1):48–55

316. Ronda C, Pedersen LE, Sommer MOA, Nielsen AT. 2016. CRMAGE: CRISPR Optimized MAGE Recombineering. *Sci. Rep.* 6(1):19452

317. Liang L, Liu R, Foster KEO, AlakshChoudhury, Cook S, et al. 2020. Genome engineering of E. coli for improved styrene production. *Metab. Eng.* 57:74–84

318. Liu R, Liang L, Choudhury A, Bassalo MC, Garst AD, et al. 2018. Iterative genome editing of Escherichia coli for 3-hydroxypropionic acid production. *Metab. Eng.* 47:303–13

319. Liu R, Liang L, Garst AD, Choudhury A, Nogué VS i, et al. 2018. Directed combinatorial mutagenesis of Escherichia coli for complex phenotype engineering. *Metab. Eng.* 47:10–20

320. Garg D, Samota MK, Kontis N, Patel N, Bala S, Rosado AS. 2023. Revolutionizing biofuel generation: Unleashing the power of CRISPR-Cas mediated gene editing of extremophiles. *Microbiol. Res.* 274:127443

321. Ke J, Robinson D, Wu Z-Y, Kuftin A, Louie K, et al. 2022. CRAGE-CRISPR facilitates rapid activation of secondary metabolite biosynthetic gene clusters in bacteria. *Cell Chem. Biol.* 29(4):696-710.e4

322. Trujillo Rodríguez L, Ellington AJ, Reisch CR, Chevrette MG. 2023. CRISPR-

Associated Transposase for Targeted Mutagenesis in Diverse Proteobacteria.
2018 *ACS Synth. Biol.* 12(7):1989–2003

323. Yang S, Zhu J, Zhou X, Zhang J, Li Q, et al. 2023. RNA-Guided DNA
2020 Transposition in Corynebacterium glutamicum and Bacillus subtilis. *ACS Synth. Biol.* 12(7):2198–2202

2022 324. Heinsch SC, Das SR, Smanski MJ. 2018. Simulation Modeling to Compare High-Throughput, Low-Iteration Optimization Strategies for Metabolic Engineering.
2024 *Front. Microbiol.* 9:

325. Perfeito L, Ghozzi S, Berg J, Schnetz K, Lässig M. 2011. Nonlinear Fitness
2026 Landscape of a Molecular Pathway. *PLOS Genet.* 7(7):e1002160

326. Jones JA, Vernacchio VR, Lachance DM, Lebovich M, Fu L, et al. 2015.
2028 ePathOptimize: A Combinatorial Approach for Transcriptional Balancing of Metabolic Pathways. *Sci. Rep.* 5:11301

2030 327. Martin VJJ, Pitera DJ, Withers ST, Newman JD, Keasling JD. 2003. Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nat.*
2032 *Biotechnol.* 21(7):796–802

328. Abdullah, Jiang Z, Hong X, Zhang S, Yao R, Xiao Y. 2020. CRISPR base editing
2034 and prime editing: DSB and template-free editing systems for bacteria and plants. *Synth. Syst. Biotechnol.* 5(4):277–92

2036 329. Zhang Y, Yuan J. 2021. CRISPR/Cas12a-mediated genome engineering in the

photosynthetic bacterium Rhodobacter capsulatus. *Microb. Biotechnol.* 14(6):2700–2710

330. Li H, Shen CR, Huang C-H, Sung L-Y, Wu M-Y, Hu Y-C. 2016. CRISPR-Cas9 for the genome engineering of cyanobacteria and succinate production. *Metab. Eng.* 38:293–302

331. Niu T-C, Lin G-M, Xie L-R, Wang Z-Q, Xing W-Y, et al. 2019. Expanding the Potential of CRISPR-Cpf1-Based Genome Editing Technology in the Cyanobacterium Anabaena PCC 7120. *ACS Synth. Biol.* 8(1):170–80

332. Wang S-Y, Li X, Wang S-G, Xia P-F. 2023. Base editing for reprogramming cyanobacterium Synechococcus elongatus. *Metab. Eng.* 75:91–99

333. McAllister KN, Sorg JA. 2019. CRISPR Genome Editing Systems in the Genus *Clostridium*: a Timely Advancement

334. Choi SY, Woo HM. 2020. CRISPRi-dCas12a: A dCas12a-Mediated CRISPR Interference for Repression of Multiple Genes and Metabolic Engineering in Cyanobacteria. *ACS Synth. Biol.* 9(9):2351–61

335. Shin J, Bae J, Lee H, Kang S, Jin S, et al. 2023. Genome-wide CRISPRi screen identifies enhanced autolithotrophic phenotypes in acetogenic bacterium *Eubacterium limosum*

336. Yunus IS, Anfelt J, Sporre E, Miao R, Hudson EP, Jones PR. 2022. Synthetic metabolic pathways for conversion of CO2 into secreted short-to medium-chain

hydrocarbons using cyanobacteria. *Metab. Eng.* 72:14–23

337. Wang Y, Yue X, Yuan S, Hong Y, Hu W, Li Y. 2021. Internal Promoters and Their Effects on the Transcription of Operon Genes for Epothilone Production in Myxococcus xanthus. *Front. Bioeng. Biotechnol.* 9:

338. Bhokisham N, VanArsdale E, Stephens KT, Hauk P, Payne GF, Bentley WE. 2020. A redox-based electrogenetic CRISPR system to connect with and control biological information networks. *Nat. Commun.* 11(1):2427

339. Fang L, Fan J, Luo S, Chen Y, Wang C, et al. 2021. Genome-scale target identification in Escherichia coli for high-titer production of free fatty acids. *Nat. Commun.* 12(1):4976

340. Tian T, Kang JW, Kang A, Lee TS. 2019. Redirecting Metabolic Flux via Combinatorial Multiplex CRISPRi-Mediated Repression for Isopentenol Production in Escherichia coli. *ACS Synth. Biol.* 8(2):391–402

341. Wang Y, Cheng H, Liu Y, Liu Y, Wen X, et al. 2021. In-situ generation of large numbers of genetic combinations for metabolic reprogramming via CRISPR-guided base editing. *Nat. Commun.* 12(1):678

342. Martin HG, Radivojevic T, Zucker J, Bouchard K, Sustarich J, et al. 2023. Perspectives for self-driving labs in synthetic biology. *Curr. Opin. Biotechnol.* 79:102881

343. Kuchina A, Brettner LM, Paleologu L, Roco CM, Rosenberg AB, et al. 2021.

Microbial single-cell RNA sequencing by split-pool barcoding. *Science*. 371(6531):eaba5257

344. Fenster JA, Werner AZ, Tay JW, Gillen M, Schirokauer L, et al. 2022. Dynamic and single cell characterization of a CRISPR-interference toolset in Pseudomonas putida KT2440 for β-ketoadipate production from p-coumarate. *Metab. Eng. Commun.* 15:e00204

345. Liu J, Liu M, Shi T, Sun G, Gao N, et al. 2022. CRISPR-assisted rational flux-tuning and arrayed CRISPRi screening of an l-proline exporter for l-proline hyperproduction. *Nat. Commun.* 13(1):891

346. Wang J, Li C, Jiang T, Yan Y. 2023. Biosensor-assisted titratable CRISPRi high-throughput (BATCH) screening for over-production phenotypes. *Metab. Eng.* 75:58–67

347. Bat-Erdene U, Billingsley JM, Turner WC, Lichman BR, Ippoliti FM, et al. 2021. Cell-Free Total Biosynthesis of Plant Terpene Natural Products Using an Orthogonal Cofactor Regeneration System. *ACS Catal.* 11(15):9898–9903

348. Fessner W-D. 2015. Systems Biocatalysis: Development and engineering of cell-free “artificial metabolisms” for preparative multi-enzymatic synthesis. *Eur. Congr. Biotechnol. - ECB 16*. 32(6):658–64

349. Schwander T, Schada von Borzyskowski L, Burgener S, Cortina NS, Erb TJ. 2016. A synthetic pathway for the fixation of carbon dioxide in vitro

350. Richardson KN, Black WB, Li H. 2020. Aldehyde Production in Crude Lysate- and Whole Cell-Based Biotransformation Using a Noncanonical Redox Cofactor System. *ACS Catal.* 10(15):8898–8903

351. Stark JC, Jaroentomeechai T, Moeller TD, Hershewe JM, Warfel KF, et al. 2021. On-demand biomanufacturing of protective conjugate vaccines

352. Vögeli B, Schulz L, Garg S, Tarasava K, Clomburg JM, et al. 2022. Cell-free prototyping enables implementation of optimized reverse β-oxidation pathways in heterotrophic and autotrophic bacteria. *Nat. Commun.* 13(1):3058

353. Claassens NJ, Burgener S, Vögeli B, Erb TJ, Bar-Even A. 2019. A critical comparison of cellular and cell-free bioproduction systems. *Pharm. Biotechnol. ● Chem. Biotechnol.* 60:221–29

354. Kelwick RJR, Webb AJ, Freemont PS. 2020. Biological Materials: The Next Frontier for Cell-Free Synthetic Biology. *Front. Bioeng. Biotechnol.* 8:

355. Moritz C, Sundaram S, Diehl C, Adam D, Borkowski O, Pandi A. 2021. Chapter 21 - Advances and applications of cell-free systems for metabolic production. In *Microbial Cell Factories Engineering for Production of Biomolecules*, ed V Singh, pp. 407–20. Academic Press

356. Miguez AM, McNerney MP, Styczynski MP. 2019. Metabolic Profiling of Escherichia coli-Based Cell-Free Expression Systems for Process Optimization. *Ind. Eng. Chem. Res.* 58(50):22472–82

357. Wagner L, Jules M, Borkowski O. 2023. What remains from living cells in bacterial lysate-based cell-free systems. *Comput. Struct. Biotechnol. J.* 21:3173–82

358. Garenne D, Beisel CL, Noireaux V. 2019. Characterization of the all-E. coli transcription-translation system myTXTL by mass spectrometry. *Rapid Commun. Mass Spectrom.* 33(11):1036–48

359. Rasor BJ, Chirania P, Rybnicky GA, Giannone RJ, Engle NL, et al. 2023. Mechanistic Insights into Cell-Free Gene Expression through an Integrated -Omics Analysis of Extract Processing Methods. *ACS Synth. Biol.* 12(2):405–18

360. Contreras-Llano LE, Meyer C, Liu Y, Sarker M, Lim S, et al. 2020. Holistic engineering of cell-free systems through proteome-reprogramming synthetic circuits. *Nat. Commun.* 11(1):3138

361. Dinglasan JLN, Doktycz MJ. 2023. Rewiring cell-free metabolic flux in E. coli lysates using a block—push—pull approach. *Synth. Biol.* 8(1):ysad007

362. Krüger A, Mueller AP, Rybnicky GA, Engle NL, Yang ZK, et al. 2020. Development of a clostridia-based cell-free system for prototyping genetic parts and metabolic pathways. *Metab. Eng.* 62:95–105

363. Moore SJ, MacDonald JT, Wienecke S, Ishwarbhai A, Tsipa A, et al. 2018. Rapid acquisition and model-based analysis of cell-free transcription–translation reactions from nonmodel bacteria. *Proc. Natl. Acad. Sci.* 115(19):E4340–49

364. Amrofell MB, Rottinghaus AG, Moon TS. 2020. Engineering microbial diagnostics

and therapeutics with smart control. *Tissue Cell Pathw. Eng.* 66:11–17

365. Omer R, Mohsin MZ, Mohsin A, Mushtaq BS, Huang X, et al. 2022. Engineered Bacteria-Based Living Materials for Biotherapeutic Applications. *Front. Bioeng. Biotechnol.* 10:

366. Schmidt F, Zimmermann J, Tanna T, Farouni R, Conway T, et al. 2022. Noninvasive assessment of gut function using transcriptional recording sentinel cells. *Science*. 376(6594):eabm6038

367. Chan CTY, Lee JW, Cameron DE, Bashor CJ, Collins JJ. 2016. “Deadman” and “Passcode” microbial kill switches for bacterial containment. *Nat. Chem. Biol.* 12(2):82–86

368. Stirling F, Bitzan L, O’Keefe S, Redfield E, Oliver JWK, et al. 2017. Rational Design of Evolutionarily Stable Microbial Kill Switches. *Mol. Cell*. 68(4):686-697.e3

369. Hartline CJ, Schmitz AC, Han Y, Zhang F. 2021. Dynamic control in metabolic engineering: Theories, tools, and applications. *Tools Strateg. Metab. Eng.* 63:126–40

370. Wu Y, Chen T, Liu Y, Tian R, Lv X, et al. 2020. Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in Bacillus subtilis. *Nucleic Acids Res.* 48(2):996–1009

371. Lv Y, Qian S, Du G, Chen J, Zhou J, Xu P. 2019. Coupling feedback genetic

circuits with growth phenotype for dynamic population control and intelligent bioproduction. *Metab. Eng.* 54:109–16

372. Tian J, Yang G, Gu Y, Sun X, Lu Y, Jiang W. 2020. Developing an endogenous quorum-sensing based CRISPRi circuit for autonomous and tunable dynamic regulation of multiple targets in Streptomyces. *Nucleic Acids Res.* 48(14):8188–8202

373. Dinh CV, Prather KLJ. 2019. Development of an autonomous and bifunctional quorum-sensing circuit for metabolic flux control in engineered Escherichia coli. *Proc. Natl. Acad. Sci.* 116(51):25562–68

374. Gupta A, Reizman IMB, Reisch CR, Prather KLJ. 2017. Dynamic regulation of metabolic flux in engineered bacteria using a pathway-independent quorum-sensing circuit. *Nat. Biotechnol.* 35(3):273–79

375. Liu Y, Chen J, Crisante D, Jaramillo Lopez JM, Mahadevan R. 2020. Dynamic Cell Programming with Quorum Sensing-Controlled CRISPRi Circuit. *ACS Synth. Biol.* 9(6):1284–91

376. Gao C, Guo L, Hu G, Liu J, Chen X, et al. 2021. Engineering a CRISPRi Circuit for Autonomous Control of Metabolic Flux in Escherichia coli. *ACS Synth. Biol.* 10(10):2661–71

377. Li S, Ye Z, Moreb EA, Hennigan JN, Castellanos DB, et al. 2021. Dynamic control over feedback regulatory mechanisms improves NADPH flux and xylitol biosynthesis in engineered E. coli. *Metab. Eng.* 64:26–40

2178 378. Ye Z, Li S, Hennigan JN, Lebeau J, Moreb EA, et al. 2021. Two-stage dynamic deregulation of metabolism improves process robustness & scalability in
2180 engineered E. coli. *Metab. Eng.* 68:106–18

379. Shabestary K, Anfelt J, Ljungqvist E, Jahn M, Yao L, Hudson EP. 2018. Targeted
2182 Repression of Essential Genes To Arrest Growth and Increase Carbon Partitioning and Biofuel Titers in Cyanobacteria. *ACS Synth. Biol.* 7(7):1669–75

2184 380. Banerjee D, Yunus IS, Wang X, Kim J, Srinivasan A, et al. 2023. Genome-scale and pathway engineering for the sustainable aviation fuel precursor isoprenol
2186 production in Pseudomonas putida

381. Kozaeva E, Volkova S, Matos MRA, Mezzina MP, Wulff T, et al. 2021. Model-
2188 guided dynamic control of essential metabolic nodes boosts acetyl-coenzyme A–dependent bioproduction in rewired Pseudomonas putida. *Metab. Eng.* 67:373–86

2190 382. Shabestary K, Hernández HP, Miao R, Ljungqvist E, Hallman O, et al. 2021. Cycling between growth and production phases increases cyanobacteria
2192 bioproduction of lactate. *Metab. Eng.* 68:131–41

383. Rodrigo-Navarro A, Sankaran S, Dalby MJ, del Campo A, Salmeron-Sanchez M.
2194 2021. Engineered living biomaterials. *Nat. Rev. Mater.* 6(12):1175–90

384. Bao T, Qian Y, Xin Y, Collins JJ, Lu T. 2023. Engineering microbial division of
2196 labor for plastic upcycling. *Nat. Commun.* 14(1):5712

385. Brooks SM, Marsan C, Reed KB, Yuan S-F, Nguyen D-D, et al. 2023. A tripartite

microbial co-culture system for de novo biosynthesis of diverse plant phenylpropanoids. *Nat. Commun.* 14(1):4448

386. Johnston TG, Yuan S-F, Wagner JM, Yi X, Saha A, et al. 2020. Compartmentalized microbes and co-cultures in hydrogels for on-demand bioproduction and preservation. *Nat. Commun.* 11(1):563

387. Lawson CE, Martí JM, Radivojevic T, Jonnalagadda SVR, Gentz R, et al. 2021. Machine learning for metabolic engineering: A review. *Tools Strateg. Metab. Eng.* 63:34–60

388. Radivojević T, Costello Z, Workman K, Garcia Martin H. 2020. A machine learning Automated Recommendation Tool for synthetic biology. *Nat. Commun.* 11(1):4879

389. Roohani Y, Huang K, Leskovec J. 2023. Predicting transcriptional outcomes of novel multigene perturbations with GEARS. *Nat. Biotechnol.*

390. Hao N, Shearwin KE, Dodd IB. 2017. Programmable DNA looping using engineered bivalent dCas9 complexes. *Nat. Commun.* 8(1):1628

391. Hu C, van Beljouw SPB, Nam KH, Schuler G, Ding F, et al. 2022. Craspase is a CRISPR RNA-guided, RNA-activated protease. *Science*. 377(6612):1278–85

392. Joshi SH-N, Yong C, Gyorgy A. 2022. Inducible plasmid copy number control for synthetic biology in commonly used E. coli strains. *Nat. Commun.* 13(1):6691

393. Li C, Zou Y, Jiang T, Zhang J, Yan Y. 2022. Harnessing plasmid replication

mechanism to enable dynamic control of gene copy in bacteria. *Metab. Eng.* 70:67–78

394. Mao X, Ma Q, Liu B, Chen X, Zhang H, Xu Y. 2015. Revisiting operons: an analysis of the landscape of transcriptional units in E. coli. *BMC Bioinformatics*. 16(1):356

395. Bar-Yaacov D, Pilpel Y, Dahan O. 2018. RNA editing in bacteria: occurrence, regulation and significance. *RNA Biol.* 15(7):863–67

396. Pan C, Li G, Malzahn AA, Cheng Y, Leyson B, et al. 2022. Boosting plant genome editing with a versatile CRISPR-Combo system. *Nat. Plants*. 8(5):513–25

397. Kirkpatrick RL, Lewis K, Langan RA, Lajoie MJ, Boyken SE, et al. 2020. Conditional Recruitment to a DNA-Bound CRISPR–Cas Complex Using a Colocalization-Dependent Protein Switch. *ACS Synth. Biol.* 9(9):2316–23

Chapter 2:

Multi-Layer CRISPRa/i Circuits for Dynamic Genetic Programs in Cell-Free and Bacterial Systems

Benjamin I. Tickman[+,1], Diego Alba Burbano[+,1,2], Venkata P. Chavali[1], Cholpisit Kiattisewee[1], Jason Fontana[1], Aset Khakimzhan[3], Vincent Noireaux[3], Jesse G. Zalatan[*,1,4], James M. Carothers[*,1,2,5]

1: Molecular Engineering & Sciences Institute
and Center for Synthetic Biology
University of Washington
Seattle, WA 98195
United States

2: Department of Chemical Engineering
University of Washington
Seattle, WA 98195
United States

3: School of Physics and Astronomy
University of Minnesota
Minneapolis, MN 55455, USA

4: Department of Chemistry
University of Washington
Seattle, WA 98195
United States

5: Lead contact

+: These authors contributed equally

*: Corresponding authors
zalatan@uw.edu
206-543-1670

jcaroth@uw.edu
206-221-4902

Published on *Cell Systems* on March 16th, 2022**.**
DOI: /10.1016/j.cels.2021.10.008

Summary

CRISPR-Cas transcriptional circuits hold great promise as platforms for engineering metabolic networks and information processing circuits. Historically, prokaryotic CRISPR control systems have been limited to CRISPRi. Creating approaches to integrate CRISPRa for transcriptional activation with existing CRISPRi-based systems would greatly expand CRISPR circuit design space. Here, we develop design principles for engineering prokaryotic CRISPRa/i genetic circuits with network topologies specified by guide RNAs. We demonstrate that multi-layer CRISPRa/i cascades and feedforward loops can operate through the regulated expression of guide RNAs in cell-free expression systems and *E. coli*. We show that CRISPRa/i circuits can program complex functions by designing type 1 incoherent feedforward loops acting as fold-change detectors and tunable pulse-generators. By investigating how component characteristics relate to network properties such as depth, width, and speed, this work establishes a framework for building scalable CRISPRa/i circuits as regulatory programs in cell-free expression systems and bacterial hosts.

Keywords

CRISPRa, CRISPRi, Transcriptional Circuits, I1-FFL, *E. coli*, Cell-Free, Expression Dynamics

Introduction

Inspired by nature, synthetic biologists seek to dynamically regulate gene expression in biological systems to conserve resources, respond to stimuli, and generate complex, time-dependent behavior (Brockman and Prather, 2015; Dinh and Prather, 2020; Fontana et al., 2018a; Santos-Moreno and Schaerli, 2020). However, there are limited examples of synthetic, dynamic transcriptional regulatory networks capable of complex, multi-gene regulation. This rarity can be attributed to the limited number of suitable components for implementing scalable regulatory networks (English et al., 2021; Jeong et al., 2019; Nielsen et al., 2016), and to the difficulty of sequentially combining components into multi-layered operations (Brophy and Voigt, 2014; Gander et al., 2017; Lucks et al., 2008; Qian et al., 2017). Hence, a scalable framework enabling rational design and tuning of dynamic regulatory programs would constitute a significant advance.

CRISPR-Cas transcriptional controls have emerged as a promising route for building gene regulatory networks enabling programmable and orthogonal control at many loci simultaneously (Banerjee et al., 2020; Landberg et al., 2020; Reis et al., 2019; Tian et al., 2019). In these systems, nuclease defective Cas proteins such as *S. pyogenes* dCas9 are combined with guide RNAs (gRNAs) that specify DNA targets. Targeting of this complex to promoters or open reading frames generates gene repression (CRISPRi). Scalable multi-gene circuits can thus be implemented simply through the programmed expression of multiple gRNAs (Gander et al., 2017; Huang et al., 2021; Nielsen and Voigt, 2014; Santos-Moreno et al., 2020). Recent efforts have demonstrated the construction of CRISPRi circuits capable of performing a diverse set of Boolean logic evaluations (Gander et al., 2017; Nielsen and Voigt, 2014; Tan and Ng, 2021; Xiang et al., 2018), and

dynamic expression programs (Dinh and Prather, 2019; Tian et al., 2020; Westbrook et al., 2019; Wu et al., 2020b). The recent discovery of new transcriptional activators and promoter design rules for effective CRISPR activation (CRISPRa) in bacteria raised the possibility of circuits combining CRISPRa and CRISPRi to form dynamic gene regulatory networks in prokaryotic systems (Dong et al., 2018; Fontana et al., 2020; Kiattisewee et al., 2021; Liu et al., 2019). Such circuits would enable network topologies and functional capabilities not possible with CRISPRi alone. Implementation of simultaneous CRISPRa and CRISPRi has been demonstrated on both multiple and individual genes (Table S2) (Dong et al., 2018; Kiattisewee et al., 2021; Wu et al., 2020a). In this work, we endeavoured to apply CRISPRa in genetic circuits that go beyond elementary operations in a single layer (Bikard et al., 2013; English et al., 2021; Fontana et al., 2020; Liu et al., 2019).

In this work, we develop genetic components and design strategies allowing CRISPRa to be combined with CRISPRi to generate a multi-layer CRISPRa/i transcriptional control system operating in *E. coli* and an *E. coli*-derived cell-free expression system (CFS). We show that the strength and dynamics of control actions can be tuned through the regulated expression of guide RNAs, with CRISPR-activation ratios as high as 25 fold in CFS and 40 fold in *E. coli*. We combine components into multi-layered operations by level-matching the output expression levels of upstream components to the acceptable input range of downstream components (McDaniel and Weiss, 2005; Wang et al., 2013). We report the successful construction and tuning of multi-guide CRISPRa/i cascades and type 1 incoherent feed forward loops (I1-FFLs) in CFS and *E. coli* to programmably achieve complex behaviors such as pulse generation

and fold-change detection. Together, a set of generalizable design rules and an expandable toolbox of orthogonal components provide a framework for rapid and scalable implementation of higher order CRISPRa/i regulatory networks. We envision that these capabilities will prove useful for the next generation of dynamically-regulated metabolic engineering efforts, multiplexed biosensing, and self-adaptive biocomputation (Bartoli et al., 2020; Wan et al., 2019; Wu et al., 2020b).

Results

CRISPRa/i Circuits in CFS

Bacterial CRISPRa is Functional in *E. coli* CFS

Cell-free systems have become an attractive platform for prototyping of genetic circuits, construction of synthetic cells and engineered biosynthetic pathways (Adamala et al., 2017; Dudley et al., 2015; Garamella et al., 2016; Karim et al., 2016; Marshall and Noireaux, 2018). However, there are limited examples of genetic circuits capable of dynamic, multi-gene regulation in CFS. CRISPRi-based genetic control is well established in CFS (Marshall and Noireaux, 2018; Westbrook et al., 2019). Incorporating CRISPRa into CFS could enable more facile circuit engineering by increasing the number of realizable network topologies (Figure S1), and could overcome challenges that limit the utility of multi-layer CRISPRi repression circuitry in CFS.

A unique feature of CFS is that component turnover is greatly diminished compared to *in vivo* systems. CFS do not undergo dilution due to cell division and experience characteristically low protein and RNA turnover rates compared to cellular systems (Garamella et al., 2016). While component turnover can be accelerated via the addition of degradation machinery, this approach is inefficient and consumes valuable, finite resources (Garamella et al., 2016). In practice, this limited turnover makes repression circuits difficult to implement because even if transcription is halted the gene product is already present. In contrast, circuits based on activation are not dependent on turnover. Implementation of activation based regulatory circuits would allow the high volumetric productivities resulting from limited component turnover to be combined with complex and dynamic multi-layer regulatory circuitry in a cell free setting (Garamella et

al., 2016). Thus, our first challenge was to adapt the CRISPRa system developed in *E. coli* for use in CFS. In this system, CRISPRa is applied using a 3'-extended guide RNA (scaffold RNA or scRNA) to direct dCas9 upstream of σ^{70} promoters. The 3' extension of the guide RNA contains an RNA hairpin (MS2) which binds an RNA binding protein (MCP) fused to a transcriptional activator (SoxS) (Figure 1A). In this system scRNAs encode information for targeting of dCas9 to precise locations along DNA as well as recruitment of a functional effector (Dong et al., 2018; Fontana et al., 2020; Kiattisewee et al., 2021; Zalatan et al., 2015). These scRNAs, J106, J206, and J306, are targeted via the spacer sequence directing CRISPRa to an expandable set of orthogonal synthetic promoters J1, J2, and J3 (Fontana et al., 2020).

To understand the portability of the CRISPRa system between E. *coli* and CFS, we tested whether basal expression levels and gene activation in CFS corresponded to previously-observed trends in E. *coli*. In cells, CRISPRa can produce high levels of gene expression from a broad range of promoter strengths, but the fold-activation decreases as basal promoter strength increases. We tested a set of synthetic minimal promoters (BBa_J231XX) (Figure S2) of varying strength in CFS. We observed a high correlation between CFS and E. *coli* for both basal promoter strength and fold-activation by CRISPRa, providing Spearman correlation coefficients of 0.91 and 0.88, respectively (Figure S3). This observed correspondence between component function in *E. coli*-derived CFS and *E. coli* is consistent with previous reports (Garamella et al., 2016; Marshall et al., 2018; Shin and Noireaux, 2012), allowing exchange of individual genetic components between the two systems.

Next, we sought to formalize a framework for the construction of higher order CRISPRa/i circuits operating through interconnected CRISPRa/i nodes (Figure 1B). In this framework, CRISPRa/i nodes are discrete transcriptional units containing target sequences for CRISPRa- and/or CRISPRi-directed transcriptional regulation (Figure 1A). To characterize CRISPRa/i nodes, we isolated dCas9, sc/sgRNAs, and the MCP-SoxS activator onto individual plasmids (Figure 1C), allowing independent titration of expression levels for all CRISPRa/i components. CRISPRa/i node characterization is conducted by measuring the output response of each node to varying levels of component transcriptional inputs provided by titrating component plasmid concentrations. We found that increasing the concentration of sc/sgRNA expressing plasmid resulted in higher overall levels of activation and repression (Figure 1D, 1E), as well as faster control (Figure S4). Titrations of dCas9 expressing plasmid revealed no differences in the strength of CRISPRa across a 40-fold range of dCas9 expression levels generated by 0.05 nM to 2 nM dCas9 expressing plasmid (Figure S5, left), likely due to saturation of active CRISPRa complexes at these concentrations of DNA target. For all levels of dCas9 expression, we observed a ~40 min delay between initiation of the cell-free reaction and the onset of CRISPRa/i control, consistent with previously reported time of dCas9 maturation and CRISPR complex formation in CFS (Marshall et al., 2018; Westbrook et al., 2019). Titrations of plasmid expressing MCP-SoxS revealed a relatively wide region between 1 nM and 24 nM over which no significant differences in endpoint measurements of CRISPRa mediated outputs were observed (Figure S5, right). Expression levels for dCas9 and MCP-SoxS activator were held constant throughout the work at 2nM and 4nM plasmid concentrations respectively.

Tuning CRISPRa/i through the Regulated Expression of Guide RNAs in CFS

The ability to easily vary plasmid concentration in CFS, combined with the multi-component nature of CRISPRa/i regulatory complexes (Figure 1C), enables tuning of all component expression levels independently. Some tuning actions are global, for instance varying dCas9 expression levels impacts both CRISPRa and CRISPRi. Other tuning actions, such as varying the level of activator protein, are expected to influence scRNA-mediated activation but not sgRNA-mediated repression. Likewise, output levels for individual nodes in a circuit can be linearly scaled by changing the concentration of plasmid at that node (Figure 1B). To provide simultaneous and independent control over both timing and expression levels of multiple target genes, we tuned CRISPRa/i control actions through the regulated expression of guideRNAs. Here, the specificity provided by guideRNA targeting allows tuning actions to be applied locally to individual CRISPRa/i nodes.

To quantify time-varying CRISPRa/i-directed changes in gene expression, we calculate production rates of CRISPR-regulated RFP expression relative to unregulated, basal expression of RFP. (relative RFP production rate). At saturating levels of scRNA expression, CRISPRa achieved constant levels of activation over the course of 4-6 hrs providing a ~20-fold ± 2 fold increase in RFP production rate relative to an off-target control (Figure 1D, right). Likewise, relative production rates from CRISPRi with saturating levels of sgRNA (Figure 1E, right) achieved steady state levels of repression by 3 hours. Guide RNA titrations revealed that increasing levels of scRNA decreased the time to 2-fold activation by CRISPRa by up to ~5 hours, and increasing levels of sgRNA decreased the time to 50% repression by CRISPRi by up to ~10 hours (Figure S4, left). We found

that sgRNA titrations were able to significantly affect the overall timing of gene expression as determined by the time to half maximum endpoint RFP values, with strong CRISPRi providing a ~3 hour shift to earlier time points as compared to a no sgRNA control (Figure S4, right). Qualitatively, we observed that increasing sgRNA expression levels resulted in a higher fraction of total expression occurring at early time points. In contrast, scRNA titrations primarily provided a scaling factor to CRISPRa output levels without greatly affecting the timing of expression (Figure 1D, 1E, right; Figure S4, right). These data suggest that under these conditions CRISPRa kinetics may be dominated by the time required for MCP-SoxS expression and maturation.

Level-Matching of Multi-Layer CRISPR Circuitry in CFS

To enable the construction of multi-layer circuits, we built activation and activation-repression cascades by level-matching the input/output dynamic ranges between sequential CRISPRa/i nodes. That is, we matched the output transcription levels of an upstream node encoding scRNA to the relevant transcriptional input range of a downstream node encoding another sc/sgRNA. From sc/sgRNA plasmid titrations, we observed that both CRISPRa and CRISPRi respond to changes in input guide RNA expression levels spanning approximately 2 orders of magnitude (Figure 2A, 2C). Across this responsive range of sc/sgRNA inputs, CRISPRa controlled outputs vary by ~24-fold at endpoint. While the dynamic range of CRISPRa generated outputs does not fully span the dynamic range of sc/sgRNA inputs, this characterization suggests that CRISPRa/i nodes can be sequentially combined by careful matching of upstream output ranges to downstream input ranges to form layered operations.

By tuning the concentration of plasmid expressing sc/sgRNA in the second layer of the cascades we were able to control the degree of overlap between response curves of upstream and downstream layers in the circuit. Based on the scRNA dose-response curves for CRISPRa by two different scRNAs in isolation (Figure 2A), we decided to build CRISPR activation-activation cascades to probe the composability of CRISPRa circuits from components characterized in isolation. Using these curves, we predicted how a 24-fold increase in transcription provided by CRISPRa in the first layer of the cascade affects cascade output. We made this prediction for four different concentrations of plasmid expressing scRNA in the second layer (Figure 2B). Upon construction of these CRISPRa cascades, we observed a strong agreement (R^2 = 0.985) between measured and predicted fold increases in outputs of CRISPRa cascades (Figure 2B). As expected, overlap between layers was maximized at 2 nM scRNA plasmid in the second layer, with the cascade providing a 16.3 ± 3.0 fold increase in measured RFP at endpoint compared to CRISPRa alone. Both lower, 0.5 nM, and higher, 4 nM, concentrations of scRNA expressing plasmid in the second layer of the cascade resulted in decreased fold changes in cascade output, at 7.6 ± .6 fold and 15.2 ± 2.3 fold respectively. From these data, we calculated the efficiency of signal propagation through the activation cascade by comparing the observed fold change in cascade output to the fold change provided by CRISPRa in the input layer. At present, given optimal level-matching we observe 67.9% ± 18.1% signal propagation for the two layer CRISPRa cascade. Together, these results suggest that we can predictably tune the degree of overlap between layers of CRISPRa/i circuits to propagate signals, satisfy input requirements of potential downstream layers, and tailor absolute gene expression levels.

Next, we constructed an activation-repression cascade with CRISPRa in the input layer activating transcription of RR2 sgRNA which targets the coding sequence of *mRFP1* for CRISPRi in the second layer of the cascade (Figure 2D, top). As in construction of the activation-activation cascade, level-matching was informed by the sc/sgRNA dose-response curves for CRISPRa and CRISPRi obtained in isolation (Figure 2C). As expected, when the overlap between layers was maximized, the CRISPRa/i cascade generated 4.6-fold ± 0.7 more repression than CRISPRi alone (Figure 2D, right). The importance of matching upstream outputs to the responsive range of downstream inputs was illustrated by overexpression of sgRNA in the second layer of the activation-repression cascade, intentionally minimizing the overlap between the upstream and downstream layers of the cascade. Under these conditions, the activation-repression cascade reduced RFP expression by 1.4-fold ± 0.3 compared to CRISPRi (Figure 2D, left).

CRISPRa/i Circuits Encode Dynamic Gene Expression Programs

Next, we sought to investigate the ability of multi-layer CRISPRa/i circuitry to encode dynamic gene expression programs inaccessible to simpler single layer controllers. As a first step, we explored the influence of level-matching on CRISPRa/i cascade dynamics by comparing relative RFP production rates arising from an activation-repression cascade to those generated by CRISPRi. When there is a high degree of overlap between the response curves of the layers in the cascade, significant changes in the timing of gene expression are observed (Figure 2E, right; 2F). For instance, an activation-repression cascade with 0.5 nM of sgRNA plasmid results in repression of the RFP output at a comparable rate, but delayed onset compared to that of CRISPRi

repression alone with 4 nM of sgRNA plasmid. This delay is interpreted as the time required for CRISPRa to activate sgRNA expression in the second layer of the circuit. We identify a ~10-fold range of sgRNA plasmid concentrations over which a CRISPR activation-repression cascade can generate significant differences in expression dynamics compared to single-layer CRISPRi (Figure 2F).

As expected, when sgRNA expression levels are mismatched we observe negligible differences in expression dynamics. If the concentration of plasmid expressing sgRNA in the second layer of the activation-repression cascade is too low, e.g. 0.01 nM, no repression is observed. Conversely, high concentrations of sgRNA expressing plasmid in the second layer of the cascade effectively result in CRISPRi applied in a single layer, producing expression dynamics identical to that of the CRISPRi control (Figure 2E, left). Above 2 nM of sgRNA expressing plasmid, we observed no difference in the time to 50% repression for the CRISPRa/i cascade as compared to single-layer CRISPRi (Figure 2F). Together, these results underscore that gene expression dynamics can be tuned by multi-layer CRISPRa/i circuits when there is sufficient overlap between the response curves of the CRISPRa/i circuit components.

After establishing the rules governing construction of layered CRISPRa/i circuitry, we endeavored to create more complex transcriptional programs to explore the scalability and composibility of CRISPRa/i regulatory networks. We combined the CRISPR activation-repression cascade with CRISPRa to form an incoherent type 1 feed forward loop (I1-FFL), a classic pulse generating circuit that is significantly overrepresented in natural systems (Alon, 2007; Kaplan et al., 2008; Mangan and Alon, 2003; Shen-Orr et al., 2002). When level-matching is taken into consideration, we see that circuit topology

298 determines the timing of gene expression (Figure 3A). As expected, we observe no difference in expression dynamics between CRISPRa and CRISPRa+CRISPRi
300 (CRISPRa+i) at low concentrations of sgRNA expressing plasmid (Figure 3A, blue, orange). When expression of sgRNA at node Y is activated by CRISPRa to form an I1-
302 FFL (Figure 3A, red), we observe a gene expression pulse, qualitatively different from expression generated by CRISPRa or CRISPRa+i. Upon addition of an orthogonal I1-
304 FFL controlling expression of GFP to the same reaction (Figure 3A, green), no differences in the timing of gene expression are observed at the output of the RFP I1-FFL. This result
306 indicates that we can operate multiple circuits simultaneously without compromising the respective expression dynamics.

308 We were able to tune the timing of the gene expression pulse generated by the I1-FFL by varying the concentration of sgRNA expressing plasmid. The maximum RFP
310 production rate occurred ~110 min earlier in the cell-free reaction when we increased the sgRNA expressing plasmid concentration 10-fold from 0.1 nM to 1 nM (Figure 3B). More
312 generally, we observed that the time of the maximum gene expression pulse could be continuously tuned over a 4-fold change in sgRNA plasmid levels, shifting expression
314 maxima earlier by up to 2 hours compared to unregulated expression (Figure 3C). To capture the I1-FFL expression dynamics and evaluate the feasibility of rationally tuning
316 CRISPRa/i circuits *in silico,* we constructed a coarse-grained mechanistic model of CRISPRa/i gene regulation. We defined first order chemical reactions for protein and
318 guide RNA production, CRISPR complex assembly, and DNA targeting (Figure 3B, bottom; Table S3). When an initial experimental observation was provided, the model was
320 capable of predicting the effects of tuning actions applied to the I1-FFL on gene

expression dynamics (Figure 3B, fit and simulate). Here, the model is fit to the experimental data for an I1-FFL with 0.1 nM sgRNA expressing plasmid and used to predict the expression dynamics for an I1-FFL with 1 nM sgRNA expressing plasmid. Similar results were obtained when fitting to the 1 nM condition and predicting the 0.1 nM condition (Figure S8). We observed a ~10 min difference between the measured and predicted timing of maximum RFP production rate, corresponding to the time resolution of our measurements. Combined with the observed predictability of level-matching in CRISPRa/i cascades, these results suggest that high-fidelity CRISPRa/i circuits could be designed and tuned *in silico* given component characterization data.

CRISPRa/i Circuits in *E. coli*

To form dynamic, multi-layer circuitry in E. coli, CRISPRa/i circuits are encoded on two plasmids. One plasmid contains dCas9, MCP-SoxS, and scRNAs acting as inputs to the first layer of a circuit, while the second plasmid contains a fluorescent reporter as well as sc/sgRNAs acting in the second layer of a circuit (Figure 4A, 5A). Unlike cell-free systems, in which the expression level of every circuit component can be precisely titrated, gene expression in cells is constrained by the expression levels achievable given different combinations of plasmid copy number and genetic parts. Level-matching of multi-layer CRISPRa/i circuits in cells is therefore more challenging, and requires attention not only to the dynamic range of components, but also the absolute expression levels and activities.

<u>Level-Matching in Multi-Layer CRISPRa Circuits</u>

To understand the level-matching requirements of scRNAs in multi-layer CRISPRa cascades in *E. coli*, we engineered the basal expression characteristics of CRISPRa

nodes in the second layer of a 2 layer activation cascade. In this circuit, scRNA expressed at node X targets CRISPRa to a promoter at node Y, activating expression of a second scRNA, targeting a fluorescent reporter at node Z for CRISPRa. Tuning of basal expression levels was accomplished through the use of synthetic minimal promoters (BBa_J231XX), as well as modifications to the 5' sequence proximal to the minimal promoter at node Y, driving scRNA expression in the second layer (Figure 4B). At the highest basal levels of scRNA expression in the second layer, CRISPRa cascades yielded 1.4x higher output levels than a comparable single-layer circuit. Decreasing basal scRNA expression levels in the second layer of the cascade by ~10x increased the output dynamic range of the CRISPRa cascade to 5.9x. Decreasing basal scRNA expression levels by a further 3.4x increased the output dynamic range by an additional 2x, resulting in an overall activation ratio of 12.3x for the CRISPRa cascade as compared to the single-layer circuit. Output levels of CRISPRa cascades at all tested scRNA expression levels were comparable to output of single-layer CRISPRa with saturating levels of scRNA expression. We observed that CRISPRa cascades were sensitive to scRNA expression, with 32% compression of the output dynamic range observed even at the lowest basal expression level of scRNA at node Y. Compression of the output dynamic range in cascades can be attributed to basal scRNA expression in the second layer of the circuit. These results suggest that engineered promoters capable of lower basal scRNA expression levels would minimize compression of activation cascade dynamic ranges.

Inducible CRISPRa by expressing MCP-SoxS from an inducible promoter

To provide an input for dynamic CRISPRa/i circuitry in *E. coli*, we chose to apply control over CRISPRa through inducible expression of the MCP-SoxS activator protein

(Figure 5B). We observed that output levels generated by CRISPRa were titratable through aTc induction of MCP-SoxS activator (Figure 5B, right). These output levels were similar to CRISPRa employing constitutive expression of MCP-SoxS (Figure 4B, left) as well as aTc induction of scRNA (Figure S9, left). Compared to CRISPRa with constitutively expressed activator and off/on target scRNA, the aTc-inducible system provided 40% lower basal levels and 16.5% lower activated levels of reporter expression (Figure 5B). A 40.4 ± 0.77 fold increase in expression was observed for the constitutive CRISPRa system supplied with on versus off-target scRNAs, whereas aTc induction of MCP-SoxS with on-target scRNA yielded a 56.3 ± 0.65 -fold increase. A similar, 43 fold increase was observed when comparing aTc inducible MCP-SoxS with on/ off-target scRNA (Figure S9, right). Together, these results establish aTc-inducible expression of the MCP-SoxS activator as a means of generating titratable levels of activation.

Level-Matching of Multi-Layer CRISPRa/i Circuits in *E. coli*

To understand the level-matching requirements of sgRNAs in multi-layer CRISPRa/i circuits in *E. coli*, we constructed an inducible CRISPR activation-repression cascade. Here, we titrated the CRISPRa input in the first layer and tuned both the expression characteristics of the promoters and the activities of sgRNAs in the second layer of the cascade. Tuning of CRISPRa inputs in the first layer of the cascade was provided by the previously described inducible MCP-SoxS activator system. In the second layer of the cascade, expression characteristics of promoters were tuned via modifications to the 5' sequence proximal to the minimal promoter, while sgRNA activities were modified through the use of 5' spacer truncations (Fontana et al., 2018b; Qi et al., 2013) (Figure 5C, Figure S11). When the J2 promoter was used to express RR2 sgRNA

390 targeting RFP (Fontana et al., 2018b), we observed 70% repression in the absence of activation (Figure 5C, left). A ~20-fold increase in sgRNA expression provided by
392 CRISPRa (Figure S10) resulted in an output dynamic range of 4-fold, spanning 23% of accessible expression levels. Decreasing the strength of CRISPRi via truncation of the
394 RR2 sgRNA spacer to 14 nucleotides decreased repression in the absence of activation of sgRNA expression to 20% as compared to an off-target control. However, truncated
396 guide RNAs were not able to achieve high levels of repression at maximal levels of activation (Figure 5C, center) resulting in a compressed output dynamic range of 2.5-fold,
398 spanning 48% of accessible expression levels. Tuning of sgRNA expression levels via modifications to the 5' sequence proximal to the minimal promoter resulted in 17% and
400 92% repression in the absence and presence of CRISPRa applied to RR2 sgRNA respectively, yielding an output dynamic range of ~10-fold, spanning 76% of the
402 accessible expression space (Figure 5C, right). Taken together, the inducible CRISPRa/i cascade and the CRISPRa cascade indicate that guide RNA expression levels produced
404 by CRISPRa are sufficient to saturate downstream layers of CRISPRa/i circuits, and that both CRISPRa and CRISPRi are highly sensitive to basal expression of sc/sgRNAs.

406 Tunability of CRISPRa/i Enables Interrogation of Complex Behavior in *E. coli*

The ability to tune CRISPRa/i circuits through both promoter engineering and guide
408 RNA truncations allows control over the abundance and strength of individual sc/sgRNAs. Such control enables independent tuning of nodes in multi-layer CRISPRa/i circuitry.
410 Paired with the ease of circuit construction, the CRISPRa/i system is suited for rapid circuit function interrogation. To showcase multi-guide tuning and circuit function

412 exploration, we constructed and characterized three different I1-FFLs in which application of repression by node Y is varied.

414 To enable observation of dynamic circuit behaviors and provide titratable levels of input activation, we chose to use inducible expression of the activator protein MCP-SoxS 416 (Figure 5B). To understand the effect of tuning actions on I1-FFL output, we constructed three different network topologies (Figure 6A. top), an I1-FFL, CRISPRa+i, and CRISPRa 418 with an off-target sgRNA. We compared the response of these three circuits to increasing levels of MCP-SoxS induction for three different tunings of sgRNA expression (Figure 6A, 420 bottom). To determine how much of the repression in the I1-FFL could be attributed to the basal expression from node Y, we compared the CRISPRa+i circuit to CRISPRa with 422 off-target sgRNA (Figure S12). We observed 7500 RFU from CRISPRa+i with off-target sgRNA. CRISPRa+i expressing full-length on-target sgRNA from the J2 promoter 424 provided 2200 and 71 RFU at 200 nM and 0 nM aTc, respectively. Thus, basal expression of the sgRNA from node Y has the effect of reducing, or compressing, the output range 426 of the MCP-SoxS CRISPRa titration by 70%. In an I1-FFL with the same sgRNA expression tuning, the output dynamic range was only 3-fold (Figure S12). We reasoned 428 that decreasing the strength of repression at node Z would reduce compression in the CRISPRa+i circuit and increase the dynamic range of the I1-FFL. Tuning repression 430 through sgRNA truncation to 14nt decreased compression of the output range by the CRISPRa+i circuit from 70% to 24%, and increased the output dynamic range of the I1-432 FFL from 3 fold to 15 fold (Figure S12). Tuning repression through modifications to the sequence 5’ of the minimal promoter expressing sgRNA resulted in 4% compression of 434 the output range by CRISPRa+i. However, counterintuitively the I1-FFL output dynamic

range was decreased to 2 fold with the same modifications to the sequence 5' of the minimal promoter expressing sgRNA (Figure S12). While the latter result was unexpected, the current suite of tuning actions available to the CRISPRa/i system in *E. coli* nonetheless does allow independent tuning of interactions between nodes in multi-guide circuits.

I1-FFLs are used in many naturally-occurring sensory systems as fold-change detectors to generate dynamic outputs determined by relative, as compared to absolute, differences in inputs to the system (Adler and Alon, 2018). Formally, fold-change detection (FCD) can be defined as a logarithmic relationship between inputs "I" and outputs "O", i.e., an input/output response curve, satisfying the equation $O = a \cdot ln(I) + b$. Theoretical work has shown that a transcriptional I1-FFL is capable of fold-change detection only under specific ratios of component expression levels and strengths (Goentoro et al., 2009). Experimentally, we can test for fold-change detection in these circuits by evaluating the variance explained by a logarithmic fit to outputs taken as a function of aTc-induced MCP-SoxS inputs. Consistent with expectation, we observed that only I1-FFLs with specific sgRNA tunings were capable of detecting fold changes of aTc over the linear range of MCP-SoxS induction (Figure 5B, right). Over this linear range, we observed an R^2 of 0.975 for a logarithmic fit between the inputs and outputs, as compared to an R^2 of 0.853 for a linear fit to the data for the I1-FFL with 20nt RR2 sgRNA expressed from the J2 promoter (Figure 6A, bottom right). Using these R^2 values we can calculate Akaike Information Criterion scores (AICc)(Methods 6) for each model. The relative likelihood of the data being described by a logarithmic as opposed to a linear model can then be computed using the AICc scores for each model. For the IFFL with RR2 sgRNA

458 expressed from the J2 promoter, we find that the logarithmic model is 197 times more likely to describe the data than a linear one. We can extend the test for fold-change 460 detection beyond the linear range of aTc-induced MCP-SoxS inputs by converting these inputs into the corresponding CRISPRa responses. We linearized the CRISPRa response 462 to aTc induction by dividing CRISPRa output levels at a given aTc induction level by the CRISPRa output at saturating concentrations of aTc. Plotting I1-FFL outputs against this 464 percent induction of the CRISPRa response provided an R^2 of 0.989 for a logarithmic fit, as opposed to only 0.896 for a linear fit resulting in a relative likelihood of 880 for the data 466 being described by a logarithmic model. By comparison, the corresponding CRISPRa+i circuit exhibited an R^2 for the input-output relationship of 0.852 for a logarithmic fit, and 468 0.997 for a linear fit resulting in a relative likelihood of less than .001 or, stated differently, that the linear model is at least 1000 times more likely than the logarithmic one to describe 470 the observed data (Figure S13). Taken together, these data show that CRISPRa/i circuits assembled into I1-FFLs can be tuned to achieve fold-change detection. This 472 demonstrated capacity of CRISPRa/i circuits to perform non-linear mapping between inputs and outputs expands the utility of the CRISPRa/i system, allowing complex 474 relationships to be encoded as network topologies.

To investigate the composability of CRISPRa/i controlled gene expression 476 dynamics, we tested three inducible circuits under continuous dilution: CRISPRa, an activation-repression cascade, and an I1-FFL tuned for fold-change detection (Figure 6B, 478 left). We observed increases in RFP/OD_{600} roughly 1 hr post induction for both the I1-FFL and CRISPRa corresponding to the action of the first layer in both circuits (Figure 6B, 480 right). For both the I1-FFL and the activation repression cascade, repression onset was

observed at ~5 hours, corresponding to the action of the second layer of each circuit. While I1-FFLs are recognized as a classic pulse generating circuit, in order to achieve fold-change detection gene expression pulses must display perfect adaptation (Adler and Alon, 2018; Goentoro et al., 2009), meaning gene expression must return to the basal level after completing the pulse. Here we observed an adaptive pulse of gene expression from the I1-FFL that starts at the baseline and ends at the baseline, corroborating the ability of the CRISPRa/i I1-IFFL to function as a fold-change detector. Overall, these results indicate that CRISPRa/i circuits are composable in that the observed dynamic behaviors can be understood from the functions of the parts.

Discussion

We have developed a set of components and a unifying framework for building dynamic CRISPRa/i gene regulatory networks that are scalable, composable, and tunable. These networks are built from CRISPRa/i nodes, which we define as transcriptional units that can be targeted for regulation by both CRISPRa and CRISPRi. The CRISPRa/i framework leverages an expandable set of synthetic promoters and orthogonal guide RNAs to specify arbitrary transcriptional regulatory topologies. The characteristics and limitations of the network are therefore determined by the properties of the constituent components. Particularly, promoter dynamic range and guide RNA function specify the transcriptional input-output relationship of each node. Understanding these relationships will be fundamental for building deep, wide, and fast regulatory networks.

We can estimate the upper bound for the maximum depth of activation cascades based on the observed cascade performance in this work. At present we observe 68% ± 18% signal propagation in a two layer activation cascade (Figure 2B). The total fraction of signal propagated in deeper cascades can be calculated by raising the fraction of signal propagated between two layers to the total number of internal layers in the cascade. This calculation indicates the current CRISPRa/i system in CFS can support cascades up to 6 layers deep before output activation ratios fall below 2.5 fold. Similarly, we achieve 68% ± 2.7% signal propagation in *E. coli* (Figure 4B), suggesting activation cascades of up to 6 layers could be built with the current implementation of the CRISPRa system *in-vivo*. With the components presented In this work, transcriptional activation generated by CRISPRa does not fully span the input dynamic range of sc/sgRNA expression in

downstream layers (Figure 2), resulting in degradation of signals as they are propagated through multi-layer CRISPRa/i circuits. The most general solution to increase fidelity of signal propagation in both CFS and *E. coli* is through engineering improved system components. Promoters with lower basal expression, leading to larger output dynamic ranges, would span a higher fraction of the input dynamic range of downstream nodes, resulting in less signal degradation between layers and deeper CRISPRa/i circuitry. We estimate that engineered promoters with a mere 5-fold increase in output dynamic range would allow CRISPRa-directed outputs to fully span the input dynamic range of sc/sgRNAs in downstream layers. In this system, modest improvements in signal propagation efficiency between layers would enable drastically deeper CRISPRa/i networks. For instance, increasing the fraction of signal propagated between layers by 12%, and output dynamic ranges by two fold would, in theory, enable cascades up to 14 layers deep before output activation ratios fall below 2.5 fold. While deeply layered cascades remain beyond the scope of current engineered regulatory networks, the large activation ratios and high fidelity of signal propagation observed in the CRISPRa/i system contribute to the robust operation of shallower networks.

Natural systems coordinate the expression of many outputs with few internal layers of computation using wide, highly-interconnected networks such as dense overlapping regulons (Rosenfeld and Alon, 2003; Shen-Orr et al., 2002; Thieffry et al., 1998). Our results indicate the CRISPRa/i system is well suited to design of wide control circuits for simultaneous and independent multi-gene regulation. In CFS, CRISPRa levels and kinetics are unchanged with respect to scRNA expression across at least an order of magnitude (Figure S6). Additionally, RFP expression levels are unchanged over a ~40-

fold range of dCas9 plasmid concentration (Figure S5). This indicates that the cell-free reaction has the resources to express high levels of scRNA and dCas9 without hindering system performance. Construction of two orthogonal I1-FFLs in the same reaction showcases the ability of CFS to harbour large circuits expressing many different sc/sgRNAs to execute multiple independent programs simultaneously. Likewise, *E. coli* are capable of expressing high levels of sc/sgRNA without experiencing growth defects or retroactivity due to guide RNA competition for dCas9 (Huang et al., 2021). We observe that modest overexpression of off-target scRNA has minimal effects on CRISPRa levels (Figure S15), consistent with recent modeling work suggesting favorable scaling for CRISPRa networks as compared to CRISPRi (Clamons and Murray, 2019). Taken together, these results suggest that the CRISPRa/i system could support the operation of programs containing up to 20 independent sc/sgRNAs with minimal impact to system performance in both CFS and *E. coli*. As CRISPRa/i circuits become larger, it may be necessary to incorporate improvements from the larger CRISPR community with our approach to overcome potential limitations imposed by expression burden and gRNA competition (Huang et al., 2021; Schmidt et al., 2021)

In both natural and engineered systems RNA-based regulatory approaches provide a means for fast and metabolically-efficient control of gene expression (Bobrovskyy and Vanderpool, 2013; Chappell et al., 2017; Stevens and Carothers, 2015; Takahashi et al., 2015; Westbrook et al., 2019). In multi-layer CRISPRa/i circuits, the speed of signal propagation is a tunable parameter depending on the relative expression levels of all components involved in both CFS and *E. coli*. Our analysis of relative production rates shows that guide RNA mediated information propagation through

internal layers of CRISPRa/i circuits can be fast (~30 mins/layer) (Figure S14) compared to the time required for initial expression of functional CRISPRa components. Under saturating expression of dCas9, sgRNA titrations reveal an initial 40-50 min delay to the onset of CRISPRi, which could be attributed to maturation of dCas9 and formation of active CRISPR complexes (Figure 1E, right). Likewise, at high expression levels of dCas9, we observe small differences in the timing of gene expression across a wide range of scRNA expression levels (Figure 1D; Figure S5, right). In CFS, CRISPRa controlled production rates reach steady state with respect to constitutive expression over the course of several hours (Figure 1D, right). In contrast, CRISPR activation generated by cascades can experience minimal additional delays compared to CRISPRa in a single layer (Figure S14). For many applications in biocomputing and metabolic control, successful operation is determined not only by the fidelity but also the speed at which information is propagated through the regulatory network. CFS containing pre-expressed dCas9 and activator protein could accelerate the onset of CRISPRa/i regulation in these systems. Extrapolation of these CFS observations to inform the speed of propagation in *E. coli* is difficult because *E. coli* experience dilution due to cell division as well as dCas9 eviction due to DNA replication. While the cell free experiments in this work were conducted as batch mode reactions, introduction of component turnover either through continuously diluted cell free reactions(Dubuc et al., 2019; Karzbrun et al., 2014; Niederholtmeyer et al., 2013), or programmable degradation(Garamella et al., 2016) presents a bridge between cellular and cell free settings. Future work to characterize the correspondence between tuning actions and the speed of signal propagation in cell free systems with component turnover could provide a promising test bed to inform the

predictive design of CRISPRa/i programs controlling gene expression timing in a cellular setting.

Given the previously demonstrated orthogonality of CRISPR-based regulation and the independence of CRISPRa/i nodes observed in this work, we find the CRISPRa/i system to be readily composable into larger motifs. CRISPRa/i circuits can be built by level-matching the response curves of different nodes, without complications arising from retroactivity and crosstalk. Indeed, upon construction of two orthogonal I1-FFLs in the same CFS reaction (Figure 3A), expression dynamics of the first I1-FFL were independent from both the presence and action of the second I1-FFL, and nearly identical to expression dynamics observed in isolation. Such orthogonality enables design of circuits with deterministic functions based solely on proper implementation of network topologies. In CFS, we built circuits capable of generating distinct gene expression profiles determined by the specific network topologies (Figure 3A). We showed that expression dynamics are tunable through component expression levels within a given circuit topology when there is overlap between upstream and downstream circuit layers (Figure 2E, 3B). In *E. coli*, dilution experiments revealed I1-FFL gene expression dynamics to be an almost perfect superposition of the dynamics of CRISPRa and an activation-repression cascade (Figure 6B). This composability is also captured by the relatively simple CFS CRISPRa/i model, which only specifies CRISPR complex assembly and node targeting reactions for each sc/sgRNA (Table S3). Paired with advancements in high-throughput component characterization in CFS and state-of-the-art modeling frameworks (Lehr et al., 2019; Moore et al., 2018; Poole et al., 2020), the CRISPRa/i

system presents a route towards scalable computer-aided design and implementation of dynamic gene regulatory networks in CFS.

Overall, this work establishes a paradigm in which CRISPRa/i system components can be easily combined to form scalable, dynamic gene regulatory networks in CFS and *E. coli*. The CRISPRa/i system has proven capable of building layered operations and simultaneously executing multiple regulatory programs without compromising guideRNA-encoded expression dynamics. The dynamic gene expression profiles arising from CRISPRa/i regulation are composable, in that network expression dynamics can be understood as the aggregate of the constituent components in both CFS and *E. coli* (Figure 3B, Figure 6B). These attributes allow rational design of CRISPRa/i circuits to tailor expression dynamics of multiple genes independently and simultaneously. We anticipate broad-ranging applications in engineered bacterial hosts, cell-free systems, and the next generation of artificial cells. Specifically, we foresee applications in metabolic engineering, with feedforward motifs providing time-ordered enzyme expression, and tunable delays enabling phenotype switching in multi-phase reactions. For biosensing applications, the scalable and versatile nature of the CRISPRa/i system will allow combinatorial logical responses and complex input-output relationships to be specified, increasing the ease of connecting sensing to reporting stages. Finally, as synthetic biology efforts transition from repurposing natural systems to the bottom-up construction of fully-artificial cells, the CRISPRa/i system could be a foundational technology capable of implementing the complex dynamic control of gene expression observed in nature, while remaining compact, robust, and engineerable.

Acknowledgments

We thank members of the Carothers and Zalatan groups for advice, materials, and comments on the manuscript. This work was supported by US National Science Foundation (NSF) Award CBET 1844152 (to J.M.C. and V.N.), NSF Award MCB 1817623 (to J.G.Z. and J.M.C.) and NSF Award EF-1935087 (to J.M.C.).

Author Contributions

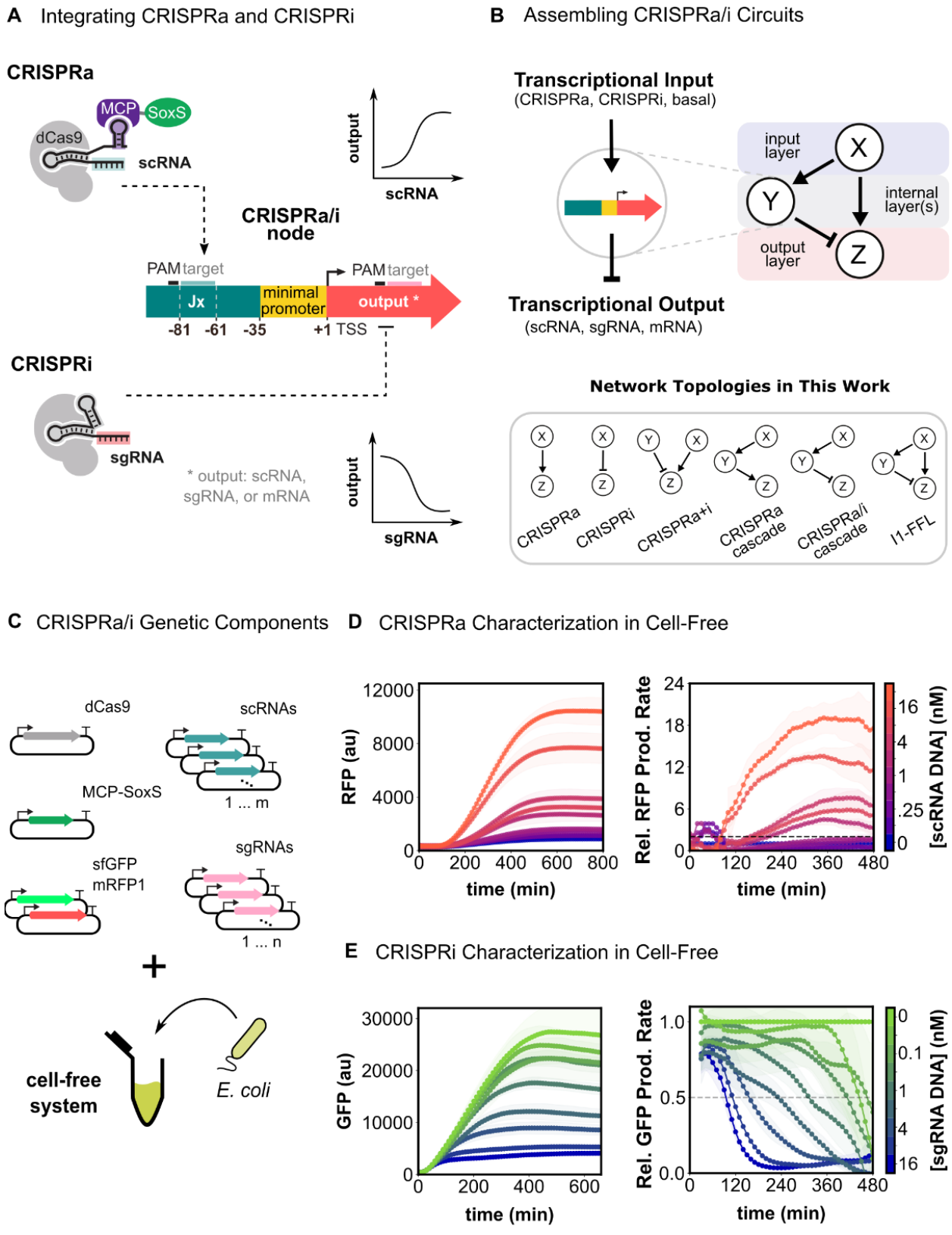
B.I.T., D.A.B., V.P.C. and J.M.C. designed the research, B.I.T., D.A.B., V.P.C., and A.K. performed experiments, B.I.T., C.K., and J.F. engineered components for CRISPRa, B.I.T., D.A.B., V.P.C., C.K. and A.K. analyzed the data. B.I.T., D.A.B., J.G.Z., and J.M.C. wrote the manuscript with input from all of the authors.

Declaration of Interests

The Noireaux laboratory receives research funds from Arbor Biosciences, a distributor of the myTXTL cell-free protein synthesis kit.

Figures

Figure 1: Combining CRISPRa with CRISPRi expands CRISPR circuit design space



A. Schematic of CRISPRa/i nodes. Modified guide RNAs (scaffold RNAs, or scRNAs) include a 3' MS2 hairpin to recruit a transcriptional activator fused to the MS2 coat protein (MCP). The expression of scRNAs targeted to an appropriate site upstream of the promoter results in CRISPR transcriptional activation (CRISPRa) from the node. The expression of a small guide RNAs (sgRNAs), which lack the MS2 hairpin, targeted to the coding sequence (CDS) results in CRISPR transcriptional repression (CRISPRi) from the node.

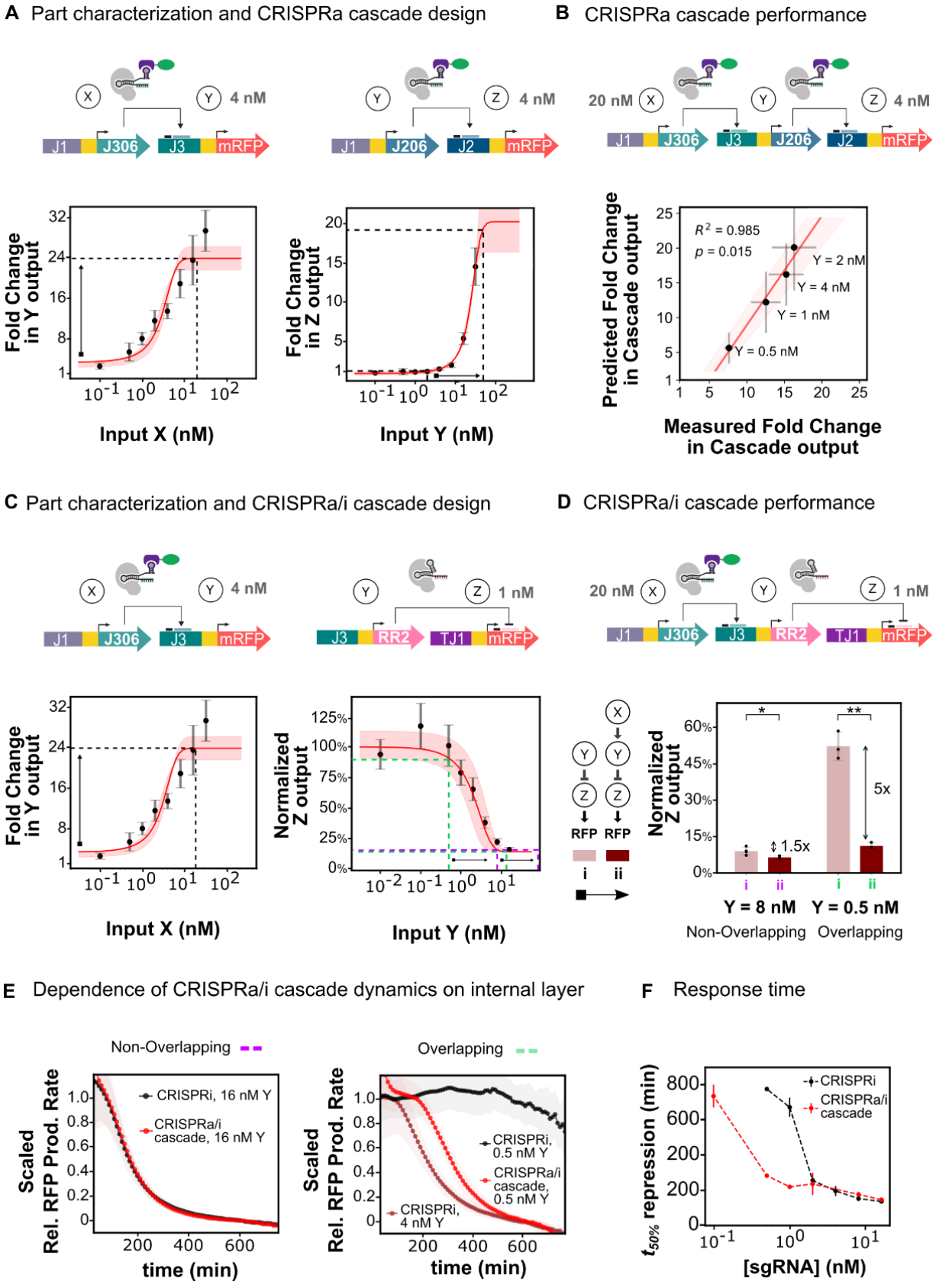
B. Schematic of CRISPRa/i circuits. CRISPRa/i nodes can be combined to form multi-layer CRISPRa/i networks when the guide RNA output generated by one node directs CRISPRa or CRISPRi at one or more other nodes in the network.

C. Schematic presentation of CRISPRa/i genetic components for use in a Cell-Free System (CFS). CRISPRa/i system components are encoded on individual plasmids and assembled into networks by mixing with *E. coli*-derived CFS.

D. Time course of CRISPRa in CFS. **Left:** CRISPRa-directed red fluorescent protein expression levels (mRFP1) from the J3 promoter are plotted as a function of time and J306 scRNA plasmid concentration. **Right:** Relative CRISPRa-directed RFP production rates (CRISPRa-directed production rates divided by unregulated production rates, Methods 5.1) are plotted as a function of time and J306 scRNA plasmid concentration. Dashed line represents 2-fold activation compared to the basal expression control (0 nM J306 scRNA plasmid). Values represent the mean ± standard deviation of three technical replicates.

E. Time course of CRISPRi in CFS. **Left:** CRISPRi-directed repression of green fluorescent protein (sfGFP) expression is plotted as a function of time and SF1 sgRNA plasmid concentration. SF1 targets the sfGFP CDS. **Right:** Relative GFP production rates are plotted as a function of time and SF1 sgRNA plasmid concentration. Dashed line represents 50% repression compared to the basal expression control (0 nM SF1 sgRNA plasmid). Values represent the mean ± standard deviation of three technical replicates.

Figure 2: Level-matching enables construction of multi-layer CRISPRa/i circuits in CFS



A. CRISPRa transcriptional input-output response curves. **Top:** Expression of J306 scRNA (X) directs CRISPRa from the J3 promoter (Y) producing an RFP output. Expression of J206 scRNA (Y) directs CRISPRa from the J2 promoter (Z), producing an RFP output. **Bottom:** Fold change in transcriptional output is plotted as a function of transcriptional input, specified by the scRNA plasmid concentration. Fold change is calculated as the ratio of RFP expression level at the reaction endpoint in the presence versus the absence of scRNA plasmid. Red line indicates a logistic fit to the data. **Bottom Right:** Pink and green dashed lines are guides showing the predicted effect of 25-fold increases in scRNA transcriptional input generated by CRISPRa in the first layer on the output of CRISPRa in the second layer of a two-layer activation cascade. Values represent the mean ± standard deviation of three technical replicates.

B. Two-layer CRISPRa cascade. **Top:** Expression of J306 scRNA at node X directs CRISPRa from the J3 promoter at node Y. Expression of J206 scRNA from the J3 promoter (Y) directs CRISPRa from the J2 promoter, which expresses mRFP1 (Z). **Bottom:** Scatter plot comparing measured and predicted fold change in cascade output. The measured fold change in cascade output is calculated as the ratio of measured RFP outputs with and without CRISPRa in the first layer of the circuit (Methods, 5.1), and are presented as the mean ± standard deviation of three technical replicates. Predicted cascade fold changes and uncertainties are calculated from the fits to the scRNA plasmid titrations shown in 2A. Statistical significance of non-zero slope was calculated using two-tailed unpaired Welch's *t* test.

C. CRISPRa/i transcriptional input-output response curves. **Top:** Expression of J306 scRNA (X) directs CRISPRa from the J3 promoter (Y), producing an RFP output.

Expression of the RR2 sgRNA (Y) directs CRISPRi of mRFP1 (Z). **Bottom:** Fold change in transcriptional output is plotted as a function of transcriptional input, specified by the scRNA and sgRNA plasmid concentration. Red line indicates a logistic fit to the data. **Bottom left:** Response curve for J306 scRNA presented as in 2A **Bottom Right:** CRISPRi data are represented as percent expression of a no repression control (Methods, 5.1). Pink and green dashed lines show the expected effect of a 25-fold increase in sgRNA transcription provided by CRISPRa in the first layer on the CRISPRi-directed output in the second layer of a two-layer CRISPR activation-repression cascade. Values represent the mean ± standard deviation of three technical replicates.

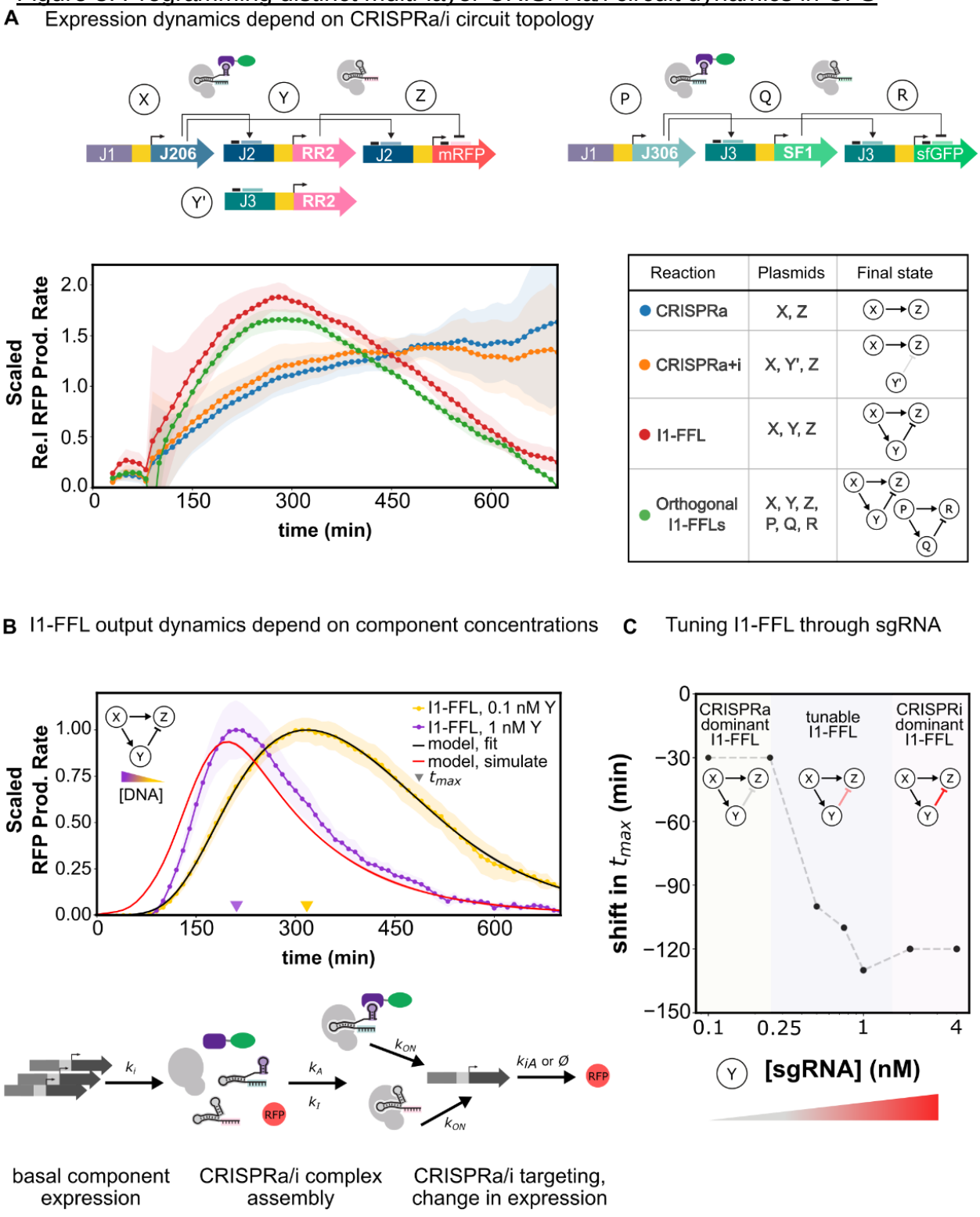
D. Two-layer CRISPRa/i cascade. **Top:** Expression of J306 scRNA (X) directs CRISPRa from the J3 promoter. RR2 sgRNA expression from the J3 promoter (Y) directs CRISPRi of mRFP1 (Z). **Bottom:** Percent of maximum expression is calculated as in Fig 2C comparing CRISPRi in one layer (I, light red) to two-layer CRISPR activation-repression cascades (II, dark red) at two different concentrations of sgRNA in the second layer. Values represent the mean ± standard deviation of three technical replicates. Asterisks indicate a statistically significant difference using a two-tailed unpaired Welch's *t* test (**p*-value < 0.05, **p-value < 0.01, ****p*-value < 0.001).

E. CRISPRi and CRISPRa/i circuit dynamics. **Left:** Gene expression over time for RFP controlled by CRISPRi and CRISPRa/i cascade at a concentration of plasmid Y that falls outside the range where input-output levels between the first and second layers overlap (no level-matching). Black line represents the scaled relative RFP production rate due to CRISPRi with 16 nM of plasmid encoding RR2 sgRNA while the red line

represents the scaled relative RFP production rate generated when CRISPRa is applied to the same concentration of sgRNA plasmid. **Right:** Comparison of CRISPRi and CRISPRa/i cascade at a concentration of plasmid Y that permits level-matching between the first and second layers of the circuit. Black line represents the scaled relative RFP production rate generated by CRISPRi with 0.5 nM of plasmid encoding RR2 sgRNA while the red line represents expression generated when CRISPRa is applied to the same concentration of sgRNA plasmid. The dark red line shows the scaled relative expression generated by CRISPRi with 4 nM of sgRNA plasmid and is provided as a point of reference. Data are presented as mean ± standard deviation of three technical replicates.

F. Time to 50% repression is plotted against the concentration of sgRNA plasmid for both CRISPRi (black line) and a CRISPRa/i cascade (red line). Showing that when output/input ranges of the first and second layer overlap, multi layer CRISPRa/i circuits can be used to tune the timing of gene expression. Data are presented as mean ± standard deviation of three technical replicates.

Figure 3: Programming distinct multi-layer CRISPRa/i circuit dynamics in CFS

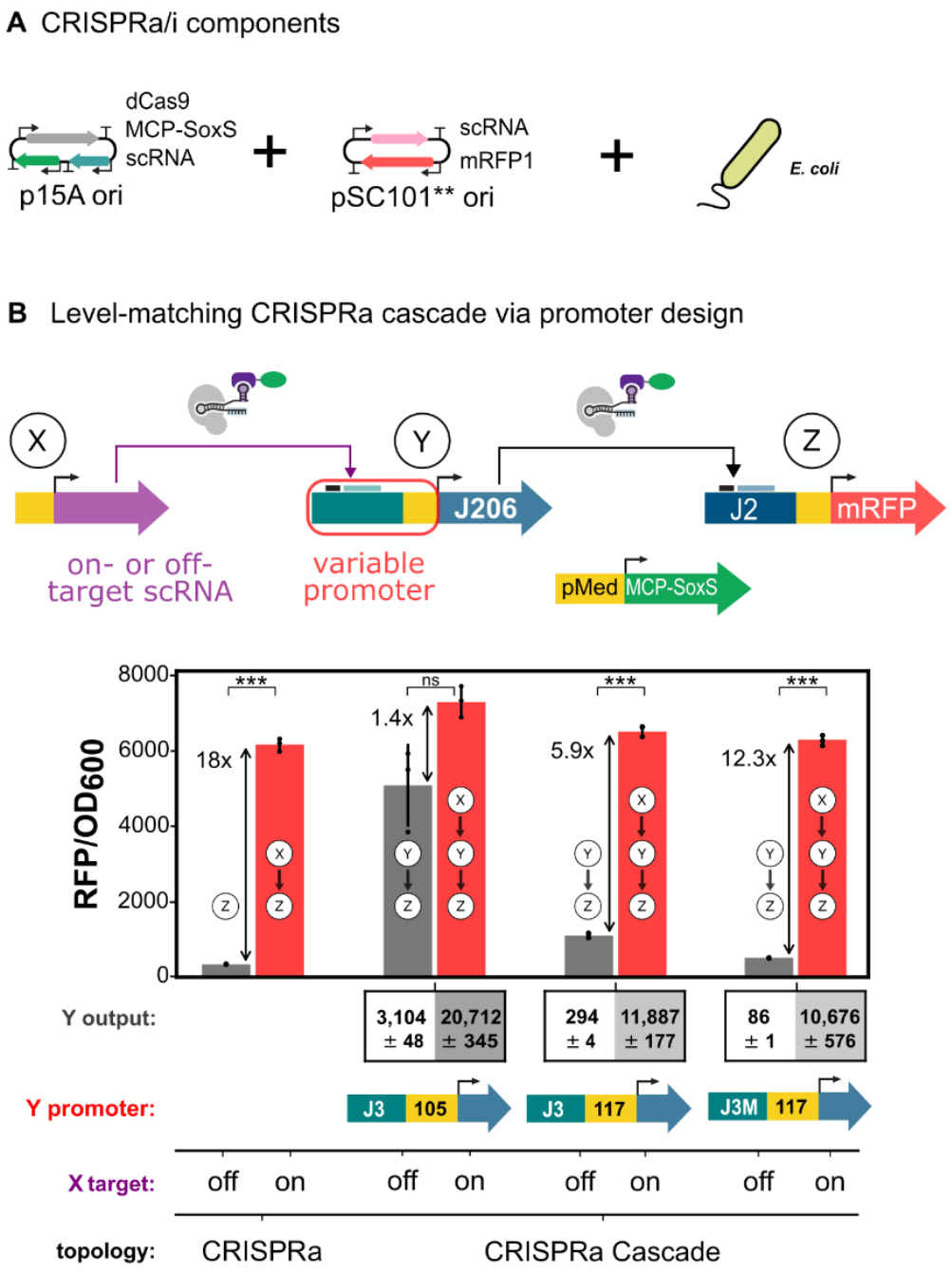


A. Time course for four different CRISPRa/i networks including orthogonal type 1 incoherent feed forward loops (I1-FFL) operating in CFS. Normalized relative RFP production rates are plotted as a function of network topology, as indicated. Outputs are scaled by the respective endpoints to fit on a common axis. Values represent the mean ± standard deviation of three technical replicates.

B. Time course for I1-FFL variants shows dependence of output expression dynamics on sgRNA concentration. Scaled RFP production rates are plotted as a function of time for two different RR2 sgRNA plasmid concentrations; 0.1 nM and 1 nM shown in yellow and purple respectively. Fits to the illustrated I1-FFL are used to predict expression dynamics as a function of sgRNA concentration. The model fit to the measured 0.1 nM data is shown in black, and the model prediction for expression dynamics from an I1-FFL with 1 nM sgRNA expressing plasmid is shown in red. Yellow and purple triangles along the x-axis denote the time of maximal expression of the pulse generated by the I1-FFL. Values represent the mean ± standard deviation of three technical replicates.

C. Shift in time of the I1-FFL expression maxima depends on sgRNA plasmid concentration, which controls the strength of the CRISPRi connection between nodes Y and Z. The time at which the maximum output rate is achieved (t_{max}) is plotted as a function of RR2 sgRNA plasmid concentration, which is expressed from node Y and acts on node Z.

Figure 4: Level-matching of CRISPRa cascades in *E. coli*

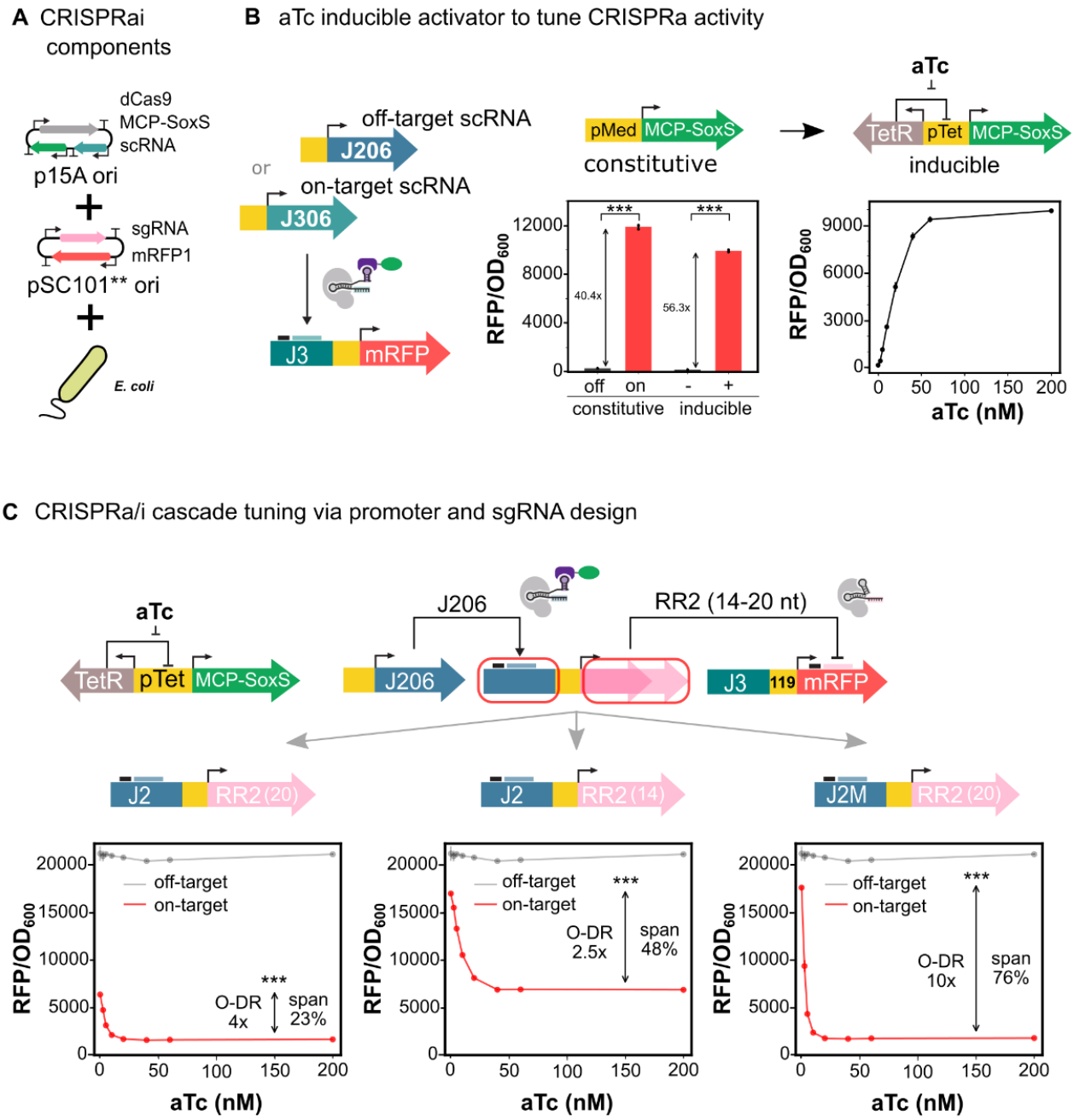


A. Schematic of CRISPRa genetic components and assembly to form CRISPRa circuits in *E. coli*. Circuits are assembled by transformation of different combinations of plasmids into *E. coli*. Activator, dCas9, and scRNAs in the first layer of a circuit are expressed from a p15A plasmid while reporters and scRNAs

in the second layer of a circuit are expressed from a pSC101** origin of replication plasmid. Data are collected in *E. coli* MG1655 grown overnight at 37 °C with shaking in EZ MOPS with 0.2% glucose and appropriate antibiotic selection.

B. **Top:** schematic of CRISPRa cascade. Tuning actions are applied by changing expression characteristics of the activatable promoter in the second layer of the circuit. **Bottom:** CRISPRa on the J2 promoter is compared to the output at node Z of the activation cascade with (red bars) and without (grey bars) input provided by node X. Y output denotes the expression levels obtained at node Y in the presence and absence of input activation from X for the tuning variant provided by Y promoter. Data are represented as mean ± standard deviation of RFP/OD_{600}. Asterisks indicate a statistically significant difference using a two-tailed unpaired Welch's *t* test (**p*-value < 0.05, ***p*-value < 0.01, ****p*-value < 0.001).

Figure 5: Level-matching of CRISPRa/i cascades in E. coli with titratable input



A. Schematic of CRISPRa/i genetic components and assembly to form CRISPRa/i circuits in *E. coli*. Circuits are assembled by transformation of different combinations of plasmids into *E. coli*. Activator, dCas9, and scRNAs in the first layer of a circuit are expressed from a p15A plasmid while reporters, and

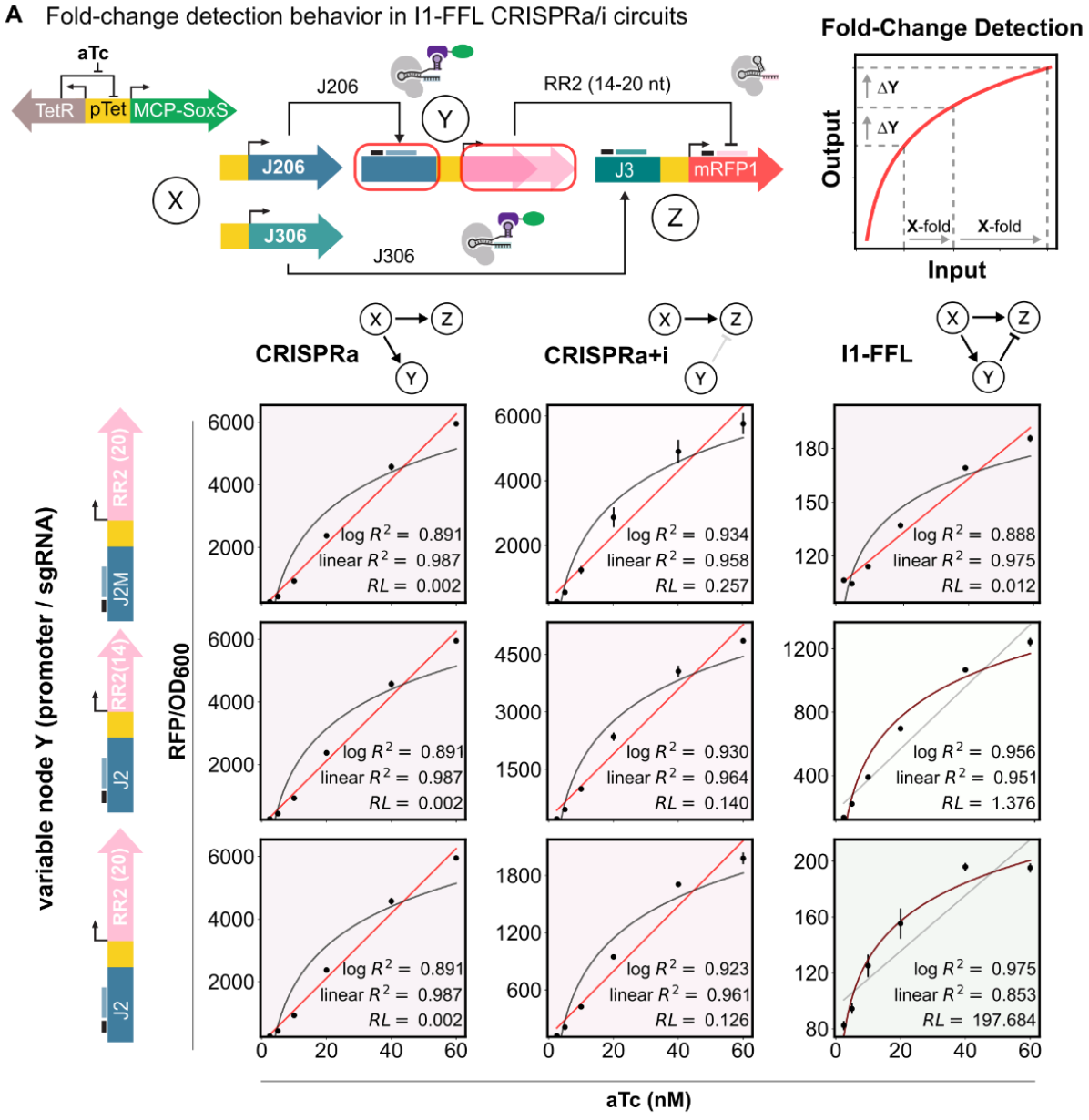
sc/sgRNAs in the second layer of a circuit are expressed from a pSC101** origin of replication plasmid. Inducible CRISPRa/i data are collected in *E. coli* MG1655 grown O/N, diluted 1:40, grown for 3 hours, and diluted 1:40 into inducing media, all at 37 °C with shaking in EZ MOPS with 0.2% glucose and appropriate antibiotic selection.

B. Dependence of CRISPRa on aTc-induced MCP-SoxS expression. **Left:** Circuit schematic showing how inducible CRISPRa is provided by pTet controlled expression of MCP-SoxS activator. For both constitutive and inducible CRISPRa, J306 scRNA directs dCas9 to the J3 promoter. Constitutive CRISPRa is compared with on/off target scRNA while inducible CRISPRa is compared ± inducer with on-target scRNA. **Center:** mRFP1 output of aTc-inducible CRISPRa ± aTc is compared to constitutive CRISPRa with on/off-target scRNA. Data are represented as mean of measured RFP/OD_{600} ± standard deviation of three biological replicates. **Right:** aTc induction of activator protein provides titratable levels of activation. Data are represented as the mean ± standard deviation of measured RFP fluorescence/OD_{600} for three biological replicates. Asterisks indicate a statistically significant difference using a two-tailed unpaired Welch's *t* test (**p*-value < 0.05, ***p*-value < 0.01, ****p*-value < 0.001).

C. Two-layer CRISPRa/i cascade output dependence on promoter and sgRNA. **Top:** schematic of aTc-inducible CRISPRa activating expression of sgRNA targeting mRFP1 for CRISPRi. Level-matching is achieved through engineering basal expression characteristics of the activatable promoter driving sgRNA expression as well as through modifications of sgRNA activity via 5' sgRNA truncations that

reduce sgRNA spacer-target DNA complementary from 20 to 14 nucleotides. **Bottom:** CRISPRi response to increasing levels of activation of the promoter driving sgRNA expression for three different tunings. On-target sgRNA is plotted in red while off-target sgRNA is plotted in grey. Data are represented as mean RFP/OD_{600} ± standard deviation. Calculations for Span and Output Dynamic range (O-DR) can be found in Methods 5.2. Asterisks indicate a statistically significant difference using a two-tailed unpaired Welch's *t* test (**p*-value < 0.05, ***p*-value < 0.01, ****p*-value < 0.001).

806 Figure 6: Tunable and composable dynamic CRISPRa/i circuits in *E. coli*



A. Fold-change detection behavior in I1-FFL CRISPRa/i circuits. **Top:** circuit schematic of CRISPRa/i incoherent type 1 feed forward loop (I1-FFL) implemented in *E. coli*. Here tuning of repression by node Y is achieved by changes to the transcriptional properties of the promoter at node Y or through modifications to the sgRNA expressed at node Y. aTc induction of MCP-SoxS activator serves as a titratable input to the system. At right, a graphical definition of fold-change detection: the same fold-change in input "X" produces the same change in output "Y". **Bottom:** aTc response for 9 combinations of network topology and sgRNA tuning showing that connectivity between nodes and tuning specifies input/output relationship of complex CRISPRa/i circuits. Network topologies are indicated along the top while tuning variants are represented along the left. Greyed out arrow in CRISPRa+i indicates on-target sgRNA from an unactivated promoter. I1-FFLs must have specific component relationships to achieve fold-change detection seen here as logarithmic input/output response curves. Data are represented as the mean of measured RFP/OD_{600} ± standard deviation of three biological replicates. Data for off-target sgRNA is the same in all three plots and is collected in the J2 context (bottom left). Statistical significance was assessed by calculating the relative likelihood (*RL*) of the log fit best describing the data as compared to the linear fit, Methods 5.

B. CRISPRa/i circuit expression dynamics were measured in a continuous dilution experiment. **Left:** schematic of continuous dilution experimental design and procedure. **Right:** Expression dynamics of CRISPRa, CRISPRa/i cascade, and an I-1FFL tuned for fold-change detection are compared. Data are presented as relative RFP/OD_{600} ± standard deviation of three biological replicates (Methods, 5.2). For

CRISPRa and the IFFL, RFP is expressed from the standard J3 reporter with a BBa_J23117 minimal promoter while for the CRISPRa/i cascade the strong minimal promoter BBa_J23119 is used to enable observation of inhibition in this assay. Relative RFP/OD_{600} scales data to the respective output response ranges to place them on a common scale for comparison of output expression dynamics.

STAR Methods

Resource Availability

Lead Contact:

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, James M. Carothers (jcaroth@uw.edu).

Materials Availability:

Plasmids generated in this study have been deposited at Addgene (https://www.addgene.org/plasmids/articles/28222993/).

Data and Code Availability:

- All CFS and *E. coli* data have been deposited at https://github.com/carothersresearch/CRISPRai_Circuits_2021 and are publicly available as of the date of publication. DOIs are listed in the key resources table.
- All models and scripts used in this work have been deposited at https://github.com/carothersresearch/CRISPRai_Circuits_2021 and are publicly available as of the date of publication.DOIs are listed in the key resources table.
- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Method Details

Plasmid Design and Preparation

Plasmid design and sequencing analyses were performed using Benchling sequence designer. All PCR amplification of plasmids and fragments used Phusion DNA polymerase in GC buffer. Primers were synthesized by IDT and resuspended into nuclease-free water. All PCR reactions were treated with DpnI for longer than 1 hour and

purified using Qiagen gel extraction kits. Plasmid assembly was achieved using 5X In-Fusion HD mastermix (Takara). Assembled plasmids were transformed into chemically competent NEB Turbo *E. coli* and plated onto LB-agar plates with either 100 µg/mL carbenicillin or 25 µg/mL chloramphenicol and grown overnight ~16 hours at 37 °C. Single colonies were picked from plates and grown overnight in LB shaking at 37 °C with appropriate concentrations of relevant antibiotics. Plasmids were isolated from subcultures using a DNA miniprep kit (QIAprep Spin Miniprep Kit) and Sanger sequenced (Genewiz inc.) to identify correctly assembled plasmids. Plasmids intended for use in CFS were grown in culture volumes ~20 mL to ensure adequate yields for multiple cell-free reactions. Plasmids intended for cell-free expression were further purified using a PCR purification kit (Invitrogen PureLink, Cat. K310001), eluted into nuclease-free water. Plasmid concentrations were quantified via spectrophotometry (Nanodrop 2000c, Cat. ND-2000C).

Cell-Free System Preparation

The cell-free system was prepared according to previously published procedures (Garamella et al., 2016; Sun et al., 2013). The cell-free system used for an experiment was thawed on ice and pooled into a 1.5 ml eppendorf tube, vortexed, and spun-down using a mini benchtop centrifuge to ensure homogeneity across samples.

Cell-Free Gene Expression Reaction

Cell-free gene expression reactions were assembled on ice from the CFS and purified DNA. A master mix with common plasmids across reactions was prepared, and 1.5 µL per reaction allocated into PCR tubes. Plasmids which were varied across reactions were added in the remaining 1 µL. The CFS was pipette mixed and added to each PCR tube

in 7.5 µL for a final volume of 10 µL. PCR tubes were vortexed, spun-down using a mini benchtop centrifuge, and placed on ice. Triplicates of 2.5 µL for each reaction were pipetted into individual wells of a 96-well V-bottom plate (Costar, Cat. 3363). The plate was sealed (Costar, Cat. 3080) and analyzed on a BioTek Synergy HTX plate reader at 29 °C. sfGFP fluorescence (ex. 485 nm, em. 528 nm) and mRFP1 fluorescence (ex. 540 nm, em. 600 nm) of cell free reactions were measured every 10 min from the bottom of the plate. All reactions were run in batch mode.

***E. coli* Experiments**

Circuits were assembled in *E. coli* through transformation of plasmid pairs. In all *E. coli* experiments CRISPRa system components (dCas9, MCP-SoxS, input sc/sgRNAs) are located on a p15A ori (copy number ~10) plasmid while sc/sgRNAs forming the second layer of a circuit were cloned into the reporter plasmid with a pSC101** ori (copy number ~5) due to the size difference between vectors. The p15A plasmid used in constitutive CRISPRa experiments was pCK085. aTc-inducible CRISPRa experiments use pJF182 in which pTet controls expression of TetR and MCP-SoxS. Plate reader measurements were conducted using a BioTek Synergy HTX with a black flat bottom plate (Ref# 3631) using 100 µL of culture.

Constitutive CRISPRa experiments:

Endpoint CRISPRa experiments are conducted using constitutive expression of all CRISPRa components (pCK085). Circuits were assembled by transformation of CRISPRa and reporter plasmids into chemically competent MG1655 E. coli. Transformed *E. coli* were outgrown for 1 hour shaking at 37 °C and plated onto LB-agar with carbenicillin and chloramphenicol. Plates were grown overnight at 37 °C. Experiments

were conducted by picking three individual colonies into 400 µL Teknova EZ-RDM with 0.2% glucose and appropriate antibiotics in 96 well plates (round 2 ml), covering with breathable membrane (Breathe Easier cat# Z763624) and shaking overnight at 37 °C at 1200 RPM on a Heidolph titramax 1000.

aTc-inducible CRISPRa experiments

For inducible CRISPRa experiments expression of activator protein MCP-SoxS was controlled by pTet. Upon addition of aTc (anhydrotetracycline) to media, pTet becomes de-repressed which enables titratable expression of MCP-SoxS activator. As above, circuits are assembled by transformation of CRISPRa and reporter plasmids into chemically competent MG1655 *E. coli*. Transformed *E. coli* were outgrown for 1 hour shaking at 37 °C and plated onto LB-agar with carbenicillin and chloramphenicol. Plates were grown overnight at 37 °C. Experiments were conducted by picking three individual colonies into 400 uL EZ-RDM with 0.2% glucose and appropriate antibiotics in 96 well plates (2ml). Cultures are covered with breathable membrane and left shaking overnight at 37 °C at 1200 RPM on a Heidolph titramax 1000. Overnight cultures are subsequently diluted 1:40 into a fresh plate of EZ-RDM and incubated at 37 °C shaking at 1200 RPM. Before exiting exponential phase (~3 hours) cultures are diluted 1:40 into a fresh plate of EZ-RDM and supplemented with appropriate concentrations of aTc. These cultures are again covered with a breathable membrane, incubated in the dark at 37 °C shaking at 1200 RPM, and grown overnight ~18 hours. Measurements are conducted in Costar 96 well black flat bottom plates in 100 uL culture volume.

Continuous dilution *E. coli* experiments

Strains used in continuous dilution experiments were constructed as above through double transformation with appropriate selection on plates. For CRISPRa and I1-FFL experiments MG1655 was used while for CRISPRa/i cascades, MG1655 with an integrated reporter (J23119-RFP) was used (Fontana et al., 2018b). Throughout the experiment all steps including liquid handling were conducted at 37 °C. Individual colonies were picked and grown overnight in 500 µL of LB with appropriate antibiotics shaking at 1200 RPM in a 96 well deep well plate. In the morning cultures were diluted 1:50 into RZ-RDM with 0.2% glucose and grown as before for one hour at a total volume of 300 µl. At one hour, 100 µL of culture was removed and measured while 100 µL fresh pre-warmed media was added to each culture. This process was repeated every 20 minutes until OD_{600} stabilized, usually within ~1-2 hours. Once ODs stabilized, aTc inducer was added to cultures to a final concentration of 50 nM. Every 20 minutes for the remainder of the experiment (~8 hours), 100 µL culture was removed and measured, while the culture was resupplied with 100 µL pre-warmed inducing media.

CFS CRISPRa/i Modeling

The CFS CRISPRa/i model was defined as a series of first order chemical reactions for protein and guide RNA production, CRISPR complex assembly, and DNA targeting. The model was implemented using the text-based model definition language Antimony for Python 3.7 (Table S3). Here we also present the underlying system of differential equations governing the model.

The general CRISPRa/i model is based on the different transcriptional states a CRISPRa/i node can take. At the DNA level, each node D_i can either be unregulated, activated by CRISPRa (D_i^A), or repressed by CIRSPRi (D_i^R). We differentiate between two different repression states based on whether CRISPRa is also bound. The ODEs defining the change in concentration over time between node states are therefore:

$$\frac{dD_i}{dt} = -k_{on}\left[C_i^A \cdot D_i \ + C_i^R \cdot D_i \ \right]$$

$$\frac{dD_i^A}{dt} = k_{on}\left[C_i^A \cdot D_i \ - C_i^R \cdot D_i^A \right]$$

$$\frac{dD_i^R}{dt} = \frac{dD_i^{R_1}}{dt} + \frac{dD_i^{R_2}}{dt} = k_{on}\left[C_i^R \cdot D_i \ - C_i^A \cdot D_i^{R_1} \right] + k_{on}\left[C_i^R \cdot D_i^A + C_i^A \cdot D_i^{R_1} \right]$$

$$= k_{on} \cdot C_i^R\left[D_i \ + D_i^A \right]$$

Where C_i^A and C_i^R represent the concentration of free CRISPRa and CRISPRi complexes targeting node i, respectively, and k_{on} is the rate of association to the DNA, assuming a one-step irreversible reaction.

The concentration of free CRISPR complexes is determined by the association rate of the different components (dCas9 C, scRNA G_i^A, and MCP-SoxS S for CRISPRa; dCas9 C and

sgRNA G_i^R for CRISPRi) as well as the rate at which they bind to the respective targets, namely:

$$\frac{dC_i^A}{dt} = k_A \cdot C \cdot G_i^A \cdot S - k_{on} \cdot C_i^A\left[D_i \ + D_i^{R_1}\right] = \mathrm{K}_A - k_{on} \cdot C_i^A\left[D_i \ + D_i^{R_1}\right]$$

$$\frac{dC_i^R}{dt} = k_R \cdot C \cdot G_i^R - k_{on} \cdot C_i^R\left[D_i \ + D_i^A\right] = \mathrm{K}_R - k_{on} \cdot C_i^R\left[D_i \ + D_i^A\right]$$

The concentration of the different protein and RNA species is based on their expression and interactions with other components. For simplicity, here protein transcription and translation are lumped.

$$\frac{dP_i^{off}}{dt} = k_T \cdot D_i \ + k_{TA} \cdot D_i^A - k_M \cdot P_i^{off} - \Gamma, \quad \Gamma = \begin{cases} \mathrm{K}_A + \mathrm{K}_R, & if \ \ P_i = C \\ \mathrm{K}_A, & if \ \ P_i = S \\ \emptyset, & otherwise \end{cases}$$

$$\frac{dP_i}{dt} = k_M \cdot P_i^{off};$$

$$\frac{dR_i}{dt} = k_T \cdot D_i \ + k_{TA} \cdot D_i^A - k_D \cdot R_i \ - \Upsilon, \quad \Upsilon = \begin{cases} \mathrm{K}_A, & if \ \ R_i = G_i^A \\ \mathrm{K}_R, & if \ \ R_i = G_i^R \end{cases}$$

Parameters were generated from log-uniform distributions spanning 4 orders of magnitude based around literature values or best estimates. Models with different parameters were evaluated using Tellurium (Choi et al., 2018; Medley et al., 2018), and the model outputs were processed in the same manner as the experimental data. The processed outputs were then fit to scaled RFP production rates of experimental data by minimizing the cumulative point-wise squared error using the Nelder-Mead simplex algorithm.

Quantification and Statistical Analysis

Cell-free data analysis

Production Rate:

Throughout this work, we define production rate as:

$$\dot{B}^{\alpha}(t) = \frac{dB^{\alpha}}{dt} = \frac{B^{\alpha}(t+30) - B^{\alpha}(t)}{30}$$

where:

B is the measured quantity: RFP or GFP

α specifies the circuit topology and relevant plasmid concentrations

Relative Production Rates:

Relative production rates of CRISPRa mediated outputs were calculated as the ratio of CRISPRa mediated production rates divided by unregulated production rates. For CRISPRa the contribution due to unregulated basal expression was subtracted from measured output levels due to CRISPRa. This was done to isolate the timing of CRISPRa mediated gene expression from the comparatively early contribution of basal expression, and to allow observation of CRISPRa mediated gene expression dynamics under conditions where basal expression of reporter constructs dominates. Throughout this work, relative production rates are abbreviated as Rel. RFP Prod. Rate, or Rel. GFP Prod. Rate, and are calculated as:

$$\dot{B}_{\Gamma}^{\alpha}(t) = \frac{\dot{B}^{\alpha}(t) - \dot{B}^{\Gamma}(t)}{\dot{B}^{\Gamma}(t)}$$

where:

α is a specific CRISPRa/i circuit

Γ is constitutive expression

For CRISPRi mediated relative production rates there is no subtraction of basal expression and relative production rates are provided as:

$$\dot{B}_{\Gamma}^{\alpha}(t) = \frac{\dot{B}^{\alpha}(t)}{\dot{B}^{\Gamma}(t)}$$

Fold change in cascade output:

In Figure 2B, fold change in cascade output was calculated as the ratio of RFP values generated by the CRISPRa cascade compared to CRISPRa in a single layer with the same concentration of scRNA expressing plasmid Y.

$$FC_{\beta}^{\alpha}(y) = \frac{B^{\alpha}(t = t_{max})/B^{\Gamma}(t = t_{max})}{B^{\beta}(t = t_{max})/B^{\Gamma}(t = t_{max})} = \frac{B^{\alpha}(t = t_{max})}{B^{\beta}(t = t_{max})}$$

where:

B is measured RFP

α is CRISPRa cascade, with *y* nM scRNA Y

β is CRISPRa, with *y* nM scRNA Y

t_{max} is the endpoint time of the cell free reaction

Predicted fold change in cascade output:

Predicted fold change in cascade output plotted in Figure 2B was generated using the fits to scRNA titration curves provided in 2A (red line). Predicted fold changes provided by CRISPRa at a given concentration of scRNA expressing plasmid *y* are denoted: $\widehat{FC}_{\Gamma}^{\alpha}(y)$. In this experiment, CRISPRa in the first layer of the cascade (2A, left) generated by x nM of scRNA at node X is expected to direct a 24-fold increase in transcription from the target. Thus, for a given concentration of scRNA expressing plasmid *y* at node Y in the second layer of the cascade, predicted fold change in cascade output is calculated as the ratio of fold changes predicted by the fit in the right panel of 2A at a scRNA concentration of 24*y*, and 1*y*.

$$\widehat{FC}_{\beta}^{\alpha}(y) = \frac{\widehat{FC}_{\Gamma}^{\alpha}(24y)}{FC_{\Gamma}^{\alpha}(y)}$$

where:

is CRISPRa cascade, with *y* nM scRNA Y

is CRISPRa, with *y* nM scRNA Y

Normalized Z output:

In Figure 2C&D normalized Z output is defined as the percent expression provided by CRISPRi or a CRISPRa/i cascade compared to an unregulated, constitutive expression control.

$$nB_{\Gamma}^{\alpha}(y) = 100 \cdot \frac{B^{\alpha}(t = t_{max})}{B^{\Gamma}(t = t_{max})}$$

where:

B is measured RFP

α is CRISPRi or a CRISPRa/i cascade with y nM sgRNA plasmid at node Y

Scaled relative production rates:

In Figure 2E relative production rates were scaled by the observed production rate of the reaction at 40 mins to place curves on a common scale before the maturation of dCas9 and the onset of CRISPRa/i control.

$$s\dot{B}_{\Gamma}^{\alpha}(t) = \frac{\dot{B}_{\Gamma}^{\alpha}(t)}{\dot{B}_{\Gamma}^{\alpha}(t = 40)}$$

where:

$\dot{B}_{\Gamma}^{\alpha}$ is relative RFP production rate

α is CRISPRi or CRISPRa/i cascade with y nM sgRNA plasmid at node Y

Time to 50% repression:

Time to 50% repression in Figure 2F and S4 was defined as the time at which relative production rates due to CRISPRa/i control first reached 50% of the initial 50 min value i.e., before maturation of dCas9 and the onset of CRISPRi.

$$t = t_{50\%} \; when \; \frac{\dot{B}_{\Gamma}^{\alpha}(t)}{\dot{B}_{\Gamma}^{\alpha}(t=50)} \geq .5$$

where:

$\dot{B}_{\Gamma}^{\alpha}$ is relative RFP production rate

α is CRISPRi, or CRISPRa/i cascade with y nM sgRNA plasmid at node Y

Normalized relative RFP production rates:

In figure 3A relative RFP production rates are scaled by their respective endpoints to place curves on a common scale:

$$n\dot{B}_{\Gamma}^{\alpha}(t) = \frac{\dot{B}_{\Gamma}^{\alpha}(t)}{B^{\alpha}(t=t_{max})}$$

where:

B is measured RFP

$\dot{B}_{\Gamma}^{\alpha}$ is relative RFP production rate

α is CRISPRa, CRISPRa+i, or an I1-FFL

Scaled RFP Production Rate:

In Figure 3B RFP production rates were scaled by their respective maxima, placing both curves on a common scale, allowing comparison of time dynamics.

$$s\dot{B}^{\alpha}(t) = \frac{\dot{B}^{\alpha}(t)}{\max\left(\dot{B}^{\alpha}(t)\right)}$$

where:

$\dot{B}(t)$ is RFP production rate

α is an I1-FFL with *y* = .1 nM or *y* = 1 nM sgRNA plasmid at node Y

Percent signal propagated by CRISPRa cascade:

The percent signal propagated by the CRISPRa cascade in CFS was calculated as the fold change in cascade output ± input divided by the fold change provided by CRISPRa in the input layer.

$$SP(y) = 100 \cdot \frac{FC_{\beta_1}^{\alpha_1}(y)}{FC_{\Gamma}^{\beta_2}(y)}$$

where:

α_1 is CRISPRa cascade with y nM of scRNA at node Y

β_1 is CRISPRa with y nM of scRNA at node Y

β_2 is CRISPRa with x nM of scRNA at node X

Time to 2x activation:

Time to 2-fold activation was defined as the time at which relative production rates are expected to first exceed 1, i.e., when CRISPRa mediated production rates first achieve twice that of unregulated expression. Time to 2-fold activation was calculated as the mean ± standard deviation from linear fits to relative RFP production rates for three technical replicates. Linear fits were calculated over a 1 hour interval between 80 and 160 mins corresponding to the initial linear increase in relative RFP production rates provided by CRISPRa.

$$t = t_{2x} \ when \ \dot{B}_{\Gamma}^{\alpha}(t) \geq 1.0$$

where:

$\dot{B(t)}$ is relative RFP production rate

α is CRISPRa with x nM scRNA at node X

<u>Time to half maximal expression:</u>

In figure S5B, time to half maximal expression was calculated differently for CRISPRa and CRISPRi. For CRISPRa the contribution due to unregulated basal expression was subtracted from measured RFP levels due to CRISPRa. This was done to isolate the

timing of CRISPRa mediated gene expression from the comparatively early contribution of leak, and to allow observation of CRISPRa mediated gene expression dynamics under conditions where basal expression of reporter constructs dominates. The time to half max is denoted as $t_{1/2}$.

For CRISPRa this was defined as the first time point at which

$$t = t_{1/2} when \frac{B^{\alpha}(t) - B^{\Gamma}(t)}{B^{\alpha}(t = t_{max}) - B^{\Gamma}(t = t_{max})} \geq .5$$

where:

B is measured RFP

α is CRISPRa with x nM scRNA at node X

For CRISPRi there was no subtraction of basal expression thus time to half maximal expression for CRISPRi is given by the first time point at which

$$t = t_{1/2} when \frac{B^{\alpha}(t)}{B^{\alpha}(t = t_{max})} \geq .5$$

where:

B is measured RFP

α is CRISPRi with x nM scRNA at node X

Relative fold change:

In Figure S12, the gene expression dynamics of CRISPRa and a CRISPRa cascade are compared to visualize the speed of signal propagation in multi-layer CRISPR circuits. For these data, fold change is a function of time:

$$rFC_{\Gamma}^{\alpha}(t,y) = \frac{FC_{\Gamma}^{\alpha}(t,y)}{FC_{\Gamma}^{\alpha}(t = t_{max},\ y)}$$

where:

B is measured RFP

α is CRISPRa or CRISPRa cascade with y nM scRNA plasmid at node Y

y is the concentration of scRNA expressing plasmid in the final layer of the circuit

Normalized fluorescence

In figure S7, RFP and GFP fluorescence were normalized by the response range for each fluorescent protein to lie on a common scale between 0 and 1.

$$n\boldsymbol{B}^{\boldsymbol{\alpha}} = \frac{\boldsymbol{B}^{\boldsymbol{\alpha}}(t = t_{max}) - \min\left(\boldsymbol{B}^{\boldsymbol{\alpha}}(t = t_{max})\right)}{\max\left(\boldsymbol{B}^{\boldsymbol{\alpha}}(t = t_{max})\right) - \min\left(\boldsymbol{B}^{\boldsymbol{\alpha}}(t = t_{max})\right)}$$

where:

$\boldsymbol{B}^{\boldsymbol{\alpha}}$ is a vector containing average RFP or GFP fluorescence for all tested conditions (CRISPRa, CRISPRa+i, I1-FFL, orthogonal I1-FFLs)

Number of possible network topologies

The number of possible network topologies presented in Figure S1 was calculated as:

$$T(M, N) = (M + 1) \cdot (N \cdot (N - 1))$$

where:

M is the number of modes of regulation

M = 1 for CRISPRi alone

M = 2 for both CRISPRa and CRISPRi together

N is the number of nodes in the network

<u>*E. coli* data analysis</u>

Throughout this work all measured RFP levels in *E. coli* were normalized by measured OD_{600}. Data are plotted as the mean RFP/OD_{600} ± standard deviation of three biological replicates with appropriate propagation of uncertainties.

Span:

In Figure 5C span was calculated in percent as the range of RFP expression values provided by CRISPRa/i cascades (0, 200nM aTc) divided by RFP expression levels obtained in the presence of off-target sgRNA.

$$S = \frac{B^{\alpha_1} - B^{\alpha_2}}{B^{\beta_1}}$$

where:

B is RFP/OD_{600} measured at endpoint

α_1 is CRISPRa/i cascade with 0 nM aTc induction

α_2 is CRISPRa/i cascade with 200 nM aTc induction

β_1 is CRISPRa/I cascade with off target sgRNA

Output dynamic range:

Output dynamic range was calculated as the ratio of measured CRISPRa/i cascade RFP outputs at 0nM aTc and 200nM aTc. The inverse quantity was used for I1-FFL output dynamic range.

$$O\text{-}DR = \frac{B^{\alpha_1}}{B^{\alpha_2}}$$

where:

B is RFP/OD_{600} measured at endpoint

α_1 is CRISPRa/i cascade with 0 nM aTc induction

α_2 is CRISPRa/i cascade with 200 nM aTc induction

Relative RFP/OD$_{600}$:

Relative RFP/OD$_{600}$ presented in Figure 6B was calculated from raw RFP/OD$_{600}$ data by subtraction of the minimum observed RFP/OD$_{600}$ value post induction, and scaling outputs by their maximum observed values to place all curves on a scale between zero and one.

$$rB^{\alpha}(t) = \frac{B^{\alpha}(t) - \min\left(B^{\alpha}(t > t_0)\right)}{\max\left(B^{\alpha}(t > t_0)\right)}$$

where:

B is RFP/OD$_{600}$ measured at endpoint

α is CRISPRa, CRISPRa/i cascade, or I1-FFl with 50 nM aTc induction

t_0 is the time at which aTc is added to the culture

Leak:

In Figure S10 leak was calculated as the percent reduction of measured CRISPRa RFP levels at 200nM aTc induction, due to unactivated sgRNA expression.

$$L = 100 \cdot \frac{B^{\alpha} - B^{\beta}}{B^{\alpha}}$$

where:

B is RFP/OD600 measured at endpoint

α is CRISPRa at 200 nM aTc

β is CRISPRa+i at 200 nM aTc

Output Range Compression:

Output range compression was defined as the output range of a CRISPRa/i circuit divided by the accessible output range. Here, output range of a CRISPRa/i cascade was defined as:

$$OR = B^{\alpha_1} - B^{\alpha_2}$$

While output range of CRISPRa was defined as:

$$OR = B^{\alpha_2} - B^{\alpha_1}$$

In a CRISPRa/i cascade, the accessible output range was calculated as the measured fluorescence provided by the CRISPRa/i cascade with an off target sgRNA

$$A = B^{\beta_1}$$

For I1-FFLs, the accessible output range was defined as the output range of the corresponding circuit with an off-target sgRNA, at 200 nM aTc and 0 nM aTc

$$A = B^{\beta_2} - B^{\beta_1}$$

While for CRISPRa+i, the accessible output range was defined as the output range with an off-target scRNA directing CRISPRa to activate sgRNA expression, at 200 nM aTc and 0 nM aTc

$$A = B^{\beta_2} - B^{\beta_1}$$

Thus, output range compression of CRISPRa/i cascades was defined in percent as:

$$ORC = 100 \cdot \left(1 - \frac{OR}{A}\right) = 100 \cdot \left(1 - \frac{B^{\alpha_1} - B^{\alpha_2}}{B^{\beta_1}}\right)$$

While output range compression of CRISPRa+I was defined as:

$$ORC = 100 \cdot \left(1 - \frac{OR}{A}\right) = 100 \cdot \left(1 - \frac{B^{\alpha_2} - B^{\alpha_1}}{B^{\beta_2} - B^{\beta_1}}\right)$$

where:

B is RFP/OD_{600} measured at endpoint

α_1 is CRISPRa+i, at 0 nM aTc

α_2 is CRISPRa+i, at 200 nM aTc

β_1 is CRISPRa+i, with off-target sgRNA at 0 nM aTc

β_2 is CRISPRa+i, with off-target sgRNA at 200 nM aTc

Percent CRISPRa induction:

Percent CRISPRa induction presented in Figure S11 was calculated by dividing measured RFP values obtained from CRISPRa by the measured RFP value provided by CRISPRa at maximal, saturating levels of aTc induction (200nM).

$$\%I(x) = 100 \cdot \frac{B^{\alpha_1}}{B^{\alpha_2}}$$

where:

B is RFP/OD_{600} measured at endpoint

α_1 is CRISPRa at *x* nM aTc

α_2 is CRISPRa at 200 nM aTc

Statistics

Statistical significance was calculated using two-tailed unpaired Welch's *t* tests. When comparing two different models (Figure 6A, S13), statistical significance was assessed by calculating the relative likelihood (*RL*) between a logarithmic model and a linear model,

based on the Akaike Information Criterion with small sample correction for each model (Akaike, 1974; Burnham and Anderson, 2002; Cavanaugh, 1997):

$$RL = exp\left(\frac{AIC_c(M_1) - AIC_c(M_2)}{2}\right)$$

$$AIC_c = 2 \cdot k + n \cdot log\left(\frac{\sum Res^2}{2}\right) + \frac{2 \cdot k \cdot (k+1)}{n - k - 1}$$

where *n* denotes the sample size, *k* denotes the number of parameters, *Res* the residuals, and M_1 and M_2 correspond to the linear and logarithmic models, respectively. Here the use of the AICc provides a second order estimation of information loss for each model. In addition to goodness of fit, the AICc takes into account the underlying model complexity and the size of the data set to penalize overfitting. For a given dataset, the model which minimizes the AICc can be said to minimize information loss. Within this paradigm, the relative likelihood is interpreted as the probability that M_2 minimizes information loss as compared to M_1.

References

Adamala, K.P., Martin-Alarcon, D.A., Guthrie-Honea, K.R., and Boyden, E.S. (2017). Engineering genetic circuit interactions within and between synthetic minimal cells. Nat. Chem. *9*, 431–439.

Adler, M., and Alon, U. (2018). Fold-change detection in biological systems. Curr. Opin. Syst. Biol. *8*, 81–89.

Akaike, H. (1974). A new look at the statistical model identification. IEEE Trans. Autom. Control *19*, 716–723.

Alon, U. (2007). Network motifs: theory and experimental approaches. Nat. Rev. Genet. *8*, 450–461.

Banerjee, D., Eng, T., Lau, A.K., Sasaki, Y., Wang, B., Chen, Y., Prahl, J.-P., Singan, V.R., Herbert, R.A., Liu, Y., et al. (2020). Genome-scale metabolic rewiring improves titers rates and yields of the non-native product indigoidine at scale. Nat. Commun. *11*, 5385.

Bartoli, V., di Bernardo, M., and Gorochowski, T.E. (2020). Self-adaptive biosystems through tunable genetic parts and circuits. Curr. Opin. Syst. Biol. *24*, 78–85.

Bikard, D., Jiang, W., Samai, P., Hochschild, A., Zhang, F., and Marraffini, L.A. (2013). Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res. *41*, 7429–7437.

Bobrovskyy, M., and Vanderpool, C.K. (2013). Regulation of Bacterial Metabolism by Small RNAs Using Diverse Mechanisms. Annu. Rev. Genet. *47*, 209–232.

Brockman, I.M., and Prather, K.L.J. (2015). Dynamic metabolic engineering: New strategies for developing responsive cell factories. Biotechnol. J. *10*, 1360–1369.

Brophy, J.A.N., and Voigt, C.A. (2014). Principles of Genetic Circuit Design. Nat. Methods *11*, 508–520.

Burnham, K.P., and Anderson, D.R. (2002). Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach (New York: Springer-Verlag).

Cavanaugh, J.E. (1997). Unifying the derivations for the Akaike and corrected Akaike information criteria. Stat. Probab. Lett. *33*, 201–208.

Chappell, J., Westbrook, A., Verosloff, M., and Lucks, J.B. (2017). Computational design of small transcription activating RNAs for versatile and dynamic gene regulation. Nat. Commun. *8*, 1051.

Clamons, S., and Murray, R. (2019). Modeling predicts that CRISPR-based activators, unlike CRISPR-based repressors, scale well with increasing gRNA competition and dCas9 bottlenecking. BioRxiv 719278.

Dinh, C.V., and Prather, K.L. (2020). Layered and multi-input autonomous dynamic control strategies for metabolic engineering. Curr. Opin. Biotechnol. *65*, 156–162.

Dinh, C.V., and Prather, K.L.J. (2019). Development of an autonomous and bifunctional quorum-sensing circuit for metabolic flux control in engineered *Escherichia coli*. Proc. Natl. Acad. Sci. *116*, 25562–25568.

Dong, C., Fontana, J., Patel, A., Carothers, J.M., and Zalatan, J.G. (2018). Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria. Nat. Commun. *9*.

Dubuc, E., Pieters, P.A., van der Linden, A.J., van Hest, J.C., Huck, W.T., and de Greef, T.F. (2019). Cell-free microcompartmentalised transcription–translation for the prototyping of synthetic communication networks. Curr. Opin. Biotechnol. *58*, 72–80.

Dudley, Q.M., Karim, A.S., and Jewett, M.C. (2015). Cell-free metabolic engineering: Biomanufacturing beyond the cell. Biotechnol. J. *10*, 69–82.

English, M.A., Gayet, R.V., and Collins, J.J. (2021). Designing Biological Circuits: Synthetic Biology Within the Operon Model and Beyond. Annu. Rev. Biochem. *90*, null.

Fontana, J., Voje, W.E., Zalatan, J.G., and Carothers, J.M. (2018a). Prospects for engineering dynamic CRISPR–Cas transcriptional circuits to improve bioproduction. J. Ind. Microbiol. Biotechnol. *45*, 481–490.

Fontana, J., Dong, C., Ham, J.Y., Zalatan, J.G., and Carothers, J.M. (2018b). Regulated Expression of sgRNAs Tunes CRISPRi in E. coli. Biotechnol. J. *13*, 1800069.

Fontana, J., Dong, C., Kiattisewee, C., Chavali, V.P., Tickman, B.I., Carothers, J.M., and Zalatan, J.G. (2020). Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements. Nat. Commun. *11*, 1618.

Gander, M.W., Vrana, J.D., Voje, W.E., Carothers, J.M., and Klavins, E. (2017). Digital logic circuits in yeast with CRISPR-dCas9 NOR gates. Nat. Commun. *8*, 15459.

Garamella, J., Marshall, R., Rustad, M., and Noireaux, V. (2016). The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. ACS Synth. Biol. *5*, 344–355.

Goentoro, L., Shoval, O., Kirschner, M., and Alon, U. (2009). The incoherent feedforward loop can provide fold-change detection in gene regulation. Mol. Cell *36*, 894–899.

Huang, H.-H., Bellato, M., Qian, Y., Cárdenas, P., Pasotti, L., Magni, P., and Del Vecchio, D. (2021). dCas9 regulator to neutralize competition in CRISPRi circuits. Nat. Commun. *12*, 1692.

Jeong, D., Klocke, M., Agarwal, S., Kim, J., Choi, S., Franco, E., and Kim, J. (2019). Cell-Free Synthetic Biology Platform for Engineering Synthetic Biological Circuits and Systems. Methods Protoc. *2*.

Kaplan, S., Bren, A., Dekel, E., and Alon, U. (2008). The incoherent feed-forward loop

can generate non-monotonic input functions for genes. Mol. Syst. Biol. *4*.

Karim, A.S., Dudley, Q.M., and Jewett, M.C. (2016). Cell-Free Synthetic Systems for Metabolic Engineering and Biosynthetic Pathway Prototyping. In Industrial Biotechnology, (John Wiley & Sons, Ltd), pp. 125–148.

Karzbrun, E., Tayar, A.M., Noireaux, V., and Bar-Ziv, R.H. (2014). Programmable on-chip DNA compartments as artificial cells. Science *345*, 829–832.

Kiattisewee, C., Dong, C., Fontana, J., Sugianto, W., Peralta-Yahya, P., Carothers, J.M., and Zalatan, J.G. (2021). Portable bacterial CRISPR transcriptional activation enables metabolic engineering in Pseudomonas putida. Metab. Eng. *66*, 283–295.

Landberg, J., Wright, N.R., Wulff, T., Herrgård, M.J., and Nielsen, A.T. (2020). CRISPR interference of nucleotide biosynthesis improves production of a single-domain antibody in Escherichia coli. Biotechnol. Bioeng. *117*, 3835–3848.

Lehr, F.-X., Hanst, M., Vogel, M., Kremer, J., Göringer, H.U., Suess, B., and Koeppl, H. (2019). Cell-Free Prototyping of AND-Logic Gates Based on Heterogeneous RNA Activators. ACS Synth. Biol. *8*, 2163–2173.

Liu, Y., Wan, X., and Wang, B. (2019). Engineered CRISPRa enables programmable eukaryote-like gene activation in bacteria. Nat. Commun. *10*, 1–16.

Lucks, J.B., Qi, L., Whitaker, W.R., and Arkin, A.P. (2008). Toward scalable parts families for predictable design of biological circuits. Curr. Opin. Microbiol. *11*, 567–573.

Mangan, S., and Alon, U. (2003). Structure and function of the feed-forward loop network motif. Proc. Natl. Acad. Sci. *100*, 11980–11985.

Marshall, R., and Noireaux, V. (2018). Synthetic Biology with an All E. coli TXTL System: Quantitative Characterization of Regulatory Elements and Gene Circuits. In Synthetic Biology: Methods and Protocols, J.C. Braman, ed. (New York, NY: Springer New York), pp. 61–93.

Marshall, R., Maxwell, C.S., Collins, S.P., Jacobsen, T., Luo, M.L., Begemann, M.B., Gray, B.N., January, E., Singer, A., He, Y., et al. (2018). Rapid and Scalable Characterization of CRISPR Technologies Using an E. coli Cell-Free Transcription-Translation System. Mol. Cell *69*, 146-157.e3.

McDaniel, R., and Weiss, R. (2005). Advances in synthetic biology: on the path from prototypes to applications. Curr. Opin. Biotechnol. *16*, 476–483.

Moore, S.J., MacDonald, J.T., Wienecke, S., Ishwarbhai, A., Tsipa, A., Aw, R., Kylilis, N., Bell, D.J., McClymont, D.W., Jensen, K., et al. (2018). Rapid acquisition and model-based analysis of cell-free transcription–translation reactions from nonmodel bacteria. Proc. Natl. Acad. Sci. U. S. A. *115*, E4340–E4349.

Niederholtmeyer, H., Stepanova, V., and Maerkl, S.J. (2013). Implementation of cell-free biological networks at steady state. Proc. Natl. Acad. Sci. *110*, 15985–15990.

Nielsen, A.A., and Voigt, C.A. (2014). Multi-input CRISPR/Cas genetic circuits that interface host regulatory networks. Mol. Syst. Biol. *10*.

Nielsen, A.A.K., Der, B.S., Shin, J., Vaidyanathan, P., Paralanov, V., Strychalski, E.A., Ross, D., Densmore, D., and Voigt, C.A. (2016). Genetic circuit design automation. Science *352*.

Poole, W., Pandey, A., Shur, A., Tuza, Z.A., and Murray, R.M. (2020). BioCRNpyler: Compiling Chemical Reaction Networks from Biomolecular Parts in Diverse Contexts (Synthetic Biology).

Qi, L.S., Larson, M.H., Gilbert, L.A., Doudna, J.A., Weissman, J.S., Arkin, A.P., and Lim, W.A. (2013). Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. Cell *152*, 1173–1183.

Qian, Y., Huang, H.-H., Jiménez, J.I., and Del Vecchio, D. (2017). Resource Competition Shapes the Response of Genetic Circuits. ACS Synth. Biol. *6*, 1263–1272.

Reis, A.C., Halper, S.M., Vezeau, G.E., Cetnar, D.P., Hossain, A., Clauer, P.R., and Salis, H.M. (2019). Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays. Nat. Biotechnol. *37*, 1294–1301.

Rosenfeld, N., and Alon, U. (2003). Response Delays and the Structure of Transcription Networks. J. Mol. Biol. *329*, 645–654.

Santos-Moreno, J., and Schaerli, Y. (2020). CRISPR-based gene expression control for synthetic gene circuits. Biochem. Soc. Trans. *48*, 1979–1993.

Santos-Moreno, J., Tasiudi, E., Stelling, J., and Schaerli, Y. (2020). Multistable and dynamic CRISPRi-based synthetic circuits. Nat. Commun. *11*, 2746.

Schmidt, M.J., Gupta, A., Bednarski, C., Gehrig-Giannini, S., Richter, F., Pitzler, C., Gamalinda, M., Galonska, C., Takeuchi, R., Wang, K., et al. (2021). Improved CRISPR genome editing using small highly active and specific engineered RNA-guided nucleases. Nat. Commun. *12*, 4219.

Shen-Orr, S.S., Milo, R., Mangan, S., and Alon, U. (2002). Network motifs in the transcriptional regulation network of Escherichia coli. Nat. Genet. *31*, 64–68.

Shin, J., and Noireaux, V. (2012). An E. coli Cell-Free Expression Toolbox: Application to Synthetic Gene Circuits and Artificial Cells. ACS Synth. Biol. *1*, 29–41.

Stevens, J.T., and Carothers, J.M. (2015). Designing RNA-Based Genetic Control Systems for Efficient Production from Engineered Metabolic Pathways. ACS Synth. Biol. *4*, 107–115.

Sun, Z.Z., Hayes, C.A., Shin, J., Caschera, F., Murray, R.M., and Noireaux, V. (2013). Protocols for implementing an Escherichia coli based TX-TL cell-free expression system for synthetic biology. J. Vis. Exp. JoVE e50762.

Takahashi, M.K., Hayes, C.A., Chappell, J., Sun, Z.Z., Murray, R.M., Noireaux, V., and Lucks, J.B. (2015). Characterizing and prototyping genetic networks with cell-free transcription–translation reactions. Methods *86*, 60–72.

Tan, S.-I., and Ng, I.-S. (2021). CRISPRi-Mediated NIMPLY Logic Gate for Fine-Tuning the Whole-Cell Sensing toward Simple Urine Glucose Detection. ACS Synth. Biol. *10*, 412–421.

Thieffry, D., Huerta, A.M., Pérez-Rueda, E., and Collado-Vides, J. (1998). From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in Escherichia coli. BioEssays *20*, 433–440.

Tian, J., Yang, G., Gu, Y., Sun, X., Lu, Y., and Jiang, W. (2020). Developing an endogenous quorum-sensing based CRISPRi circuit for autonomous and tunable dynamic regulation of multiple targets in Streptomyces. Nucleic Acids Res. *48*, 8188–8202.

Tian, T., Kang, J.W., Kang, A., and Lee, T.S. (2019). Redirecting Metabolic Flux via Combinatorial Multiplex CRISPRi-Mediated Repression for Isopentenol Production in Escherichia coli. ACS Synth. Biol. *8*, 391–402.

Wan, X., Volpetti, F., Petrova, E., French, C., Maerkl, S.J., and Wang, B. (2019). Cascaded amplifying circuits enable ultrasensitive cellular sensors for toxic metals. Nat. Chem. Biol. *15*, 540–548.

Wang, Y.-H., Wei, K.Y., and Smolke, C.D. (2013). Synthetic biology: advancing the design of diverse genetic systems. Annu. Rev. Chem. Biomol. Eng. *4*, 69–102.

Westbrook, A., Tang, X., Marshall, R., Maxwell, C.S., Chappell, J., Agrawal, D.K., Dunlop, M.J., Noireaux, V., Beisel, C.L., Lucks, J., et al. (2019). Distinct timescales of RNA regulators enable the construction of a genetic pulse generator. Biotechnol. Bioeng. *116*, 1139–1151.

Wu, F., Shim, J., Gong, T., and Tan, C. (2020a). Orthogonal tuning of gene expression noise using CRISPR–Cas. Nucleic Acids Res. *48*, e76–e76.

Wu, Y., Chen, T., Liu, Y., Tian, R., Lv, X., Li, J., Du, G., Chen, J., Ledesma-Amaro, R., and Liu, L. (2020b). Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in Bacillus subtilis. Nucleic Acids Res. *48*, 996–1009.

Xiang, Y., Dalchau, N., and Wang, B. (2018). Scaling up genetic circuit design for cellular computing: advances and prospects. Nat. Comput. *17*, 833–853.

Zalatan, J.G., Lee, M.E., Almeida, R., Gilbert, L.A., Whitehead, E.H., La Russa, M., Tsai, J.C., Weissman, J.S., Dueber, J.E., Qi, L.S., et al. (2015). Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. Cell *160*, 339–350.

Chapter 3:

Engineering Activatable Promoters for Scalable and Multi-Input CRISPRa/i Circuits

Diego Alba Burbano[+,1,2], Ryan Cardiff[+1], Benjamin I. Tickman[1,2], Cholpisit Kiattisewee[1], Cassandra Maranas[1], Jesse G. Zalatan*[1,3], James M. Carothers*[,1,2]

1: Molecular Engineering & Sciences Institute
and Center for Synthetic Biology
University of Washington
Seattle, WA 98195
United States

2: Department of Chemical Engineering
University of Washington
Seattle, WA 98195
United States

3: Department of Chemistry
University of Washington
Seattle, WA 98195
United States

+: These authors contributed equally

*: Corresponding authors
zalatan@uw.edu
206-543-1670

jcaroth@uw.edu
206-221-4902

Published on *PNAS* on July 18th, 2023.
DOI:/10.1073/pnas.2220358120

Classification: Biological Sciences, Applied Biological Sciences

Keywords: Activatable promoter engineering, CRISPR activation, gene regulatory networks, genetic circuits, cell-free, optogenetics

Significance statement

Engineered gene regulatory networks in cell-free systems hold great promise for investigating the limits of biological information processing and for developing next-generation biosensing and bioproduction platforms. We address the challenge of engineering gene regulatory networks with dynamically-specified expression levels at many different targets. The work described here enables new classes of deep, wide, and multi-input CRISPR-based gene regulatory networks with complexities approaching those found in nature.

Abstract

Dynamic, multi-input gene regulatory networks are ubiquitous in nature. Multi-layer CRISPR-based genetic circuits hold great promise for building gene regulatory networks akin to those found in naturally-occurring biological systems. We develop an approach for creating high-performing activatable promoters that can be assembled into deep, wide, and multi-input CRISPR-activation and -interference (CRISPRa/i) gene regulatory networks. By integrating sequence-based design and *in-vivo* high-throughput screening, we design activatable promoters that achieve up to 1000-fold dynamic range in an *E. coli*-based cell-free system. These new components enable CRISPRa gene regulatory networks that are six layers deep and four branches wide. We show the generalizability of the promoter engineering workflow by improving the dynamic range of the light-dependent EL222 optogenetic system from 6-fold to 34-fold. Additionally, high dynamic range promoters enable conditional CRISPRa systems mediated by small molecules and protein-protein interactions. We apply these tools to build input-responsive CRISPRa/i gene regulatory networks, including feedback loops, logic gates, multi-layer cascades, and dynamic pulse modulators. Our work provides a generalizable approach for the design of high dynamic range activatable promoters and enables new classes of gene regulatory functions in cell-free systems.

Introduction

Natural biological systems employ complex gene regulatory networks (GRNs) to sense diverse environmental cues and respond to them through the coordinated expression of multiple genes [1]–[3]. Cell-free systems (CFS) have emerged as an attractive chassis for building synthetic biological systems as they allow for rapid prototyping of genetic parts and circuits [4]–[8]. To build increasingly complex cell-free systems that can sense and respond to diverse inputs, new approaches for increasing the capabilities of synthetic GRNs are needed [9]–[12]. The ability to construct complex GRNs will enable investigating the limits of biological circuit design [10], [13], building multiplexed molecular biosensors [14]–[16], deploying on-demand bioproduction platforms [17]–[19], and the bottom-up construction of synthetic cells [20]–[23].

CRISPR-Cas transcriptional regulation has proven a promising framework for building sophisticated genetic circuits across a variety of biological systems [24]–[29]. Transcriptional units containing target sequences for CRISPR activation (CRISPRa) and/or CRISPR interference (CRISPRi), termed CRISPRa/i nodes, can be assembled into circuits with network topologies specified by a set of guide RNAs (gRNAs). Experimental and theoretical analysis indicates that the CRISPRa/i system is well suited to design deep and wide control circuits, defined as circuits containing two or more internal nodes connected in series or parallel through orthogonal gRNAs [30]–[32]. Large CRISPRi GRNs with up to 7 sgRNAs have been constructed in yeast by implementing low leak promoters and high dynamic range repressors [33]. In *E. coli*-based CFS, CRISPRi-based genetic control is well established [29], [34], and CRISPRa has recently been incorporated [32], greatly expanding the circuit design space. Integration of

CRISPRa and CRISPRi has enabled two-layer activation and activation-repression cascades, as well as two orthogonal I1-FFLs in the same CFS. However, CRISPRa circuitry is limited by a lack of promoter-gRNA pairs that can be interconnected with minimal signal degradation [32]. Hence, a generalizable approach for engineering activatable promoters with low basal and high activated expression levels would significantly improve CRISPRa and enable the assembly of complex, input-responsive synthetic GRNs.

Promoter engineering efforts have traditionally focused on designing constitutive and inducible promoters with predictable expression characteristics [35], [36]. Tuning the strength of constitutive promoters involves designing promoter sequences that maintain either weak or strong RNA polymerase (RNAP) recruitment to the promoter [37], [38]. Inducible promoters contain recognition sites for transcriptional activators or repressors that modulate transcriptional levels upon binding [39], [40]. Efforts to develop inducible transcription systems with high dynamic ranges have relied on engineering de-repression based systems such as the Lac or Tet promoters [40]–[44], largely due to the difficulty of rationally designing activatable promoters [45], [46]. For effective activation, RNAP recruitment to the promoter should be weak in the absence of an activator, however transcription initiation should be strong upon activator-mediated RNAP recruitment [39], [45]. Hence, higher dynamic ranges with CRISPRa and other transcriptional activation systems can be achieved through promoter sequence engineering to tune RNAP interactions.

We develop an approach integrating sequence-based design and *in-vivo* high-throughput screening to generate an expandable set of high-performing promoters that

exhibit both low basal and high activated expression levels. Through a sequential selection approach, we design activatable promoters with up to 1000-fold dynamic range, constituting a 33-fold improvement from previous synthetic promoters when ported to CFS [32], [45]. These promoters enable network topologies not previously accessible, including a six-layer deep cascade and a four-branch parallel circuit. Additionally, by engineering activatable promoters, we successfully incorporate different inputs into CRISPRa/i circuits through the blue light-responsive EL222 transcriptional activator as well as protein-protein interaction (PPI)-dependent CRISPRa assembly, including SYNZIP and the small-molecule responsive abscisic acid (ABA) and gibberellic acid (GA) systems. Through the engineering of high-performing CRISPRa/i nodes as well as input signal processing modules, we show that CRISPRa/i circuits built with these components can be tuned to achieve a broad range of functions, including deep multi-layer and wide multi-branch activation cascades, positive feedback loops, AND logic gates, and dynamic two-input pulse modulators. Overall, this work describes a new workflow for engineering activatable promoters and provides a toolbox of versatile components with immediate utility for implementing CRISPRa/i circuits. Together, these developments dramatically expand the ability to assemble large, multi-input GRNs in CFS.

Results

We first sought to characterize the impact of RNAP recruitment on both basal and activated expression levels of synthetic activatable promoters. RNAP recruitment is dependent on the affinity of the RNAP sigma subunit (σ) for the -10 and -35 hexamers of the minimal promoter, as well as the affinity of the RNAP α-CTD subunits for the UP-element sequence upstream of the minimal promoter [47], [48] (Figure 1A). We systematically designed libraries of these discrete promoter regions and screened them in *E. coli*. The libraries were co-transformed with an aTc-inducible CRISPRa plasmid to enable parallel screening of basal and activated expression levels (Figure 1B, Methods 2.2) [32]. In the absence of aTc, RNAP recruitment is determined by the promoter basal strength. Upon aTc induction, the MCP-SoxS activator is expressed and localized to the promoter via CRISPRa complexes with scaffold RNAs (scRNAs) containing an MS2 hairpin. MCP-SoxS then recruits RNAP to the promoter through α-CTD interactions, activating transcription [49] (Figure 1A). This approach allows us to characterize the impact of individual promoter regions on basal and activated expression simultaneously, and combine variants with low basal and high activated expression to construct high-performing activatable promoters.

Functional interrogation of promoter regions with CRISPRa

<u>Impact of Minimal Promoter Region on Activability</u>

Previous work has demonstrated the importance of the minimal promoter region in determining promoter basal and activated expression levels [39], [45], [49]–[51]. We designed two minimal promoter libraries mutagenizing the -10/-35 hexamers and the

intervening sequence of the previously identified best-performing minimal promoter (BBa_J23117) within the J3 synthetic CRISPRa promoter (Figure 1C, Methods 3.1) [45]. These libraries were co-transformed with the aTc-inducible CRISPRa plasmid expressing the J306 scRNA that targets the cognate J3 promoter. Both libraries yielded high promoter diversity with basal and activated expression levels ranging from that of a no-reporter control to a strong constitutive promoter (BBa_J23119) (Figure 1D). The set of minimal promoter variants that maintain both low basal and high activated expression levels can be conceptualized as a Pareto optimal front. In multi-objective optimization, a Pareto front defines the best-performing solutions for which no further improvements in either objective can be achieved without compromising the other [52]. Three variants from this Pareto front exhibited up to 32% lower basal and 4% higher activated expression levels compared to the original BBa_J23117, indicating the original minimal promoter was not a part of the Pareto front. This finding suggests that promoter mutagenesis can yield improved activatable promoters beyond previous screening methods based on promoter basal strengths alone [45]. By mutagenizing the minimal promoter of CRISPRa promoters, we generated sequences with >100-fold dynamic range in CRISPRa-mediated gene expression.

Impact of UP-Element Region on Activability

RNAP promoter recognition is enhanced by the AT-rich UP-elements upstream of the minimal promoter, which anchor the α-subunits of RNAP [39], [47], [53]. For effective CRISPRa, RNAP should only be recruited to the promoter in the presence of an on-target scRNA. Hence, for transcriptional activation with SoxS, improvements in dynamic range could be achieved by minimizing RNAP-UP-element interactions and lowering basal

expression levels. We designed five UP-element libraries mutagenizing the AT-rich *E. coli* consensus sequence with increasing GC-content (Figure 2A, Methods 3.2). As expected, the consensus UP-element and the AT-rich library had the highest basal expression levels (Figure 2B, left). On average, these libraries showed only 3-fold activation, as compared to 37- to 44-fold activation for the more GC-rich libraries (Figure S1). Qualitatively, we observed a monotonic decrease in basal expression levels and no impact on maximum activated levels with increased GC-content (Figure S1). We identified the optimal variants from each library and found a shift in the Pareto front towards lower basal and higher activated expression levels with increasing UP-element GC-content (Figure 2B, right). Specifically, the median basal and median activated expression levels of the GC-rich optimal variants were 59.7-fold lower and 1.7-fold higher than that of the AT-rich optimal variants. The original J3 synthetic promoter sat in between fronts consisting of high and low GC-content variants. By mutagenizing the UP-element, we generated promoter variants with >350-fold dynamic range in CRISPRa-mediated gene expression.

Impact of the scRNA target site Region on Activability

Transcriptional activators bind upstream of the minimal promoter region to recruit RNAP to the transcription start site (TSS) [39]. For CRISPRa, the optimal scRNA target site location for SoxS-mediated activation is -81 bp upstream of the TSS [45]. Changing the scRNA target sequence enables rapid generation of orthogonal CRISPRa promoters [54], [55]. Due to the proximity to the UP-element region, we reasoned that the sequence composition of the scRNA target site may have an impact on basal expression levels. We designed three scRNA target site libraries composed of varying GC-content (Methods

3.3) and measured the basal expression of each library. We found that GC-rich libraries had 4.3-fold lower median basal expression compared to AT-rich libraries (Figure 2C). Additionally, the spread of the basal expression decreased monotonically with increasing GC-content of the scRNA target site sequence (Figure 2C). Together, these results indicate GC-rich scRNA target site sequences lead to low basal expression CRISPRa promoters. To validate the CRISPRa activity at these low basal expression scRNA target sites, we then selected 10 GC-rich variants and constructed the corresponding scRNAs. All variants produced a higher fold-activation than the original J306 scRNA (Figure S2), with 3.5-fold average increase in fold-activation.

Combining promoter regions to engineer high-performing CRISPRa promoters

Engineering activatable promoters by combining optimized promoter regions

We proceeded to test if the highest performing variants from the UP-element and minimal promoter screens could be combined to yield activatable promoters with improved performance. We selected three high-performing variants from both the UP-element and minimal promoter screens, as well as the starting J3 UP-element and BBa_J23117 minimal promoter, and constructed a combinatorial set of 16 promoters. Surprisingly, we found that promoter regions that gave the largest improvements in the original context did not necessarily give the largest improvements when tested in different contexts (Figure S3). For instance, an UP-element that gave 400-fold activation when tested with the BBa_J23117 minimal promoter only gave 270-fold activation when tested with a different high-performing minimal promoter, largely due to a decrease in activated expression (Figure S3, right). These results highlight that promoter region context effects play an important role in the design of high-performing activatable promoters.

Engineering activatable promoters through sequential promoter region screening

We tested whether promoters with improved basal and activated expression levels could be achieved by selecting minimal promoters in the context of high-performing UP-elements. We first screened the UP-element region as these libraries had a larger impact than the minimal promoter libraries on the location of the Pareto front (Figures 1D, 2B). We mutagenized the UP-element of a promoter containing the minimal promoter BBa_J23117 and a high GC-rich scRNA target. We selected three UP-element variants from the Pareto front, which had on average 90% lower basal and 12% higher activated expression than the J3 promoter (Figure 2D). We then screened minimal promoter libraries in the context of these three selected UP-element variants. We again selected three new promoter variants from the Pareto front, which had on average 83% lower basal and 56% higher activated expression than the J3 promoter (Figure 2D). Notably, with this sequential screening approach we were able to shift the Pareto front towards lower basal and higher activated expression in both the UP-element and minimal promoter screens.

Engineering deep and wide circuits with high-performing CRISPRa promoters

CRISPRa promoters expressing scRNAs for CRISPRa or sgRNAs for CRISPRi, termed CRISPRa/i nodes, can be assembled into multi-layer circuits with network topologies specified by the guide RNAs. Proper circuit function requires level-matching the input/output dynamic ranges between sequential CRISPRa/i nodes to minimize signal degradation [31]–[33]. Promoters with lower basal and higher activated expression levels should span a higher fraction of the input dynamic range of downstream nodes, resulting in less signal degradation [32]. Therefore, we sought to evaluate how dynamic range improvements achieved at the promoter level translate into fold-activation and signal

propagation improvements at the circuit level and enable construction of increasingly deep and wide CRISPRa circuits. We implement these circuits in CFS due to the ease of rapidly prototyping genetic parts and circuits [6], [32]. Given the strong correspondence between component function in bacteria and CFS [32], [56], [57], we reasoned that the high-performing CRISPRa promoters designed in *E. coli* could be used to build larger and more complex GRNs in CFS.

Engineering functional CRISPRa/i nodes

We first characterized the three promoter variants selected above (Figure 2D, HP1-3) in CFS and observed up to ~1000-fold dynamic range (Figure S4). We then generated a set of orthogonal CRISPRa/i nodes to be assembled into multi-layer circuits following previously-described methods [32]. We combined the highest dynamic range activatable promoter (HP3, Table S1) with previously-screened scRNA target sites to generate orthogonal CRISPRa/i nodes (Figure 3A). We characterized the dose-response curves for each orthogonal scRNA-promoter pair and found these new nodes gave an average activation of 890-fold (Figures 3B, S5; Table S1).

Deep CRISPRa Circuits

We investigated whether deep multi-layer cascades could be implemented using the improved CRISPRa/i nodes. We first built a two-layer CRISPRa cascade by tuning the expression levels of the input and internal CRISPRa/i nodes and achieved up to 127-fold activation (Figures S6, S7). Next, we assembled four-layer activation cascades. To compare circuit performance and dynamics in response to a scRNA input, we measured RFP expression and time to maximum expression rate (t_{max}) (Figure 3C, insert). If the input signal propagates faster than the leak from the rest of the nodes, CRISPRa-

dependent expression through the network accelerates, reducing t_{max}. Therefore, a larger t_{max} between the +/- input conditions (taken as Δt_{max}) corresponds to an improved circuit function. For a cascade with equal node concentration at each layer, we found that the circuit was not input responsive (Figure 3C, middle). We then tuned node concentrations by either decreasing or increasing the concentration of each subsequent node as depth increased (denoted "*D*" and "*I*", respectively) (Figure 3C, left). Assemblies *D* and *I* both had their t_{max} significantly accelerated compared to the no input conditions (Δt_{max} of 85 min and 165 min, respectively) (Figure 3C, right). Assemblies *D* and *I* also had the lowest and highest leak and absolute expression levels (Figure 3C, middle), indicating that timing and expression level of multi-layer, input-responsive circuits can be controlled through node concentration tuning.

We then used individual scRNA-promoter dose-response curves to inform circuit assembly. This strategy achieved 3-fold higher activated expression than assembly *D* and 4.5-fold lower basal expression than assembly *I*, and a Δt_{max} of 105 min (Figure 3C). We tuned the concentrations of the first and third internal nodes of the four-layer cascade for a 8-fold improvement in fold-activation (Figure 3D). Qualitatively, higher concentrations of the third node resulted in higher activated states, while lower concentrations of the first node minimized basal expression (Figure S8). We then changed the high-performing promoter of the second internal node for a worse performing promoter with higher basal expression. We observed no difference in expression level or t_{max} with or without input scRNA, indicating the circuit was no longer input-responsive. This result underscores the importance of high-performing promoters for building deep transcriptional circuits.

We investigated how input signals are propagated through increasingly deep circuits (Figure 3E, left). We define signal propagation as the ratio of the fold-activation between the cascade output and input layer. Signal propagation was sustained above 80% until the 4th layer was added, after which it decreased rapidly (Figure 3E, right). Nevertheless, we observed measurable output differences in circuits of up to six layers. We quantify the dynamics of signal propagation in terms of signal delay, or the difference in time to reach the maximum fold-activation of the cascade between the cascade output and input layers. The two-layer cascade gave no significant difference in signal delay compared to a single-layer CRISPRa reaction. This may suggest there is a slow step in output production, such as fluorophore maturation, that masks the effect of the second layer. Beyond two layers, the signal delay showed a consistent increase of ~50 min with subsequent additions of the third, fourth, and sixth layers (Figure 3E, right), suggesting that the response characteristics are maintained at each layer in deep circuits.

Wide CRISPRa Circuits

To identify conditions under which multiple orthogonal nodes can compose wide CRISPRa circuits, we constructed one, two, three and four parallel three-layer cascades operating in the same CFS reaction. We used a single input to activate the downstream nodes, and measured circuit performance by connecting all cascade outputs to a RFP node (Figure 3F, left). We maintained the internal node concentrations constant across parallel cascades. We observed up to 66% decrease in output fold-activation as the width of the circuit increased from one to four cascades (Figure 3F, right). This decrease came largely from higher output levels in the absence of scRNA input (Figure S9), most likely due to higher overall basal expression of internal scRNAs. We then constructed the same

circuits and tuned the node concentrations proportionally to the number of parallel cascades, effectively maintaining the total node concentration constant. When constructed in this manner, we found no statistically significant difference in the fold-activation across cascades of different widths (Figure 3F, right). Hence, by tuning the concentration of orthogonal CRISPRa/i nodes, we show an arbitrary number of parallel circuits with as many as nine nodes may be regulated.

Developing activatable promoters for blue-light responsive CRISPRa/i circuits

Above we demonstrated that high-performing CRISPRa promoters can be generated through parallel screening of the basal and activated levels of promoter variants. In this section, we show that the same approach can be used to generate high-performing activatable promoters for blue-light responsive gene expression using the EL222 transcriptional activator. Furthermore, by placing guide RNA expression under the control of the engineered high-performing EL222 promoter, we demonstrate light-responsive CRISPRa/i signal transduction modules and signal processing CRISPRa/i circuits.

<u>High-performing blue-light responsive promoters</u>

The EL222 transcriptional activator interactions with the RNAP and the DNA binding site are well-characterized, making it a suitable model system for developing optogenetic inputs for CRISPRa/i circuits (Figure 4A) [58]–[61]. Briefly, EL222 binds an 18 bp sequence upstream of the -35 region of the *luxI* promoter and subsequently recruits RNAP through interactions with the α and σ subunits [61]. We mutagenized the *luxI* minimal promoter (Figure 4A, Methods 3.4), and screened variants in *E. coli* in both dark and light to select for high dynamic range (Figure 4B). Starting with a dynamic range of

less than 2-fold, we observed up to 4-fold dynamic range in response to blue-light. Similar to our CRISPRa promoter screens, minimal promoters with very low (BBa_J23113) or very high (BBa_J23119) basal expressions exhibited low dynamic range in response to blue-light. We selected 4 variants with >2-fold higher dynamic range than the *luxI* minimal promoter and characterized them in CFS. These variants yielded a 34.1-fold difference in expression between light and dark, compared to just 6.2-fold for the original *luxI* minimal promoter (Figure S10). This improvement comes largely from a reduction in the basal expression from the blue-light promoter, suggesting we successfully minimized the RNAP-minimal promoter interactions without weakening EL222-promoter interactions. More importantly, these results demonstrate this approach for engineering actionable promoters is applicable to other transcriptional activation systems.

Blue-light responsive CRISPRa/i circuits

We evaluated whether the engineered blue-light promoter transcription levels were suitable for expressing gRNAs for CRISPRa/i circuits. We titrated gRNA-expressing plasmid concentrations and compared RFP expression across dark and light conditions. For CRISPRi, the highest light-dependent change in repression was 50% (Figure S11). For CRISPRa, the highest light-dependent fold-activation was 14-fold (Figures 4C, S12). We then implemented a positive feedback loop to increase light-dependent CRISPRa output levels. In a CRISPRa/i circuit, positive feedback (PFB) can be achieved by including a downstream node that expresses a scRNA targeting an upstream node. We expected the degree of positive feedback in the system to be tunable by titrating the PFB node, with high concentrations of this node resulting in activation in the absence of blue-light. When optimally tuned, the positive feedback loop increases the light-dependent

CRISPRa output levels almost 2-fold (Figure 4D). Excess PFB node led to a 7.6-fold increase in basal expression, decreasing the light-dependent activation to 1.2-fold. These results highlight that rationally designed genetic circuits built from engineered activatable promoters can be used to improve the dynamic range of input-responsive signal processing modules.

Engineered activatable promoters enable conditional CRISPRa dependent on protein-protein interactions

The versatility of protein-protein interaction (PPI)-mediated genetic regulation for coupling peptide or small-molecule binding to transcriptional outputs has long attracted interest [62]–[68]. Implementation of PPI-dependent transcriptional activation has been difficult in prokaryotic systems due to strict target site requirements and low dynamic range of activatable promoters [45], [49]. Therefore, we explored if high-performing activatable promoters could permit construction of PPI-dependent conditional CRISPRa signal transduction modules.

Development of Conditional CRISPRa Systems

As an experimental testbed, we incorporated three previously characterized protein-protein heterodimerization domains into our CRISPRa system: the synthetic coiled-coil SYNZIP 5/6 pairs [69], the abscisic acid (ABA) responsive ABI–PYL1 [62], [70], [71], and the gibberellic acid (GA) responsive GID1–GAI [62], [72]. We fused these heterodimerization domains to SoxS and MCP to enable conditional recruitment of SoxS to the CRISPRa complex. We generated the MCP-SZP6 and SoxS-SZP5 domains for SYNZIP-CRISPRa, the MCP-ABI and SoxS-PYL1 domains for ABA-CRISPRa, and MCP-GAI and SoxS-GID1 for GA-CRISPRa (Figure 5A). In the original J3 promoter context, we observed a 5.7-fold activation with SYNZIP-CRISPRa when SoxS was fused to the N-

terminus of the SYNZIP domain, compared to 1.4-fold when fused in the opposite orientation (Figure S13). The fold-activation of ABA-CRISPRa was also maximized when SoxS was at the N-terminus (Figure S13, right), therefore we moved forward with all SoxS N-terminus fusions. MCP-SYNZIP gave 5.7-fold activation, while SYNZIP-MCP gave only 2.8-fold (Figure S13). For ABA- and GA-CRISPRa, only C-terminus MCP fusions were tested due to the MCP-SYNZIP result and the strong precedent for using C-terminus MCP fusions in CRISPRa systems [49], [73].

CRISPRa operates narrowly within -71:-101 bp from the TSS in a phase-dependent manner [45]. We questioned whether the introduction of an additional protein linkage into MCP-SoxS affects the relative scRNA target site requirements [74]. We designed a CRISPRa promoter with densely packed scRNA target sites and variants with 1 bp frameshifts to allow screening with single base pair resolution between -81:-111 bp from the TSS (Figure 5B, Table S1). Surprisingly, SYNZIP-CRISPRa maintains the same preference for the -81 bp from the TSS targeting site and the same stringent 10-11 bp phase dependency seen in conventional CRISPRa (Figure 5B, S14). Having identified the optimal scRNA target site, we changed the promoter with densely packed scRNA target sites for a high-performing promoter (HP3, Table S1), and observed a 5.4-fold improvement in SYNZIP-CRISPRa fold-activation (Figure S15).

We then identified the permissible small-molecule input and protein expression levels of each conditional CRISPRa system in the context of our high-performing activatable promoters. For ABA- and GA-CRISPRa, small-molecule titrations showed that ABA-CRISPRa is responsive between 0.1-10 uM, and GA-CRISPRa is responsive between 0.1-100 uM (Figure 5C), with up to 7.9- and 9.0-fold-activation, respectively. We

screened dimer stoichiometries and expression levels by surveying a range of concentrations for both the MCP and SoxS components. SYNZIP-CRISPRa performs the best of the three systems, giving a maximal activation of 67-fold compared to a reaction with no MCP or SoxS plasmid added. Even at low CRISPRa component concentrations, SYNZIP-CRISPRa still achieves 59-fold activation (Figure 5D, left). ABA-CRISPRa gives a maximum activation of 18.6-fold (Figure 5D, middle). For GA-CRISPRa, the maximum activation of 5.9-fold was accessible in a relatively narrow range of component concentrations (Figure 5D, right). For all three conditional CRISPRa systems, higher expression of heterodimers did not necessarily improve activation. In line with the behaviors of natural scaffolds [75], [76], we observed a unique optimal concentration for each conditional CRISPRa system. The differences between systems may be due to the different affinities of each protein-protein interaction.

Engineering multi-input CRISPRa/i circuits

Multi-layer and multi-input circuits with conditional CRISPRa

To use conditional CRISPRa systems as inputs into more complex CRISPRa/i circuits, we began by characterizing the scRNA dose-response curve of the novel CRISPRa systems. For all three conditional systems, the amount of scRNA needed to saturate the CRISPRa response was similar to that of direct CRISPRa (Figures S5, S16). We tested the orthogonality of the small molecule systems to evaluate if they could be used together for independent gene regulation (Figure 6A). We found ABA-CRISPRa to be highly specific to its target ligand, showing no significant activation in the presence of GA. GA-CRISPRa showed 3.1-fold cross-activation from ABA, in line with reports in eukarya [62], but maintained a 3-fold higher specificity for its target ligand, giving 10.5-

fold activation from GA. These results suggest that the ABA- and GA-CRISPRa systems can be used for effective orthogonal gene regulation.

We built two types of input-responsive circuits to demonstrate the use of conditional CRISPRa for multi-input and multi-layer input processing: an AND logic gate and a CRISPRa cascade. To accomplish AND logic, we co-expressed both ABA- and GA-CRISPRa (Figure 6B, left). The addition of either ABA or GA resulted in 2-fold activation compared to the no-ligand condition. In the presence of both ligands, 4.5-fold activation was achieved, allowing us to clearly distinguish between one- and two-inputs ($p = 0.03$) and achieve AND logic. A two-layer ABA-CRISPRa cascade with both internal layers dependent on ABA gave 2.5-fold activation upon addition of ABA, showing that conditional CRISPRa can also support multi-layer information processing (Figure 6B, right).

Two-input dynamic pulse generator

Synthetic biologists aim to recreate complex, dynamic signaling networks that use multiple input-responsive regulators to tightly regulate the expression timing and magnitude of downstream targets [2], [3], [77]. Towards this goal, we sought to couple blue-light CRISPRi with conditional CRISPRa into a tunable pulse generator with two-input control over the level and timing of gene expression (Figure 6C, left). Based on our system characterizations, SYNZIP-CRISPRa is well suited for integration with blue-light CRISPRi due to its high-fold activation. To simulate relative production rates as a function of CRISPRa and CRISPRi inputs, we built upon a coarse-grained mechanistic model of CRISPRa/i regulation [32] by introducing blue-light pulses regulating sgRNA expression (Methods 6). By simulating changes in the pulse width as CRISPRi inputs as well as

changes in the scRNA concentration as CRISPRa inputs, we were able to capture changes in both the levels and timing of RFP production rates (Figure 6C, left). This simulation analysis indicates that there are regimes where blue-light CRISPRi and conditional CRISPRa can be coupled to modulate the production rate of a desired gene target.

Experimentally, we first kept the CFS in the dark for one hour to allow for EL222 expression and used varying blue-light exposure times to dynamically tune CRISPRi. To tune SYNZIP-CRISPRa, we changed the scRNA-expressing plasmid concentration. As predicted by the model, RFP production rate pulse was tunable by both the scRNA plasmid concentration and the exposure time to blue-light (Figure 6D). When compared across conditions with the same CRISPRa input, higher CRISPRi input led to 20-56% lower maximum production rates. When compared across conditions with the same CRISPRi input, higher CRISPRa input led to faster maximum production rates by 30 minutes. These results demonstrate that different CRISPRa/i-based signal transduction modules can be integrated and individually modulated to control gene expression, highlighting the potential of input-responsive CRISPRa/i GRNs for complex signal processing applications.

Discussion

Natural biological systems have evolved GRNs containing activatable promoters with wide ranges of expression characteristics that enable dynamic responses to changing environmental conditions. Engineering activatable promoters has traditionally been thought to involve a trade-off between basal and activated expression levels [35], [41], [45]. In this work, we show that basal and activated expression levels can largely be decoupled to generate activatable promoters with both lower basal and higher activated expression levels than previously possible. High-throughput, sequential screening of promoter regions constructed through sequence-based design allowed us to overcome context effects and identify high-performing activatable promoters. With this approach, we successfully engineered a suite of orthogonal CRISPRa promoters that match the basal and activated expression levels of the canonical Tet inducible system and exceed those of the Lac system (Figure S17) [78].

The *E. coli* transcriptional network is governed by a hierarchical structure containing nine layers of regulation [1]. Engineered activatable promoters allowed us to build multi-layer CRISPRa/i GRNs in *E. coli*-based CFS with depths and widths significantly larger than the state of the art [32], [79], approaching the complexity of natural GRNs. Specifically, a 33-fold improvement in promoter dynamic range resulted in 80% lower signal degradation in two-layer cascades (Figures S6), and enabled deep GRNs with up to six layers of regulation. Additionally, we demonstrated wide GRNs regulating up to four parallel cascades, indicating that the CRISPRa/i framework is well suited for the design of wide control circuits for parallel computing and multi-gene regulation. Further improvements in GRN complexity may be limited by resource constraints, such

as upstream gRNAs outcompeting downstream gRNAs for dCas9 binding (Figure S7). Strategies to dynamically regulate upstream gRNA expression, such as reversing CRISPRa complex binding or implementing negative autoregulation motifs [80], could enable even larger GRNs.

Biological systems continuously monitor and process environmental signals by using signal transduction modules, such as inducible transcriptional effectors, as inputs to complex GRNs [81], [82]. Our work provides a general framework for optimizing transcriptional activation systems at the promoter level and integrating them into CRISPRa/i GRNs. Promoter engineering of the optogenetic EL222 system enabled high light-dependent dynamic ranges, with relevant expression levels for downstream applications. Through inducible gRNA expression, we demonstrated input signal modulation through various GRN topologies, including positive feedback loops and CRISPRa/i cascades, as well as integration of different signal transduction modules into the same CRISPRa/i GRN. Given the high degree of tunability and composability of the CRISPRa/i framework [28], [32], other GRNs with different signal processing functions, such as thresholding [83], amplification [77], and frequency filtering [84], should be readily implementable. Overall, our work highlights the potential for building complex, input-responsive systems through CRISPRa/i GRNs.

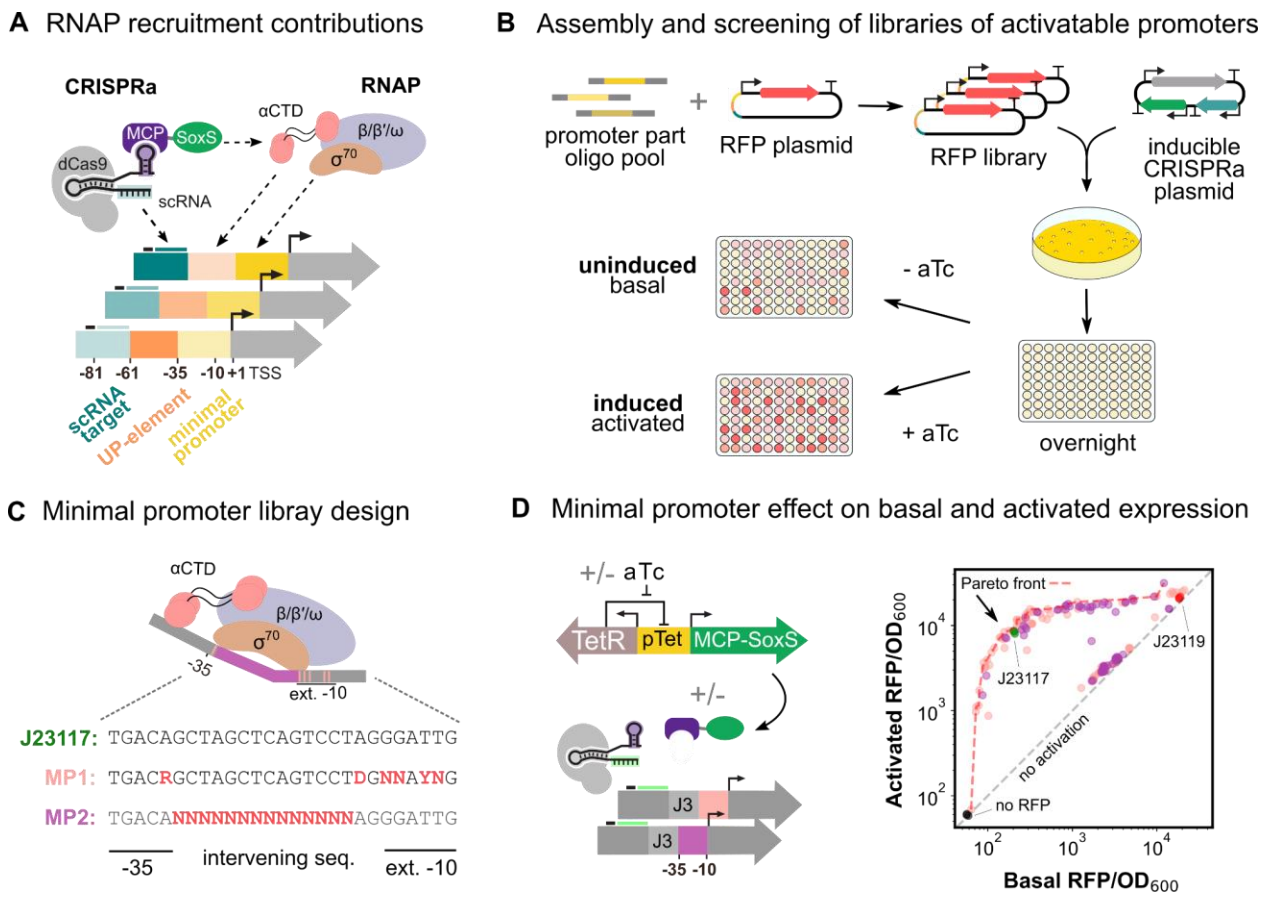
Protein-protein interactions have been used widely to execute complex, input-responsive functions in eukaryotes [85]–[89]. Implementing similar systems in prokaryotes has been difficult, and the development of high dynamic range promoters allowed us to successfully prototype and optimize conditional CRISPRa systems in *E. coli*-based CFS. Implementation of novel conditional CRISPRa systems may be

streamlined by the fact that all systems tested here are effective when targeted -81 bp from the TSS, despite the presence of additional protein-protein interactions up to 500 amino acids in length. Additionally, conditional CRISPRa fold-activation is proportional to the strength of the protein-protein interaction (Table S2) [69], [72], [90]–[94], informing the *a priori* selection of heterodimers for use in conditional CRISPRa. Collectively, our work suggests that other heterodimerization domains could be implemented, with minimal prototyping, as signal transduction modules for CRISPRa/i GRNs for multiplexed biosensing or screening of PPIs in CFS.

Our workflow for activatable promoter engineering enables the dynamic specification of expression levels for large networks of orthogonal gene targets. The new classes of deep, wide, and input-responsive CRISPRa/i GRNs developed here have immediate application in CFS for investigating the rules for genetic circuit design [10], [13] and biological information processing [20]–[23], as well as for building dynamic, multi-enzyme expression programs for self-assembling bioproduction platforms [17]–[19], [95], [96]. Moreover, CRISPRa/i GRNs could be integrated with existing field-deployable medical diagnostics and environmental monitors to enable complex, multi-input signal processing [4], [15], [16], [97]. Moving forward, this work could serve as a stepping stone for building entirely synthetic cells and engineered living materials with GRNs that match or go beyond the complexity of natural systems.

Figures

Figure 1: Functional interrogation of promoter regions with CRISPRa



A. Schematic of RNAP interactions with the CRISPRa complex and target promoter. σ^{70} affinity for the minimal promoter and α-CTD affinity for the UP-element determines RNAP recruitment to a promoter. When CRISPRa is targeted to a promoter with a complementary scRNA target site, the RNAP α-CTD domain is recruited by the SoxS transcriptional activator. RNAP-promoter and CRISPRa-promoter interactions can be modulated by modifying the DNA sequence of the different promoter regions.

B. Workflow for the assembly and characterization of libraries of activatable promoters. A library of RFP genes with varying promoters is generated through PCR (Methods

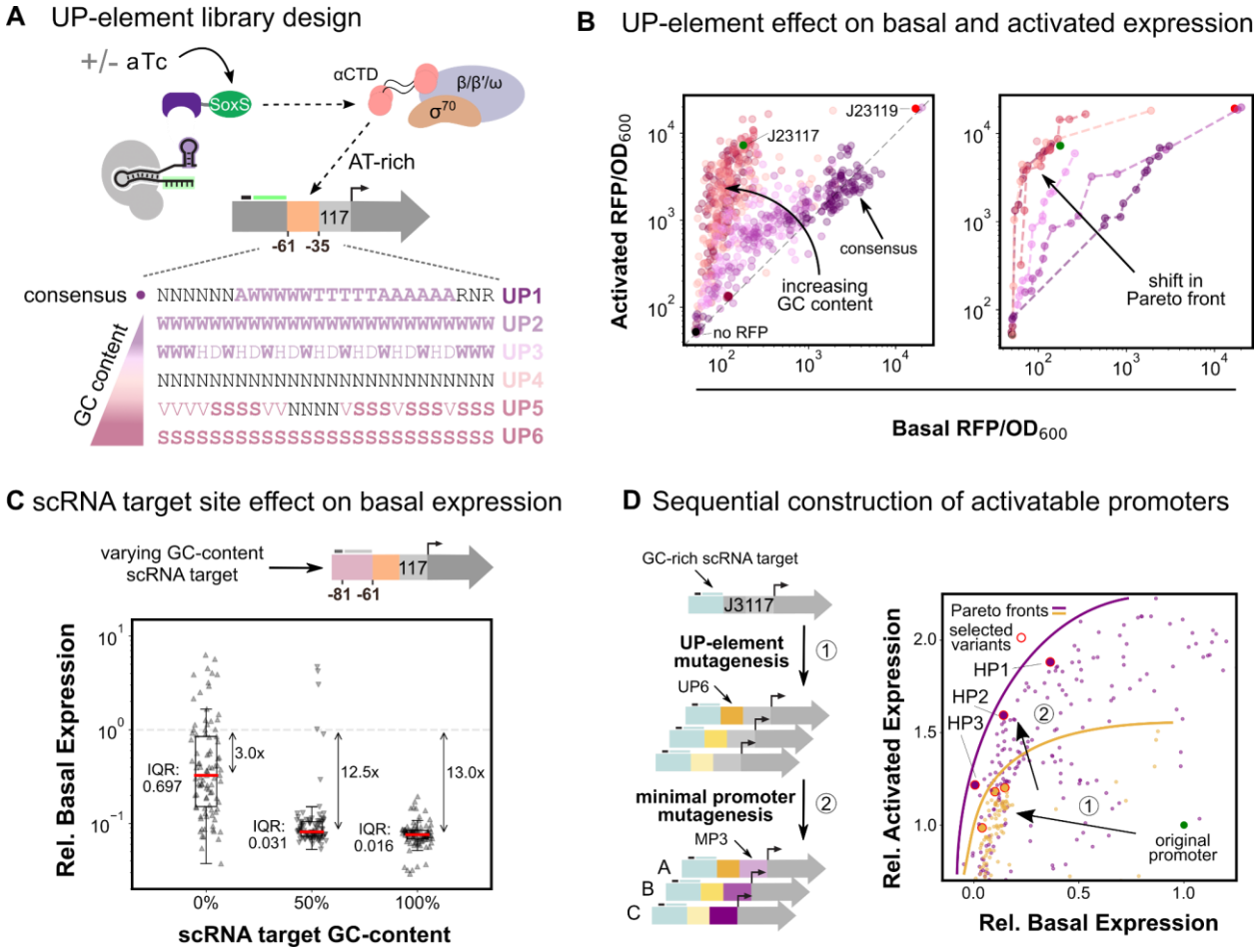
554 1.1). The library is then co-transformed into *E. coli* with an aTc-inducible CRISPRa plasmid. Colonies are then seeded overnight in EZ-RDM. Overnight cultures are
556 subsequently diluted into a fresh EZ-RDM and supplemented with appropriate concentrations of aTc. For each promoter variant in the libraries, basal and activated
558 RFP levels were measured with 0 nM and 200 nM aTc, respectively (Methods 2.2).

C. Schematic of RNAP interaction with the minimal promoter and library design. σ^{70}
560 recognizes specific positions in the extended -10 and -35 regions of the minimal promoter, which informed the design of the library MP1. σ^{70} binding is also influenced
562 by the GC-content, the length, and the $^{-15}$TGn^{-1} motif of the intervening sequence, which informed the design of library MP2 (Methods 3.1).

564 D. Minimal promoter effect on expression levels. **Left:** Schematic of inducible CRISPRa system and minimal promoter libraries of the J3 synthetic promoter. MCP-SoxS is
566 expressed from the aTc-inducible pTet promoter. dCas9 and J306 scRNA are constitutively expressed. **Right:** Activated and basal RFP/OD_{600} for the two minimal
568 promoter libraries. Red dash line defines the Pareto front containing the best performing promoter variants (Methods 7.1), for which no further improvements in
570 basal or activated levels can be achieved without compromising the other. Gray dash line defines promoter variants with equal activated and basal expression levels,
572 indicating they are not activated by CRISPRa. The J23117 minimal promoter (green) is included as a standard reference for CRISPRa efficiency. The J23119 minimal
574 promoter (red) is an example of a non-activatable promoter due to high basal

expression levels. A plasmid without RFP (black) indicates the background fluorescence of the system.

Figure 2: Combining promoter regions to engineer high-performing CRISPRa promoters



578 A. Schematic of RNAP interactions with the UP-element and library design. α-CTD affinity for AT-rich UP-elements upstream of the minimal promoter helps recruit RNAP. 580 Upon targeting with CRISPRa, UP-element RNAP recruitment contributions are largely replaced by SoxS-RNAP α-CTD interactions. UP-element libraries with 582 increasing GC-content were designed to minimize α-CTD interactions (Methods 3.2).

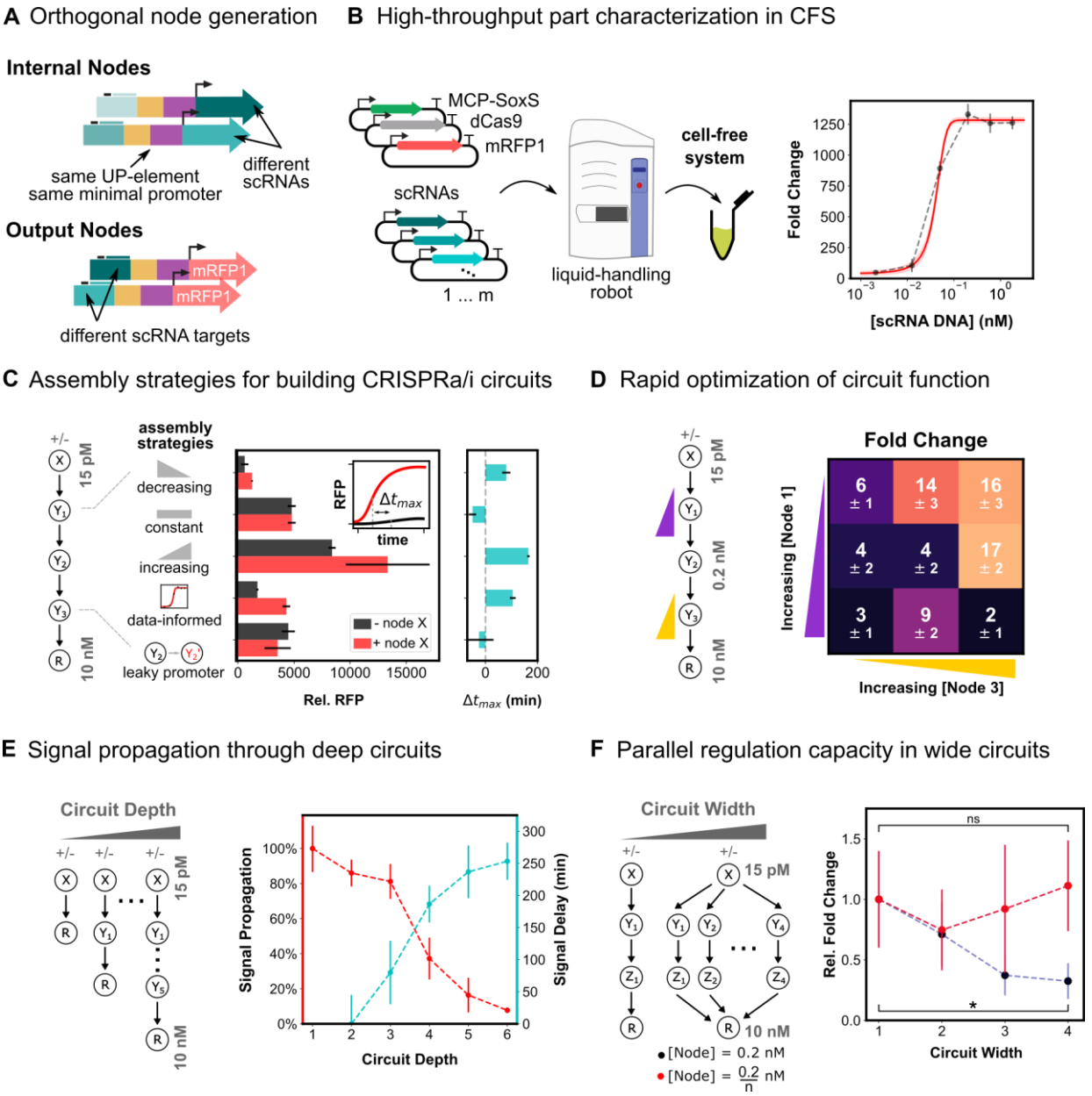
B. UP-element GC-content effect on expression levels is shown through activated and 584 basal RFP/OD_{600} for the six UP-element libraries. **Left:** Increasing GC-content in the UP-element lowers the range of basal expression level, while maintaining the full 586 range of activated expression levels. Gray dash line defines promoter variants with

equal activated and basal expression levels, indicating they are not activated by CRISPRa. **Right:** Colored dash lines define the Pareto front containing the best performing promoter variants for each UP-element library (Methods 7.1). Increasing the UP-element GC-content effectively shifts the Pareto front towards lower basal expression levels.

C. scRNA target site composition effect on basal expression. Comparison of three scRNA libraries with increasing GC-content (Methods 3.3). Basal expression levels are normalized to the standard J3 promoter basal expression level. Red lines indicate the median expression level of each distribution. The interquartile range (IQR) is calculated as the difference between the upper and lower quartiles and measures the spread of the distribution.

D. Sequential construction of activatable promoters. **Left:** Activatable promoters were constructed by sequential library mutagenesis screens starting from the J3 promoter with a GC-rich scRNA target site. Three Pareto optimal UP-elements were selected after promoter mutagenesis with a GC-rich UP-element library (1). We then mutagenized the minimal promoter of the three previously selected variants (2), and again selected three Pareto optimal variants. **Right:** Basal and activated expression levels for all mutagenesis variants normalized to the standard J3 promoter expression levels (green). Yellow points represent variants from the UP-element mutagenesis (1), while purple points represent variants from the minimal promoter mutagenesis (2). Red circles indicate selected variants from each screen, and solid lines depict the Pareto optimal fronts. Each sequential mutagenesis led to variants with both lower basal and higher activated expression levels.

Figure 3: Engineering deep and wide circuits with high-performing CRISPRa promoters



A. Schematic of orthogonal CRISPRa/i nodes for use in cell-free circuits. Internal nodes contain an orthogonal scRNA target site and express orthogonal scRNAs. Output nodes contain orthogonal scRNA target sites and express RFP. All nodes contain the same UP-element and minimal promoter (HP3).

B. High-throughput characterization of scRNA components in CFS. **Left:** Plasmids encoding each CRISPRa component are mixed using an acoustic liquid handling

robot and expressed in CFS. **Right:** scRNA-dose response curves for each node are generated by titrating the amount of scRNA plasmid from 0.5 pM to 5 nM. Values represent the mean ± standard deviation of three technical replicates.

C. Comparison of assembly strategies for building a four-layer CRISPRa cascade. **Left:** Internal node concentrations either decreased from 200 pM to 32 pM as depth increased, were held constant at 200 pM, or increased from 200 pM to 1.25 nM as depth increased. A fourth assembly method was tested in which internal node concentrations were 40, 200, and 170 pM, based on individual scRNA-dose response characteristics. A fifth cascade was included in which the high-performing promoter of the second internal node was replaced with the leaky J2 promoter. Input and output node concentrations were held constant across all strategies at 0 or 15 pM and 10 nM, respectively. **Center:** Cascade output RFP expression for each assembly strategy with scRNA input (red) and without (black), relative to RFP basal expression. Values represent the mean ± standard deviation of three technical replicates. **Right:** Change in time to maximum expression rate (Δt_{max}) for each assembly strategy. Δt_{max} is calculated by finding the difference in time to reach maximum RFP production rate between the with and without input conditions (Methods 7.1).

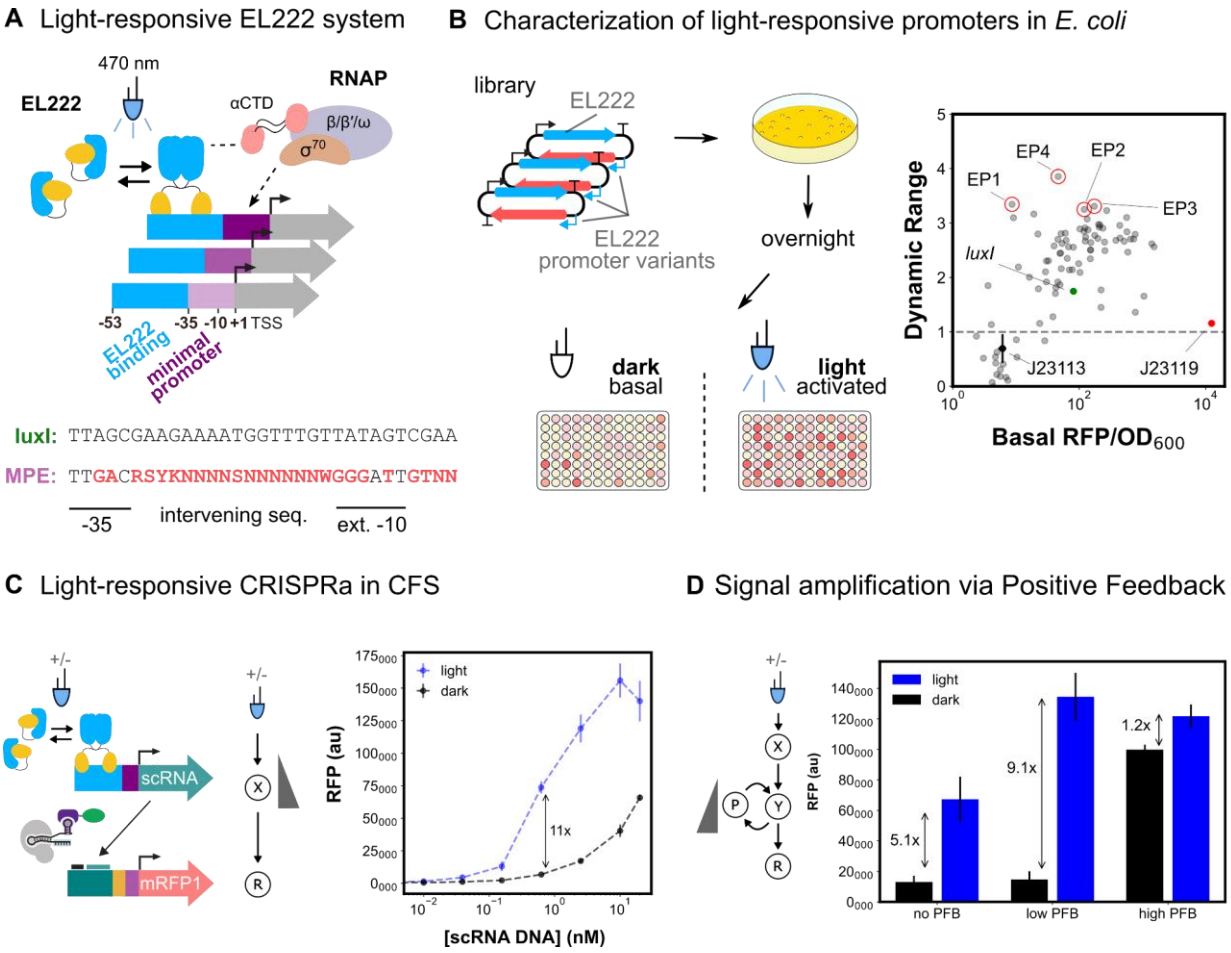
D. Rapid fold change optimization of a four-layer CRISPRa cascade. **Left:** The first and third internal nodes of the cascade were varied between 40 and 160 pM, and 85 and 340 pM, respectively. The input node, second internal node, and output node were held constant at 0 or 15 pM, 0.2 nM, and 10 nM, respectively. **Right:** Fold change between with and without scRNA input for each CRISPRa cascade.

E. Signal propagation through deep CRISPRa/i circuits. **Left:** Schematic of CRISPRa cascades with increasing depth. Input and output node concentrations were held constant across all cascades at 0 or 15 pM and 10 nM, respectively. All node concentrations are tabulated in Table S5. **Right:** Propagation efficiency and signal delay are shown as a function of circuit depth. Propagation efficiency for each cascade is calculated as the ratio between the fold change of the output RFP compared to the fold change of the input layer (Methods 7.2). Signal delay is calculated as the difference in time to reach the maximum fold-activation of the cascade between the cascade output and input layer (Methods 7.2).

F. Construction of wide CRISPRa/i circuits. **Left:** Schematic of CRISPRa cascades with increasing width. Input and output node concentrations were held constant across all cascades at 0 or 15 pM and 10 nM, respectively. **Right:** The concentration of each internal node was held at 0.2 nM as circuit width increased (blue), or the internal node concentration was scaled down proportionally to the width of the circuit (red), such that each internal node concentration is 0.2/n nM, where n is the number of parallel cascades. Fold-activation is given relative to a single CRISPRa cascade (Methods 7.2). Asterisks indicate a statistically significant difference using a two-tailed unpaired Welch's t-test ($*p$ value < 0.05, $**p$ value < 0.01, $***p$ value < 0.001).

For all panels, values represent the mean ± standard deviation of three technical replicates.

Figure 4: Developing activatable promoters for blue-light responsive CRISPRa/i circuits



A. Schematic of EL222 light-responsive promoter system and library design. EL222 transcription factor dimerizes in response to 470 nm light and binds a specific sequence upstream of the minimal promoter. EL222 then recruits RNAP through interactions with the α-CTD domain. Minimal promoter library design is based on the original *luxI* promoter and previous minimal promoter libraries (Methods 3.4).

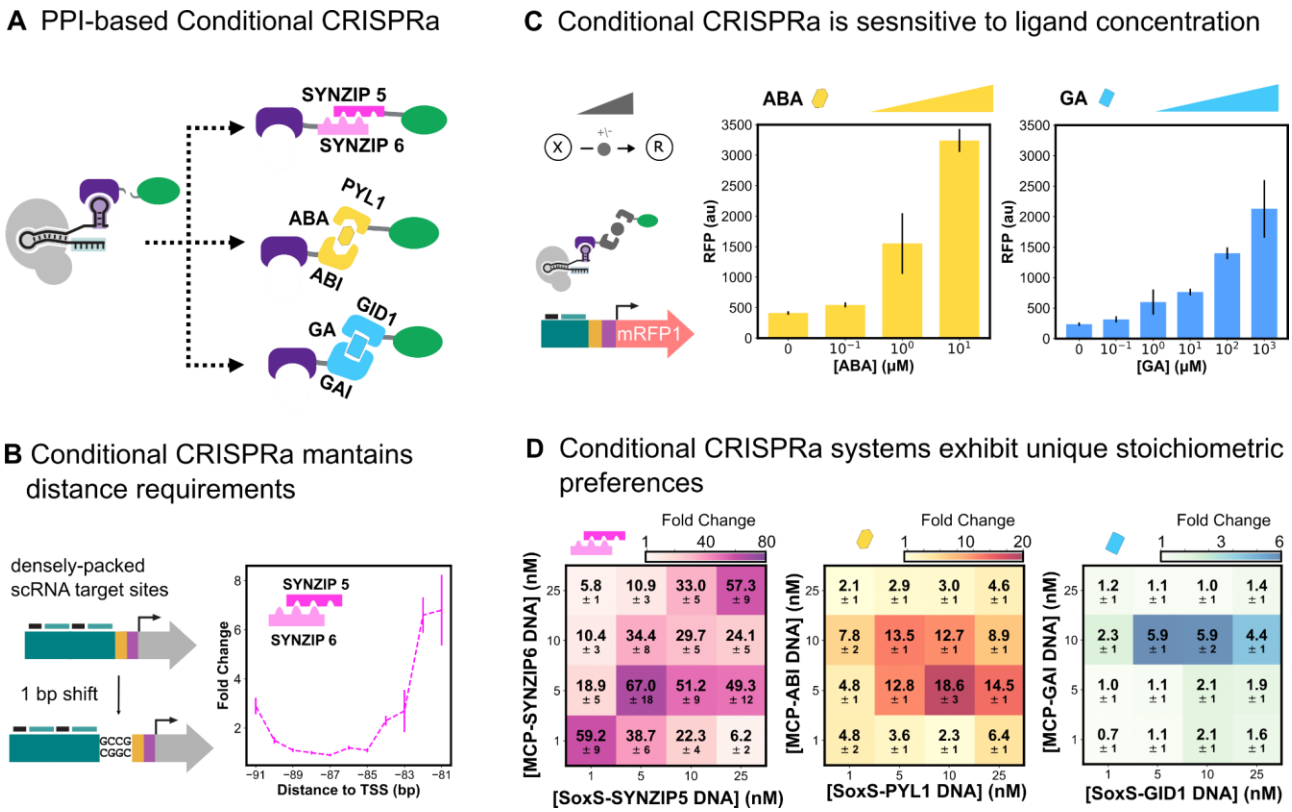
B. Characterization of light responsive promoters in *E. coli*. **Left:** Schematic of blue-light promoter screening (Methods 5). EL222 protein and promoter library are expressed from a single plasmid. Assembly and screening are carried out as previously described. Basal and activated expression levels are measured from

cultures not exposed or continuously exposed to blue-light, respectively. **Right:** Basal expression and dynamic range of blue-light promoter variants. Gray dash line defines promoter variants with equal activated and basal expression levels, indicating they are not activated by EL222. The J23119 minimal promoter (red) and J23113 (black) are examples of non-activatable promoters. Variants with improved performance (red circles) compared to the original *luxI* promoter (green) were selected for use in CFS.

C. Light-responsive CRISPRa in CFS. **Left:** Schematic of EL222 scRNA expression and downstream CRISPRa. is achieved by expressing a scRNA from an engineered blue-light promoter. **Right:** Titration of blue-light inducible scRNA plasmid concentration to maximize the fold change between blue-light dependent CRISPRa (blue) and CRISPRa due to scRNA leak in the dark (black). Values represent the mean ± standard deviation of three technical replicates.

D. Improvement of blue-light CRISPRa dynamic range through the construction of a positive feedback circuit. **Left:** Schematic of a blue-light responsive CRISPRa cascade with positive feedback (PFB). PFB is achieved by including a downstream node that expresses a scRNA targeting an upstream node. **Right:** Blue-light dependent CRISPRa (blue) and CRISPRa due to scRNA leak in the dark (black). The amount of positive feedback was tuned by adjusting the concentration of the PFB node. Values represent the mean ± standard deviation of three technical replicates.

Figure 5: Engineered activatable promoters enable PPI-dependent conditional CRISPRa



A. Schematic of different conditional CRISPRa systems. MCP-SoxS fusion is split and the two proteins are instead fused to one end of a heterodimerization domain. The heterodimerization domains used to build conditional CRISPRa systems are the SYNZIP5/SYNZIP6 pair, the abscisic acid (ABA)-responsive ABI/PYL1 domain, and the gibberellic acid (GA)-responsive GAI/GID1 domain.

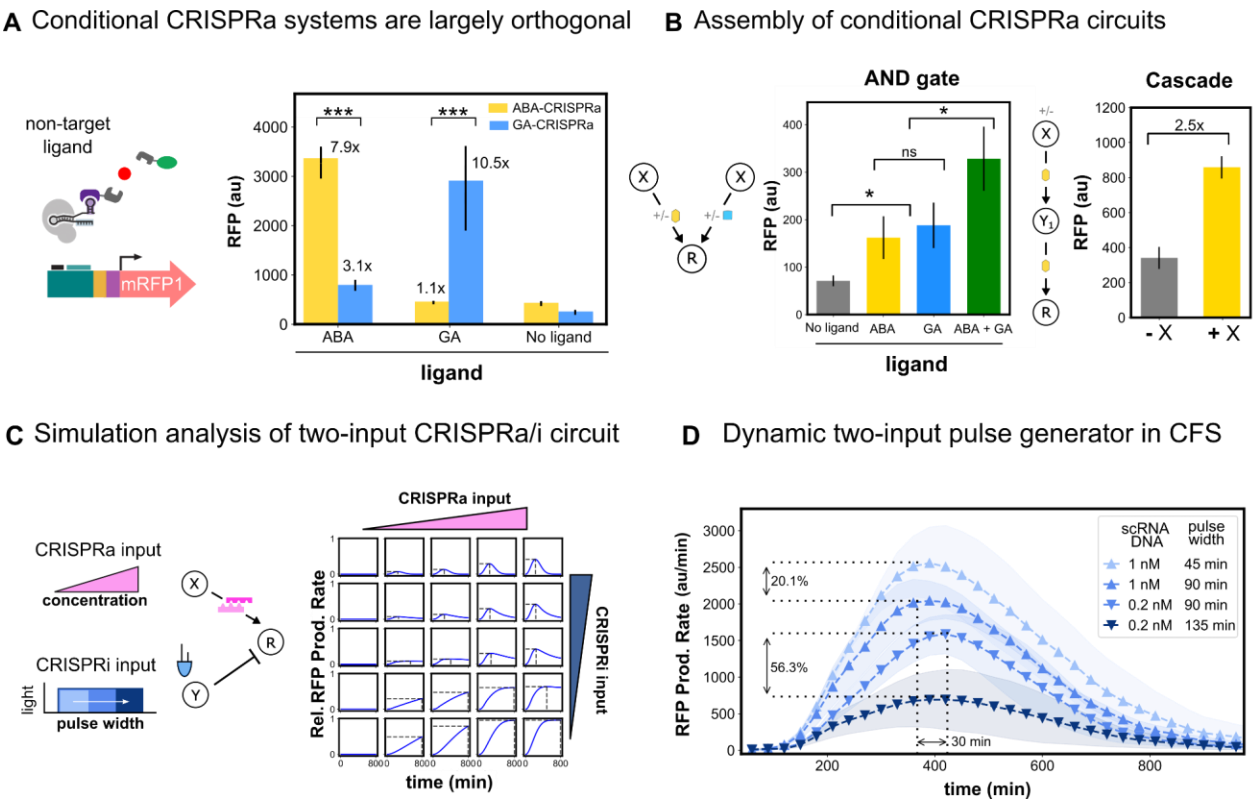
B. Distance requirements of conditional CRISPRa. **Left:** An engineered promoter containing densely packed scRNA target sites and single base pair additions allows for CRISPRa targeting with single base pair resolution between -81 and -111 bp from the TSS. Phase shifts are introduced at the 5' end of the promoter UP-element. **Right:** Testing SYNZIP-CRISPRa with single base pair resolution between -81 and -91 bp from the TSS. SYNZIP-CRISPRa components are

expressed at 5 nM. Fold change is calculated relative to an off-target scRNA for each promoter variant.

C. Tuning conditional CRISPRa response through titration of small molecule concentration. For ABA- and GA- CRISPRa, the corresponding small molecule was titrated between 0 and 10 or 0 and 10^3 μM respectively to find the optimal concentration. ABA- and GA-CRISPRa components are expressed at 5 nM.

D. Improving conditional CRISPRa response by optimizing component stoichiometries. The concentration of the plasmids expressing the MCP and SoxS components for each dimerization system were varied between 1 and 25 nM and tested combinatorially to find the best ratio of the two heterodimerization proteins. ABA is added at 10 μM and GA is added at 10^3 μM. Fold change is given relative to a reaction with no MCP and SoxS plasmids added.

Figure 6: Engineering multi-input CRISPRa/i Circuits



720 A. Conditional CRISPRa response to non-cognate ligands. The orthogonality of the small molecule-responsive conditional CRISPRa systems was tested by adding
722 either the corresponding or non-corresponding small molecule to cell-free reactions containing the machinery for abscisic acid (ABA)- or gibberellic acid
724 (GA)- CRISPRa. The components for both ABA- and GA- CRISPRa are added at their respective optimal screened concentrations. ABA is added at 10 μM and GA is added at 10^3 μM.

B. Assembly of conditional CRISPRa circuits. For both circuits, the components for
728 both ABA- and GA- CRISPRa are added at their respective optimal screened concentrations. ABA is added at 10 μM and GA is added at 10^3 μM. **Left:** The AND
730 gate was constructed by adding the components for both ABA- and GA- CRISPRa

in a cell-free reaction. **Right:** The CRISPRa cascade was assembled by using ABA-CRISPRa to activate expression of both the first and second node in an activation cascade. The first node was added at either 0.05 or 0 nM, and the internal and output nodes were added at 10 nM.

C. Simulation analysis of a two-input CRISPRa/i circuit using SYNZIP5/SYNZIP6 heterodimerization mediated-CRISPRa and blue-light CRISPRi. The simulation sampled different amounts of CRISPRa and CRISPRi input to analyze whether the two systems could be integrated to modulate the production rate of a desired gene target (Methods 6).

D. SYNZIP-CRISPRa and blue-light CRISPRi were integrated to construct a tunable pulse generator. The amount of CRISPRa was tuned by adding either 0.2 nM or 1 nM of constitutively expressed scRNA plasmid to the CFS reaction. The sgRNA targeting RFP for CRISPRi was driven from the blue-light responsive engineered EL222 promoter. The amount of CRISPRi was tuned by adjusting the time of blue-light exposure between 45 and 135 min. RFP production rates are plotted as a function of CRISPRa and CRISPRi inputs (Methods 7.2).

For all panels, values represent the mean ± standard deviation of three technical replicates. Asterisks indicate a statistically significant difference using a two-tailed unpaired Welch's t-test ($*p$ value < 0.05, $**p$ value < 0.01, $***p$ value < 0.001).

752

Method Details

1. Plasmid and Library Construction

1.1 Cloning

All PCR amplification of plasmids and fragments used Phusion DNA polymerase in GC buffer. Primers were synthesized by IDT and resuspended into nuclease-free water. All PCR reactions were treated with DpnI for longer than 1 hour and purified using Qiagen gel extraction kits. Plasmid assembly was achieved using 5X In-Fusion HD mastermix (Takara).

Assembled plasmids and libraries were transformed into chemically competent NEB Turbo *E. coli* and plated onto LB-agar plates with either 100 µg/mL carbenicillin or 25 µg/mL chloramphenicol. Transformed cells were grown overnight ~16 hours at 37 °C. Single colonies were picked from plates and grown overnight in LB shaking at 37 °C with appropriate concentrations of relevant antibiotics.

For libraries, 10 uL of the outgrowth was diluted 1:20 with LB and plated onto LB-agar with carbenicillin or gentamicin to check library complexity. The remaining outgrowth was seeded into 5 mL of LB with carbenicillin or gentamicin. Transformed cells were grown overnight ~16 hours at 37 °C. Single colonies were picked from plates and grown overnight in LB with carbenicillin or gentamicin. Both single colonies and culture were sequence verified.

Plasmids were isolated from subcultures using a DNA miniprep kit (QIAprep Spin Miniprep Kit) and Sanger sequenced (Genewiz inc.) to identify correctly assembled plasmids.

2. *E. coli* Experiments

In all *E. coli* experiments CRISPRa system components (dCas9, MCP-SoxS, scRNA) are located on a p15A ori plasmid while reporter construct is located on a pSC101** ori plasmid. Plate reader measurements were conducted using a BioTek Synergy HTX with a black flat bottom plate (Ref# 3631) using 100 µL of culture. Transformed *E. coli* were outgrown for 1 hour shaking at 37 °C and plated onto LB-agar with carbenicillin and chloramphenicol. Plates were grown overnight at 37 °C. Experiments were conducted by picking three individual colonies into 400 µL Teknova EZ-RDM with 0.2% glucose and appropriate antibiotics in 96 well plates, covering with breathable membrane (Breathe Easier cat# Z763624) and shaking overnight at 37 °C at 1200 RPM on a Heidolph Titramax 1000. For inducible experiments, overnight cultures are subsequently diluted 1:40 into a fresh plate of EZ-RDM and supplemented with appropriate concentrations of aTc. Measurements are conducted in Costar 96 well black flat bottom plates in 100 uL culture volume.

3. Design of Promoter Region Libraries

3.1 Minimal Promoter Libraries

Of the two libraries, the first was designed by rationally mutagenizing specific bases that are known contacts of RNAP within the minimal promoter. The second library was made by randomly mutagenizing within the intervening sequence. Since the libraries yielded similar Pareto fronts, we then combined these mutations into a third minimal promoter library, which was used in the sequential screening process.

3.2 UP-Element Libraries

We designed five UP-element libraries mutagenizing the AT-rich *E. coli* consensus sequence with increasing GC-content. We generated 5 libraries from 0% to 100% GC-content, and a library representing the *E. coli* consensus sequence (Table S3).

3.3 scRNA Target Site Libraries

We generated three scRNA target site libraries with varying compositions of GC-content (0%, 50%, and 100%) (Table S3). These libraries were used in tandem with a GC-rich UP-element.

3.4 EL222 Minimal Promoter Libraries

Starting with the native *luxI* minimal promoter, we introduced rational mutations to make it resemble a synthetic activatable promoter (J23117). We then randomly mutagenized within the -10:-35 region (Table S3).

4. Cell-Free Reactions

4.1 Cell-Free System Preparation

The cell-free system was acquired from Arbor Biosciences (myTXTL). The cell-free system used for an experiment was thawed on ice and pooled into a 1.5 ml Eppendorf tube, vortexed, and spun-down using a mini benchtop centrifuge to ensure homogeneity across samples. Details about plasmid preparation are provided in Methods S1.

4.2 Cell-Free Gene Expression Reaction

Cell-free gene expression reactions were assembled on ice from the CFS and purified DNA. A master mix with common plasmids across reactions was prepared, and 1.5 µL per reaction allocated into PCR tubes. Plasmids which were varied across reactions were

added in the remaining 1 µL. For reactions containing ABA (Sigma, A4906) or GA, .1 µL of the small molecules were added alongside the plasmids. For reactions involving more than 5 plasmids, plasmids were mixed with an acoustic liquid handler robot (Echo Labcyte 525). The CFS was pipette mixed and added to each PCR tube in 7.5 µL for a final volume of 10 µL. PCR tubes were vortexed, spun-down using a mini benchtop centrifuge, and placed on ice. Triplicates of 2.5 µL for each reaction were pipetted into individual wells of a 96-well V-bottom plate (Costar, Cat. 3363). The plate was sealed (Costar, Cat. 3080) and analyzed on a BioTek Synergy HTX plate reader at 29 °C. mRFP1 fluorescence (ex. 540 nm, em. 600 nm) of cell-free reactions were measured every 10 min from the bottom of the plate. All reactions were run in batch mode.

5. Optogenetic Experiments

For both *E. coli* and CFS optogenetic experiments, the cultures and reactions were prepared as described above. The incubation conditions were modified to include a blue-light illumination source (UVP Visi-Blue UV Transilluminator, 8 Watts, 460/470 nm). The samples were placed at 37 °C or 29 °C in an incubator (Thermo Forma Orbital Shaker, Model #435) with the illumination source placed atop the incubator and irradiating inwards. The distance between the illumination source and the deepwell plates is 14 cm. CFS reactions were placed inside the incubator at 29 °C at a distance of 6 cm with the bottom of the wells facing the illumination source. In both cases, the dark conditions were kept inside a cardboard box inside the incubator. Endpoint plate reader measurements were conducted using a BioTek Synergy HTX.

6. CFS Blue-light CRISPRa/i modeling

The CFS blue-light CRISPRa/i model was expanded from the previously described CFS CRISPRa/i model [32]. The model constitutes a series of first order chemical reactions for protein and guide RNA production, CRISPR complex assembly, and DNA targeting. All model details are described in Methods S2.

7. Quantification and statistical analysis

7.1 Data analysis

Throughout this work all measured RFP levels in *E. coli* were normalized by measured OD600 with appropriate propagation of uncertainties. All metrics are described in Methods S3.

7.2 Statistics

Statistical significance was calculated using two-tailed unpaired Welch's *t*-tests. Asterisks in Figures indicate a statistically significant difference ($*$: p-value < 0.05, $**$: p-value < 0.01, $***$: p-value < 0.001).

Author Information

Corresponding Authors

James M. Carothers - *Department of Chemical Engineering, Molecular Engineering & Sciences Institute and Center for Synthetic Biology, University of Washington, Seattle, WA 98195*; Email: jcaroth@uw.edu

Jesse G. Zalatan- *Department of Chemistry, Molecular Engineering & Sciences Institute and Center for Synthetic Biology, University of Washington, Seattle, WA 98195*; Email: zalatan@uw.edu

Authors

Diego Alba Burbano - *Department of Chemical Engineering, Molecular Engineering & Sciences Institute and Center for Synthetic Biology, University of Washington, Seattle, WA 98195*

Ryan Cardiff - *Molecular Engineering & Sciences Institute and Center for Synthetic Biology, University of Washington, Seattle, WA 98195*

Benjamin I. Tickman - *Department of Chemical Engineering, Molecular Engineering & Sciences Institute and Center for Synthetic Biology, University of Washington, Seattle, WA 98195*

Cholpisit Kiattisewee - *Molecular Engineering & Sciences Institute and Center for Synthetic Biology, University of Washington, Seattle, WA 98195*

Cassandra Maranas - *Molecular Engineering & Sciences Institute and Center for Synthetic Biology, University of Washington, Seattle, WA 98195*

Author Contributions

D.A.B., R.C., B.I.T., J.G.Z., and J.M.C. designed the research. D.A.B., R.C., and C.M. performed experiments. D.A.B., R.C., B.I.T., C.K., and C.M. analyzed the data. D.A.B., R.C., J.G.Z., and J.M.C. wrote the manuscript with input from all of the authors.

Acknowledgements

We thank members of the Carothers and Zalatan groups for advice, materials, and comments on the manuscript.

Competing interests

D.A.B., R.C., J.G.Z., and J.M.C are inventors on patents and/or patent applications filed by the University of Washington that describe promoter engineering, conditional CRISPRa technologies, and CRISPRa/i circuits in prokaryotic systems. J.G.Z. and J.M.C are members of the Wayfinder Biosciences scientific advisory board.

Data sharing plans

Data supporting the findings of this work are available within the paper and its supporting information files. All data and code are available on GitHub at: https://github.com/carothersresearch/CRISPRai_Circuits_2022

Funding information

This work was supported by US National Science Foundation (NSF) Award MCB 2032794 (to J.M.C. and J.G.Z.), NSF Award CBET 1844152 (to J.M.C.), NSF Award EF-1935087 (to J.M.C.), US Department of Energy (DOE) BETO DE-EE0008927 (to J.M.C. and J.G.Z.) and DOE ARPA-E DE-AR00002387-1567 (to J.M.C). This material is based upon work supported by the National Science Foundation Graduate Research Fellowship Program under Grant No. DGE-2140004 (D.A.B.). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

[1] H.-W. Ma, B. Kumar, U. Ditges, F. Gunzer, J. Buer, and A.-P. Zeng, “An extended transcriptional regulatory network of Escherichia coli and analysis of its hierarchical structure and network motifs,” *Nucleic Acids Res.*, vol. 32, no. 22, pp. 6643–6649, 2004, doi: 10.1093/nar/gkh1009.

[2] L. Cai, C. K. Dalal, and M. B. Elowitz, “Frequency-modulated nuclear localization bursts coordinate gene regulation,” *Nature*, vol. 455, no. 7212, Art. no. 7212, Sep. 2008, doi: 10.1038/nature07292.

[3] D. Benzinger, S. Ovinnikov, and M. Khammash, “Synthetic gene networks recapitulate dynamic signal decoding and differential gene expression,” *Cell Syst.*, p. S2405471222000825, Mar. 2022, doi: 10.1016/j.cels.2022.02.004.

[4] J. K. Jung *et al.*, “Cell-free biosensors for rapid detection of water contaminants,” *Nat. Biotechnol.*, vol. 38, no. 12, Art. no. 12, Dec. 2020, doi: 10.1038/s41587-020-0571-7.

[5] A. S. Karim *et al.*, “In vitro prototyping and rapid optimization of biosynthetic enzymes for cell design,” *Nat. Chem. Biol.*, vol. 16, no. 8, pp. 912–919, Aug. 2020, doi: 10.1038/s41589-020-0559-0.

[6] F.-X. Lehr *et al.*, “Cell-Free Prototyping of AND-Logic Gates Based on Heterogeneous RNA Activators,” *ACS Synth. Biol.*, vol. 8, no. 9, pp. 2163–2173, Sep. 2019, doi: 10.1021/acssynbio.9b00238.

[7] S. J. Moore *et al.*, “Rapid acquisition and model-based analysis of cell-free transcription–translation reactions from nonmodel bacteria,” *Proc. Natl. Acad. Sci.*, vol. 115, no. 19, pp. E4340–E4349, May 2018, doi: 10.1073/pnas.1715806115.

[8] Z. Swank and S. J. Maerkl, “CFPU: A Cell-Free Processing Unit for High-Throughput, Automated In Vitro Circuit Characterization in Steady-State Conditions,” *BioDesign Res.*, vol. 2021, Mar. 2021, doi: 10.34133/2021/2968181.

[9] M. C. Bassalo, R. Liu, and R. T. Gill, “Directed evolution and synthetic biology applications to microbial systems,” *Curr. Opin. Biotechnol.*, vol. 39, pp. 126–133, Jun. 2016, doi: 10.1016/j.copbio.2016.03.016.

[10] J. A. N. Brophy and C. A. Voigt, “Principles of genetic circuit design,” *Nat. Methods*, vol. 11, no. 5, Art. no. 5, May 2014, doi: 10.1038/nmeth.2926.

[11] Z. Swank, N. Laohakunakorn, and S. J. Maerkl, “Cell-free gene-regulatory network engineering with synthetic transcription factors,” *Proc. Natl. Acad. Sci.*, vol. 116, no. 13, pp. 5892–5901, Mar. 2019, doi: 10.1073/pnas.1816591116.

[12] P.-F. Xia, H. Ling, J. L. Foo, and M. W. Chang, “Synthetic genetic circuits for programmable biological functionalities,” *Biotechnol. Adv.*, vol. 37, no. 6, p. 107393, Nov. 2019, doi: 10.1016/j.biotechadv.2019.04.015.

[13] U. Alon, “Network motifs: theory and experimental approaches,” *Nat. Rev. Genet.*, vol. 8, no. 6, pp. 450–461, Jun. 2007, doi: 10.1038/nrg2102.

[14] J. K. Jung *et al.*, “Cell-free biosensors for rapid detection of water contaminants,” *Nat. Biotechnol.*, vol. 38, no. 12, pp. 1451–1459, Dec. 2020, doi: 10.1038/s41587-020-0571-7.

[15] J. K. Jung, C. M. Archuleta, K. K. Alam, and J. B. Lucks, “Programming cell-free biosensors with DNA strand displacement circuits,” *Nat. Chem. Biol.*, vol. 18, no. 4, pp. 385–393, Apr. 2022, doi: 10.1038/s41589-021-00962-9.

[16] P. Q. Nguyen *et al.*, “Wearable materials with embedded synthetic biology sensors for biomolecule detection,” *Nat. Biotechnol.*, vol. 39, no. 11, pp. 1366–1374, Nov. 2021, doi: 10.1038/s41587-021-00950-3.

[17] K. Pardee *et al.*, “Portable, On-Demand Biomolecular Manufacturing,” *Cell*, vol. 167, no. 1, pp. 248-259.e12, Sep. 2016, doi: 10.1016/j.cell.2016.09.013.

[18] J. C. Stark *et al.*, “On-demand biomanufacturing of protective conjugate vaccines,” *Sci. Adv.*, vol. 7, no. 6, p. eabe9444, Feb. 2021, doi: 10.1126/sciadv.abe9444.

[19] C. Zhou, X. Lin, Y. Lu, and J. Zhang, “Flexible on-demand cell-free protein synthesis platform based on a tube-in-tube reactor,” *React. Chem. Eng.*, vol. 5, no. 2, pp. 270–277, 2020, doi: 10.1039/C9RE00394K.

[20] K. P. Adamala, D. A. Martin-Alarcon, K. R. Guthrie-Honea, and E. S. Boyden, “Engineering genetic circuit interactions within and between synthetic minimal cells,” *Nat. Chem.*, vol. 9, no. 5, pp. 431–439, May 2017, doi: 10.1038/nchem.2644.

[21] L. Aufinger, J. Brenner, and F. C. Simmel, “Complex dynamics in a synchronized cell-free genetic clock,” *Nat. Commun.*, vol. 13, no. 1, Art. no. 1, May 2022, doi: 10.1038/s41467-022-30478-2.

[22] J. Garamella, D. Garenne, and V. Noireaux, “TXTL-based approach to synthetic cells,” *Methods Enzymol.*, vol. 617, pp. 217–239, 2019, doi: 10.1016/bs.mie.2018.12.015.

[23] E. Karzbrun, A. M. Tayar, V. Noireaux, and R. H. Bar-Ziv, “Synthetic biology. Programmable on-chip DNA compartments as artificial cells,” *Science*, vol. 345, no. 6198, pp. 829–832, Aug. 2014, doi: 10.1126/science.1255550.

[24] D. Banerjee *et al.*, “Genome-scale metabolic rewiring improves titers rates and yields of the non-native product indigoidine at scale,” *Nat. Commun.*, vol. 11, no. 1, p. 5385, Dec. 2020, doi: 10.1038/s41467-020-19171-4.

[25] H. Kim, D. Bojar, and M. Fussenegger, “A CRISPR/Cas9-based central processing unit to program complex logic computation in human cells,” *Proc. Natl. Acad. Sci.*, vol. 116, no. 15, pp. 7214–7219, Apr. 2019, doi: 10.1073/pnas.1821740116.

[26] J. Landberg, N. R. Wright, T. Wulff, M. J. Herrgård, and A. T. Nielsen, “CRISPR

interference of nucleotide biosynthesis improves production of a single-domain antibody in Escherichia coli," *Biotechnol. Bioeng.*, vol. 117, no. 12, pp. 3835–3848, 2020, doi: 10.1002/bit.27536.

[27] A. C. Reis *et al.*, "Simultaneous repression of multiple bacterial genes using nonrepetitive extra-long sgRNA arrays," *Nat. Biotechnol.*, vol. 37, no. 11, pp. 1294–1301, Nov. 2019, doi: 10.1038/s41587-019-0286-9.

[28] J. Santos-Moreno and Y. Schaerli, "CRISPR-based gene expression control for synthetic gene circuits," *Biochem. Soc. Trans.*, vol. 48, no. 5, pp. 1979–1993, Oct. 2020, doi: 10.1042/BST20200020.

[29] A. Westbrook *et al.*, "Distinct timescales of RNA regulators enable the construction of a genetic pulse generator," *Biotechnol. Bioeng.*, vol. 116, no. 5, pp. 1139–1151, May 2019, doi: 10.1002/bit.26918.

[30] S. Clamons and R. Murray, "Modeling predicts that CRISPR-based activators, unlike CRISPR-based repressors, scale well with increasing gRNA competition and dCas9 bottlenecking," Synthetic Biology, preprint, Jul. 2019. doi: 10.1101/719278.

[31] J. Nielsen and J. D. Keasling, "Engineering Cellular Metabolism," *Cell*, vol. 164, no. 6, pp. 1185–1197, Mar. 2016, doi: 10.1016/j.cell.2016.02.004.

[32] B. I. Tickman *et al.*, "Multi-layer CRISPRa/i circuits for dynamic genetic programs in cell-free and bacterial systems," *Cell Syst.*, p. S2405471221004191, Nov. 2021, doi: 10.1016/j.cels.2021.10.008.

[33] M. W. Gander, J. D. Vrana, W. E. Voje, J. M. Carothers, and E. Klavins, "Digital logic circuits in yeast with CRISPR-dCas9 NOR gates," *Nat. Commun.*, vol. 8, no. 1, Art. no. 1, May 2017, doi: 10.1038/ncomms15459.

[34] R. Marshall and V. Noireaux, "Quantitative modeling of transcription and translation of an all- E. coli cell-free system," *Sci. Rep.*, vol. 9, no. 1, Art. no. 1, Aug. 2019, doi: 10.1038/s41598-019-48468-8.

[35] H. Alper, C. Fischer, E. Nevoigt, and G. Stephanopoulos, "Tuning genetic control through promoter engineering," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 102, no. 36, pp. 12678–12683, Sep. 2005, doi: 10.1073/pnas.0504604102.

[36] J. R. Kelly *et al.*, "Measuring the activity of BioBrick promoters using an in vivo reference standard," *J. Biol. Eng.*, vol. 3, p. 4, Mar. 2009, doi: 10.1186/1754-1611-3-4.

[37] R. C. Brewster, D. L. Jones, and R. Phillips, "Tuning Promoter Strength through RNA Polymerase Binding Site Design in Escherichia coli," *PLOS Comput. Biol.*, vol. 8, no. 12, p. e1002811, Dec. 2012, doi: 10.1371/journal.pcbi.1002811.

[38] T. L. Fleur, A. Hossain, and H. M. Salis, "Automated Model-Predictive Design of Synthetic Promoters to Control Transcriptional Profiles in Bacteria," Synthetic Biology, preprint, Sep. 2021. doi: 10.1101/2021.09.01.458561.

[39] D. J. Lee, S. D. Minchin, and S. J. W. Busby, “Activating Transcription in Bacteria,” *Annu. Rev. Microbiol.*, vol. 66, no. 1, pp. 125–152, Oct. 2012, doi: 10.1146/annurev-micro-092611-150012.

[40] T. C. Yu *et al.*, “Multiplexed characterization of rationally designed promoter architectures deconstructs combinatorial logic for IPTG-inducible systems,” *Nat. Commun.*, vol. 12, no. 1, p. 325, Jan. 2021, doi: 10.1038/s41467-020-20094-3.

[41] B. F. Cress *et al.*, “Rapid generation of CRISPR/dCas9-regulated, orthogonally repressible hybrid T7-lac promoters for modular, tuneable control of metabolic pathway fluxes in Escherichia coli,” *Nucleic Acids Res.*, vol. 44, no. 9, pp. 4472–4485, May 2016, doi: 10.1093/nar/gkw231.

[42] T. M. Groseclose, R. E. Rondon, Z. D. Herde, C. A. Aldrete, and C. J. Wilson, “Engineered systems of inducible anti-repressors for the next generation of biological programming,” *Nat. Commun.*, vol. 11, no. 1, Art. no. 1, Sep. 2020, doi: 10.1038/s41467-020-18302-1.

[43] X. Liu *et al.*, “De novo design of programmable inducible promoters,” *Nucleic Acids Res.*, vol. 47, no. 19, pp. 10452–10463, Nov. 2019, doi: 10.1093/nar/gkz772.

[44] I. J. Roney, A. D. Rudner, J.-F. Couture, and M. Kærn, “Improvement of the reverse tetracycline transactivator by single amino acid substitutions that reduce leaky target gene expression to undetectable levels,” *Sci. Rep.*, vol. 6, p. 27697, Jun. 2016, doi: 10.1038/srep27697.

[45] J. Fontana *et al.*, “Effective CRISPRa-mediated control of gene expression in bacteria must overcome strict target site requirements,” *Nat. Commun.*, vol. 11, no. 1, p. 1618, Dec. 2020, doi: 10.1038/s41467-020-15454-y.

[46] Y. Chen *et al.*, “Tuning the dynamic range of bacterial promoters regulated by ligand-inducible transcription factors,” *Nat. Commun.*, vol. 9, no. 1, p. 64, Dec. 2018, doi: 10.1038/s41467-017-02473-5.

[47] W. Ross, S. E. Aiyar, J. Salomon, and R. L. Gourse, “Escherichia coli promoters with UP elements of different strengths: modular structure of bacterial promoters,” *J. Bacteriol.*, vol. 180, no. 20, pp. 5375–5383, Oct. 1998, doi: 10.1128/JB.180.20.5375-5383.1998.

[48] R. K. Shultzaberger, Z. Chen, K. A. Lewis, and T. D. Schneider, “Anatomy of Escherichia coli σ 70 promoters,” *Nucleic Acids Res.*, vol. 35, no. 3, pp. 771–788, Feb. 2007, doi: 10.1093/nar/gkl956.

[49] C. Dong, J. Fontana, A. Patel, J. M. Carothers, and J. G. Zalatan, “Synthetic CRISPR-Cas gene activators for transcriptional reprogramming in bacteria,” *Nat. Commun.*, vol. 9, no. 1, Art. no. 1, Jun. 2018, doi: 10.1038/s41467-018-04901-6.

[50] I. G. Hook-Barnard and D. M. Hinton, “Transcription Initiation by Mix and Match Elements: Flexibility for Polymerase Binding to Bacterial Promoters,” *Gene Regul.*

Syst. Biol., vol. 1, pp. 275–293, Dec. 2007.

[51] M. S. Paget and J. D. Helmann, “The σ70family of sigma factors,” *Genome Biol.*, vol. 4, no. 1, p. 203, Jan. 2003, doi: 10.1186/gb-2003-4-1-203.

[52] Y. Censor, “Pareto optimality in multiobjective problems,” *Appl. Math. Optim.*, vol. 4, no. 1, pp. 41–59, Mar. 1977, doi: 10.1007/BF01442131.

[53] S. T. Estrem, T. Gaal, W. Ross, and R. L. Gourse, “Identification of an UP element consensus sequence for bacterial promoters,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 95, no. 17, pp. 9761–9766, Aug. 1998.

[54] J. Fontana, D. Sparkman-Yager, J. G. Zalatan, and J. M. Carothers, “Challenges and opportunities with CRISPR activation in bacteria for data-driven metabolic engineering,” *Curr. Opin. Biotechnol.*, vol. 64, pp. 190–198, Aug. 2020, doi: 10.1016/j.copbio.2020.04.005.

[55] J. Lian, M. HamediRad, S. Hu, and H. Zhao, “Combinatorial metabolic engineering using an orthogonal tri-functional CRISPR system,” *Nat. Commun.*, vol. 8, no. 1, p. 1688, Nov. 2017, doi: 10.1038/s41467-017-01695-x.

[56] R. Kelwick, L. Ricci, S. M. Chee, D. Bell, A. J. Webb, and P. S. Freemont, “Cell-free prototyping strategies for enhancing the sustainable production of polyhydroxyalkanoates bioplastics,” *Synth. Biol.*, vol. 3, no. 1, p. ysy016, Jan. 2018, doi: 10.1093/synbio/ysy016.

[57] B. Vögeli *et al.*, “Cell-free prototyping enables implementation of optimized reverse β-oxidation pathways in heterotrophic and autotrophic bacteria,” *Nat. Commun.*, vol. 13, no. 1, Art. no. 1, Jun. 2022, doi: 10.1038/s41467-022-30571-6.

[58] D. Camsund, A. Jaramillo, and P. Lindblad, “Engineering of a Promoter Repressed by a Light-Regulated Transcription Factor in Escherichia coli,” *BioDesign Res.*, vol. 2021, Sep. 2021, doi: 10.34133/2021/9857418.

[59] P. Jayaraman, K. Devarajan, T. K. Chua, H. Zhang, E. Gunawan, and C. L. Poh, “Blue light-mediated transcriptional activation and repression of gene expression in bacteria,” *Nucleic Acids Res.*, vol. 44, no. 14, pp. 6994–7005, Aug. 2016, doi: 10.1093/nar/gkw548.

[60] L. B. Motta-Mena *et al.*, “An optogenetic gene expression system with rapid activation and deactivation kinetics,” *Nat. Chem. Biol.*, vol. 10, no. 3, pp. 196–202, Mar. 2014, doi: 10.1038/nchembio.1430.

[61] B. D. Zoltowski, L. B. Motta-Mena, and K. H. Gardner, “Blue light-induced dimerization of a bacterial LOV-HTH DNA-binding protein,” *Biochemistry*, vol. 52, no. 38, pp. 6653–6661, Sep. 2013, doi: 10.1021/bi401040m.

[62] Y. Gao, X. Xiong, S. Wong, E. J. Charles, W. A. Lim, and L. S. Qi, “Complex transcriptional modulation with orthogonal and inducible dCas9 regulators,” *Nat. Methods*, vol. 13, no. 12, pp. 1043–1049, Dec. 2016, doi: 10.1038/nmeth.4042.

[63] P. K. Jain *et al.*, “Development of Light-Activated CRISPR Using Guide RNAs with Photocleavable Protectors,” *Angew. Chem. Int. Ed.*, vol. 55, no. 40, pp. 12440–12444, 2016, doi: 10.1002/anie.201606123.

[64] K. Kundert *et al.*, “Controlling CRISPR-Cas9 with ligand-activated and ligand-deactivated sgRNAs,” *Nat. Commun.*, vol. 10, no. 1, Art. no. 1, May 2019, doi: 10.1038/s41467-019-09985-2.

[65] Y. Nihongaki, F. Kawano, T. Nakajima, and M. Sato, “Photoactivatable CRISPR-Cas9 for optogenetic genome editing,” *Nat. Biotechnol.*, vol. 33, no. 7, pp. 755–760, Jul. 2015, doi: 10.1038/nbt.3245.

[66] L. R. Polstein and C. A. Gersbach, “A light-inducible CRISPR-Cas9 system for control of endogenous gene activation,” *Nat. Chem. Biol.*, vol. 11, no. 3, pp. 198–200, Mar. 2015, doi: 10.1038/nchembio.1753.

[67] Z.-M. Ying, F. Wang, X. Chu, R.-Q. Yu, and J.-H. Jiang, “Activatable CRISPR Transcriptional Circuits Generate Functional RNA for mRNA Sensing and Silencing,” *Angew. Chem. Int. Ed.*, vol. 59, no. 42, pp. 18599–18604, 2020, doi: 10.1002/anie.202004751.

[68] Y. Yu *et al.*, “Engineering a far-red light–activated split-Cas9 system for remote-controlled genome editing of internal organs and tumors,” *Sci. Adv.*, vol. 6, no. 28, p. eabb1777, Jul. 2020, doi: 10.1126/sciadv.abb1777.

[69] K. E. Thompson, C. J. Bashor, W. A. Lim, and A. E. Keating, “SYNZIP Protein Interaction Toolbox: in Vitro and in Vivo Specifications of Heterospecific Coiled-Coil Interaction Domains,” *ACS Synth. Biol.*, vol. 1, no. 4, pp. 118–129, Apr. 2012, doi: 10.1021/sb200015u.

[70] D. Cunningham-Bryant, J. Sun, B. Fernandez, and J. G. Zalatan, “CRISPR–Cas-Mediated Chemical Control of Transcriptional Dynamics in Yeast,” *ChemBioChem*, vol. 20, no. 12, pp. 1519–1523, 2019, doi: 10.1002/cbic.201800823.

[71] F.-S. Liang, W. Q. Ho, and G. R. Crabtree, “Engineering the ABA Plant Stress Pathway for Regulation of Induced Proximity,” *Sci. Signal.*, vol. 4, no. 164, p. rs2, Mar. 2011, doi: 10.1126/scisignal.2001449.

[72] T. Miyamoto *et al.*, “Rapid and Orthogonal Logic Gating with a Gibberellin-induced Dimerization System,” *Nat. Chem. Biol.*, vol. 8, no. 5, pp. 465–470, Mar. 2012, doi: 10.1038/nchembio.922.

[73] J. G. Zalatan *et al.*, “Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds,” *Cell*, vol. 160, no. 1, pp. 339–350, Jan. 2015, doi: 10.1016/j.cell.2014.11.052.

[74] M. C. Villegas Kcam, A. J. Tsong, and J. Chappell, “Rational engineering of a modular bacterial CRISPR–Cas activation platform with expanded target range,” *Nucleic Acids Res.*, vol. 49, no. 8, pp. 4793–4802, May 2021, doi:

10.1093/nar/gkab211.

[75] E. F. Jr. Douglass, C. J. Miller, G. Sparer, H. Shapiro, and D. A. Spiegel, “A Comprehensive Mathematical Model for Three-Body Binding Equilibria,” *J. Am. Chem. Soc.*, vol. 135, no. 16, pp. 6092–6099, Apr. 2013, doi: 10.1021/ja311795d.

[76] A. Levchenko, J. Bruck, and P. W. Sternberg, “Scaffold proteins may biphasically affect the levels of mitogen-activated protein kinase signaling and reduce its threshold properties,” *Proc. Natl. Acad. Sci.*, vol. 97, no. 11, pp. 5818–5823, May 2000, doi: 10.1073/pnas.97.11.5818.

[77] B. Wang, M. Barahona, and M. Buck, “Engineering modular and tunable genetic amplifiers for scaling transcriptional signals in cascaded gene networks,” *Nucleic Acids Res.*, vol. 42, no. 14, pp. 9484–9492, Aug. 2014, doi: 10.1093/nar/gku593.

[78] T. S. Lee *et al.*, “BglBrick vectors and datasheets: A synthetic biology platform for gene expression,” *J. Biol. Eng.*, vol. 5, no. 1, p. 12, Dec. 2011, doi: 10.1186/1754-1611-5-12.

[79] J. Garamella, R. Marshall, M. Rustad, and V. Noireaux, “The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology,” *ACS Synth. Biol.*, vol. 5, no. 4, pp. 344–355, Apr. 2016, doi: 10.1021/acssynbio.5b00296.

[80] D. A. Specht, L. B. Cortes, and G. Lambert, “Overcoming Leak Sensitivity in CRISPRi Circuits Using Antisense RNA Sequestration and Regulatory Feedback,” *ACS Synth. Biol.*, Aug. 2022, doi: 10.1021/acssynbio.2c00155.

[81] S. Kaplan, A. Bren, A. Zaslaver, E. Dekel, and U. Alon, “Diverse Two-Dimensional Input Functions Control Bacterial Sugar Genes,” *Mol. Cell*, vol. 29, no. 6, pp. 786–792, Mar. 2008, doi: 10.1016/j.molcel.2008.01.021.

[82] S. Krishna, L. Orosz, K. Sneppen, S. Adhya, and S. Semsey, “Relation of Intracellular Signal Levels and Promoter Activities in the gal Regulon of Escherichia coli,” *J. Mol. Biol.*, vol. 391, no. 4, pp. 671–678, Aug. 2009, doi: 10.1016/j.jmb.2009.06.043.

[83] T. M. Groseclose, A. N. Hersey, B. D. Huang, M. J. Realff, and C. J. Wilson, “Biological signal processing filters via engineering allosteric transcription factors,” *Proc. Natl. Acad. Sci.*, vol. 118, no. 46, Nov. 2021, doi: 10.1073/pnas.2111450118.

[84] C. J. Bashor *et al.*, “Complex signal processing in synthetic gene circuits using cooperative regulatory assemblies,” *Science*, vol. 364, no. 6440, pp. 593–597, May 2019, doi: 10.1126/science.aau8287.

[85] J. Beltrán *et al.*, “Rapid biosensor development using plant hormone receptors as reprogrammable scaffolds,” *Nat. Biotechnol.*, pp. 1–7, Jun. 2022, doi: 10.1038/s41587-022-01364-5.

[86] G. W. Foight *et al.*, “Multi-input chemical control of protein dimerization for programming graded cellular responses,” *Nat. Biotechnol.*, vol. 37, no. 10, Art. no. 10, Oct. 2019, doi: 10.1038/s41587-019-0242-8.

[87] G. Guntas *et al.*, “Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins,” *Proc. Natl. Acad. Sci.*, vol. 112, no. 1, pp. 112–117, Jan. 2015, doi: 10.1073/pnas.1417910112.

[88] Z. Huang *et al.*, “Creating Red Light-Switchable Protein Dimerization Systems as Genetically Encoded Actuators with High Specificity,” *ACS Synth. Biol.*, vol. 9, no. 12, pp. 3322–3333, Dec. 2020, doi: 10.1021/acssynbio.0c00397.

[89] D. I. Piraner, M. H. Abedi, B. A. Moser, A. Lee-Gosselin, and M. G. Shapiro, “Tunable thermal bioswitches for in vivo control of microbial therapeutics,” *Nat. Chem. Biol.*, vol. 13, no. 1, pp. 75–80, Jan. 2017, doi: 10.1038/nchembio.2233.

[90] J. Carey, V. Cameron, P. L. De Haseth, and O. C. Uhlenbeck, “Sequence-specific interaction of R17 coat protein with its ribonucleic acid binding site,” *Biochemistry*, vol. 22, no. 11, pp. 2601–2610, May 1983, doi: 10.1021/bi00280a002.

[91] F. Dupeux *et al.*, “A thermodynamic switch modulates abscisic acid receptor sensitivity,” *EMBO J.*, vol. 30, no. 20, pp. 4171–4184, Oct. 2011, doi: 10.1038/emboj.2011.294.

[92] K. Miyazono *et al.*, “Structural basis of abscisic acid signalling,” *Nature*, vol. 462, no. 7273, Art. no. 7273, Dec. 2009, doi: 10.1038/nature08583.

[93] M. Ueguchi-Tanaka *et al.*, “GIBBERELLIN INSENSITIVE DWARF1 encodes a soluble receptor for gibberellin,” *Nature*, vol. 437, no. 7059, Art. no. 7059, Sep. 2005, doi: 10.1038/nature04028.

[94] H. Yoshida *et al.*, “Evolution and diversification of the plant gibberellin receptor GID1,” *Proc. Natl. Acad. Sci.*, vol. 115, no. 33, pp. E7844–E7853, Aug. 2018, doi: 10.1073/pnas.1806040115.

[95] W. S. Grubbe, B. J. Rasor, A. Krüger, M. C. Jewett, and A. S. Karim, “Cell-free styrene biosynthesis at high titers,” *Metab. Eng.*, vol. 61, pp. 89–95, Sep. 2020, doi: 10.1016/j.ymben.2020.05.009.

[96] N. S. Kruyer *et al.*, “Membrane Augmented Cell-Free Systems: A New Frontier in Biotechnology,” *ACS Synth. Biol.*, vol. 10, no. 4, pp. 670–681, Apr. 2021, doi: 10.1021/acssynbio.0c00625.

[97] J. A. Peruzzi, N. R. Galvez, and N. P. Kamat, “Engineering transmembrane signal transduction in synthetic membranes using two-component systems,” Synthetic Biology, preprint, Oct. 2022. doi: 10.1101/2022.10.30.514420.

Chapter 4:

Automated Assembly of Large Kinetic Metabolic Models of Prokaryotic Systems

Margaret Cook[+,1], Diego Alba Burbano[+,1,2], Ryan A.L. Cardiff[1], Herbert M. Sauro[1,3], James M. Carothers[*,1,2]

1: Molecular Engineering & Sciences Institute
and Center for Synthetic Biology
University of Washington
Seattle, WA 98195
United States

2: Department of Chemical Engineering
University of Washington
Seattle, WA 98195
United States

3: Department of Bioengineering
University of Washington
Seattle, WA 98195
United States

+: These authors contributed equally

*: Corresponding author
jcaroth@uw.edu
206-221-4902

Abstract

Cell-free systems offer a powerful platform for rapid metabolic prototyping and biosynthesis, but their inherent complexity requires predictive models to guide rational design. We developed an automated computational pipeline that integrates proteomic and metabolomic data to construct, parameterize, and simulate kinetic models of cell-free lysates. Enzyme-catalyzed reactions are mapped using KEGG annotations, kinetic parameters are sourced from BRENDA or estimated using machine learning tools, and thermodynamic constraints are incorporated via eQuilibrator. Model parameters are refined by fitting the model to time-course metabolomics data using a global optimization algorithm. We apply this workflow to model an *E. coli*-based cell-free lysate, generating a kinetic model comprising 33 enzymes, 53 reactions, 63 metabolites, and 438 parameters. The model was trained on 338 time-resolved metabolite measurements and validated on 91, accurately capturing metabolic behavior. Simulations revealed key flux distributions, cofactor dependencies, and metabolite sinks, and enabled exploration of how system-level responses emerge from perturbations such as lysate dilution, enzyme inhibition, and heterologous enzyme expression. This framework supports the scalable construction of predictive kinetic models across diverse lysates and engineered pathways and offers a valuable tool for designing and optimizing cell-free bioproduction systems.

Introduction

Synthetic biologists seek to understand and harness biological systems to engineer solutions in a wide variety of fields, ranging from medicine and agriculture to chemical bioproduction. However, optimizing living cells for bioproduction remains challenging due to multiple factors, such as the competition between the engineered biosynthesis pathway and cell viability, as well as the dense interconnectivity of metabolic networks, which hinders rational design (1–4). Importantly, crude extract-based cell-free systems retain much of the native metabolic machinery required for energy and cofactor regeneration essential for sustained biosynthesis. However, cell-free lysates also contain other enzymes that may interfere with the engineered pathway, especially if the desired product is not orthogonal to the endogenous metabolism (5–8). To address this challenge, computational models are being developed to predict metabolic fluxes, identify key bottlenecks, and guide the engineering of cell-free biosynthetic systems (9–11).

Computational metabolic models are essential to broaden our knowledge of metabolic networks (12,13). They provide a rational and systematic framework for integrating existing biochemical knowledge with experimental data, thus enabling understanding of the complex operation of metabolism. A widely used approach in metabolic modeling is constraint-based modeling, particularly Flux Balance Analysis (FBA). FBA utilizes stoichiometric constraints and optimization principles to estimate steady-state flux distributions in a metabolic system by applying constraints such as nutrient availability and growth (14,15). While constraint-based models perform well for *in vivo* models that achieve steady-state conditions, they have key shortcomings when applying such tools to *in vitro* systems which lack an obvious objective function (ie.

76 growth). Dynamic models that incorporate enzyme kinetics are better suited to represent cell-free systems (9,12). However, kinetic models require identifying appropriate rate laws
78 for each reaction and to either experimentally measure or extract kinetic parameters from literature, which makes them very time consuming to generate. Hence, current cell-free
80 kinetic models include only enzymes directly involved in the pathway of interest and sometimes those immediately surrounding it. However, there is strong evidence that the
82 endogenous metabolism of crude lysates has a significant impact on the flux of engineered pathways, especially if the targets are central metabolites (5–8).

84 To this point, we have designed a framework to construct, parameterize, fit and analyze kinetic models of cell-free lysates with the aim of examining the effect of the
86 lysate's endogenous metabolism on engineered pathways. We developed FRENDA (Fast Retrieval of ENzyme DAta), a computational workflow that integrates cell-free proteomic
88 and metabolomic data with information from curated biochemical databases into a human-readable spreadsheet. FRENDA systematically identifies enzymatic reactions
90 based on proteomic identifiers, retrieves available kinetic parameters from sources such as BRENDA (16), and estimates missing values using machine learning models trained
92 on enzyme sequence and substrate structure features (17,18). The resulting output is a structured, human-readable dataset detailing the metabolic network, including reactions,
94 metabolites, and kinetic parameters. This dataset can be automatically converted into an executable model using a companion tool that generates Antimony code, which is readily
96 translatable into SBML format for dynamic simulation and systems-level analysis (19–21). Since fitting large kinetic models to experimental data is computationally challenging, in
98 this work we also implemented a parallelizable optimization framework based on

population-based algorithms using the Python library pygmo in combination with libRoadRunner for fast model simulation (22–24). This framework, packaged in a Docker container for scalability and portability, enables efficient parameter estimation by distributing computational tasks across multiple processors or computing nodes. We applied this framework to generate multiple kinetic models of an *E. coli* cell-free and investigate how interventions—such as lysate dilution and small-molecule inhibition—affect flux distribution across the metabolic network. These case studies demonstrated that even simple perturbations can induce complex, system-wide effects, highlighting the importance of detailed kinetic models for predicting metabolic behavior. Together, this integrated workflow for assembling, fitting and analyzing large kinetic models will provide a foundation for accelerating the design and optimization of metabolic systems, with broad applications in synthetic biology, cell-free biomanufacturing, and high-throughput pathway prototyping.

Results

Developing workflows for assembling and fitting kinetic models

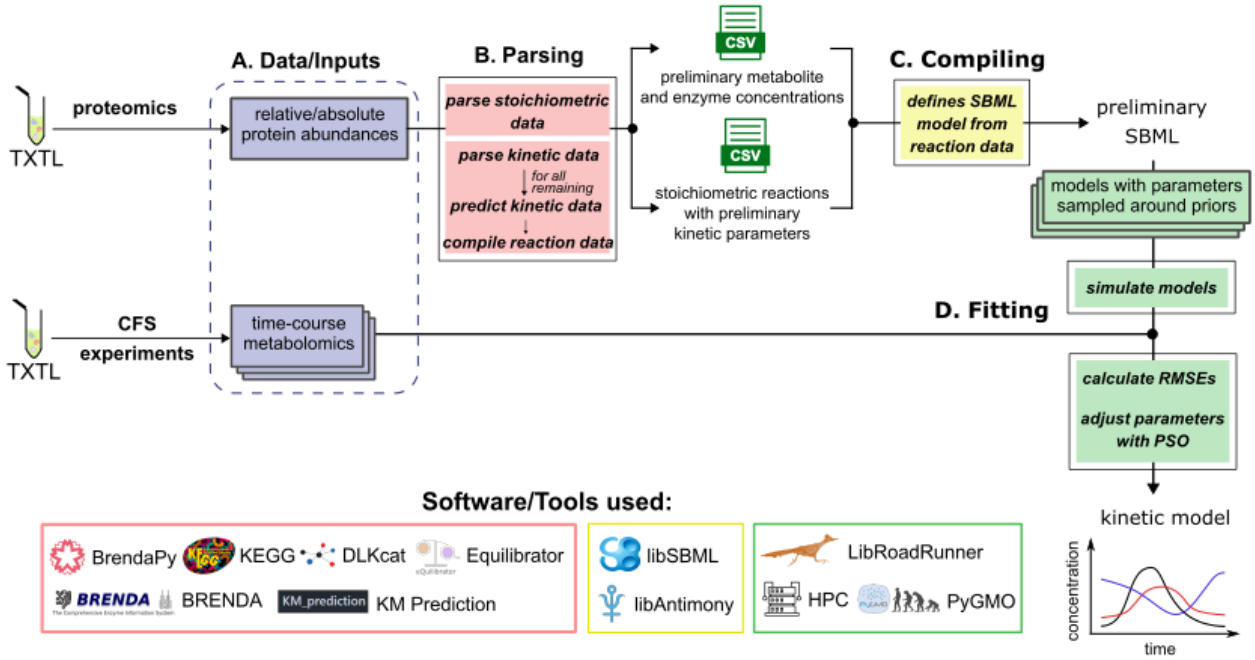
FRENDA for constructing and parameterizing cell-free models

To streamline the process of constructing and parametrizing kinetic models of metabolic systems, we developed FRENDA (Fast Retrieval of ENzyme DAta) (Figure 1). FRENDA takes proteomics and metabolomics data and parses several databases to create a list of all the enzymatic reactions in the metabolic network. First, the proteomics IDs are automatically converted into KEGG IDs, which are used to retrieve relevant information for each enzyme from the KEGG database (25), such as the set of KEGG reaction IDs catalyzed by the enzyme and the corresponding substrate and product KEGG IDs. FRENDA then queries the BRENDA database using each enzyme ID to retrieve specific kinetic parameters for the enzyme and its corresponding species, which are then compiled with other reaction details. If kinetic parameters are unavailable for a given enzyme-substrate combination, two machine learning algorithms—DLKcat and KM Prediction (17,18)—are used to estimate the necessary parameters based on the enzyme's protein sequence, fetched from UniProt (26), and the substrate's SMILES string (27), fetched from PubChem (28). Thermodynamic equilibrium constants for each reaction are then calculated using the open-source tool eQuilibrator (29). Finally, metabolite concentrations are estimated using metabolomics data, values from previous studies, or the average of other metabolites, in that order of priority.

The output of FRENDA is a human readable spreadsheet that tabulates all enzymes and metabolites present in the system, all the enzymatic reactions, and their associated kinetic parameters. At this point, the user can specify which rate law best

describes each reaction, as well as any inhibition that may affect a reaction. By default, FRENDA employs the common modular rate law as it covers all possible stoichiometries and types of regulation and is thermodynamically consistent (30). To convert the FRENDA spreadsheet into an executable model, we developed F2A (FRENDA-to-Antimony) which parses the reaction and metabolite information and converts it into the text-based model definition language Antimony. This Antimony file contains all variable definitions and ordinary differential equations describing the system and is still human readable, which facilitates troubleshooting. More importantly, Antimony can readily be converted into the standard modeling format SBML for executing and analyzing the model.

Figure 1: Computational workflow for assembling and fitting kinetic models



Schematic of workflow for assembling and fitting kinetic models. A) Proteomics and metabolomics data for lysate are collected and tabulated. B) FRENDA uses proteomics to parse for reaction information and kinetic parameters and generate human-readable spreadsheets with all the model information. C) F2A compiles the spreadsheets into

Antimony code and makes the executable SBML model. D) SBML model is fitted to the metabolomics data using high-performance computing.

We used FRENDA to assemble a kinetic cell-free model based on the proteomic analysis of cell-free lysate prepared from *E. coli* BL21 Rosetta2 grown aerobically with glucose as a carbon source (31,32). To explore how many reactions are really needed to describe the system, we adopted a modular model-building approach, utilizing KEGG modules and manual curation to incrementally incorporate a larger subset of biosynthetic pathways present in the lysate. These models contained between 29 and 104 reactions, and 169 and 508 parameters (Figure 2C).

We evaluated these different cell-free models on experimental data collected in a previous study (7). The dataset consists of concentrations of key metabolites, such as pyruvate, malate, succinate, and glyoxylate, over time, under several conditions, including adding different metabolites, diluting the lysate, and adding an enzyme inhibitor (Figure 2A). In total, this dataset contained 18 metabolites in 33 samples across 7 experiments, summing to 594 datapoints (Figure 2B). We assessed model performance with the weighted root mean square error (RMSE). Unsurprisingly, these models were not able to capture the cell-free experimental data (Figure SI), most likely due to using kinetic parameters measured in a cellular context instead of cell-free. Therefore, we implemented a global optimization scheme to fit the kinetic models to the experimental data.

Model fitting to metabolomics data with parallel particle–swarm optimization

Fitting large kinetic models to experimental data is challenging due to the non-linearity of the systems of ODEs, resulting in a non-convex optimization problem. One

class of algorithms designed to solve non–convex, non–linear optimization problems is
176 population–based algorithms, which explore the parameter space by iteratively improving a pool of candidate solutions. This class includes methods such as genetic algorithms,
178 differential evolution, particle–swarm optimization (PSO), and many others, all of which have been shown to have good performance when fitting different kinetic models (24,33–
180 36). An advantage of population-based algorithms is that they can efficiently be parallelized to achieve linear speedups by simply distributing the population update step
182 across different computing nodes. To efficiently implement optimization algorithms in parallel and parametrize our CFS models, we combined pygmo (24), a Python library that
184 implements optimization algorithms in parallel using ipyparallel for task distribution, with libRoadRunner (22,23), a library for compiling SBML models directly into machine code
186 for faster execution (Figure 2B). We packaged these libraries into a Docker image for easier deployment, creating a flexible framework for running many model simulations
188 across multiple processors or even multiple machines. In this work, model simulations were performed on the University of Washington HYAK high-performance computing
190 cluster, across 10 Intel Xeon Gold 6230 CPU @ 2.10GHz with 40 logical processors each.

192 We fit the CFS models to previously published experimental data supplemented with additional key metabolite measurements, totaling 18 metabolite concentrations
194 (malate, pyruvate, serine, glycine, NADH, CoA, acetyl-CoA, succinate, fumarate, citrate, ATP, glyoxylate, alanine, aspartate, lysine, valine, oxobutanoate, and hydroxyglutarate)
196 across 26 conditions with varying concentrations of added metabolites, enzyme inhibitors, and dilution factors (7). For some conditions we collected time course data. When

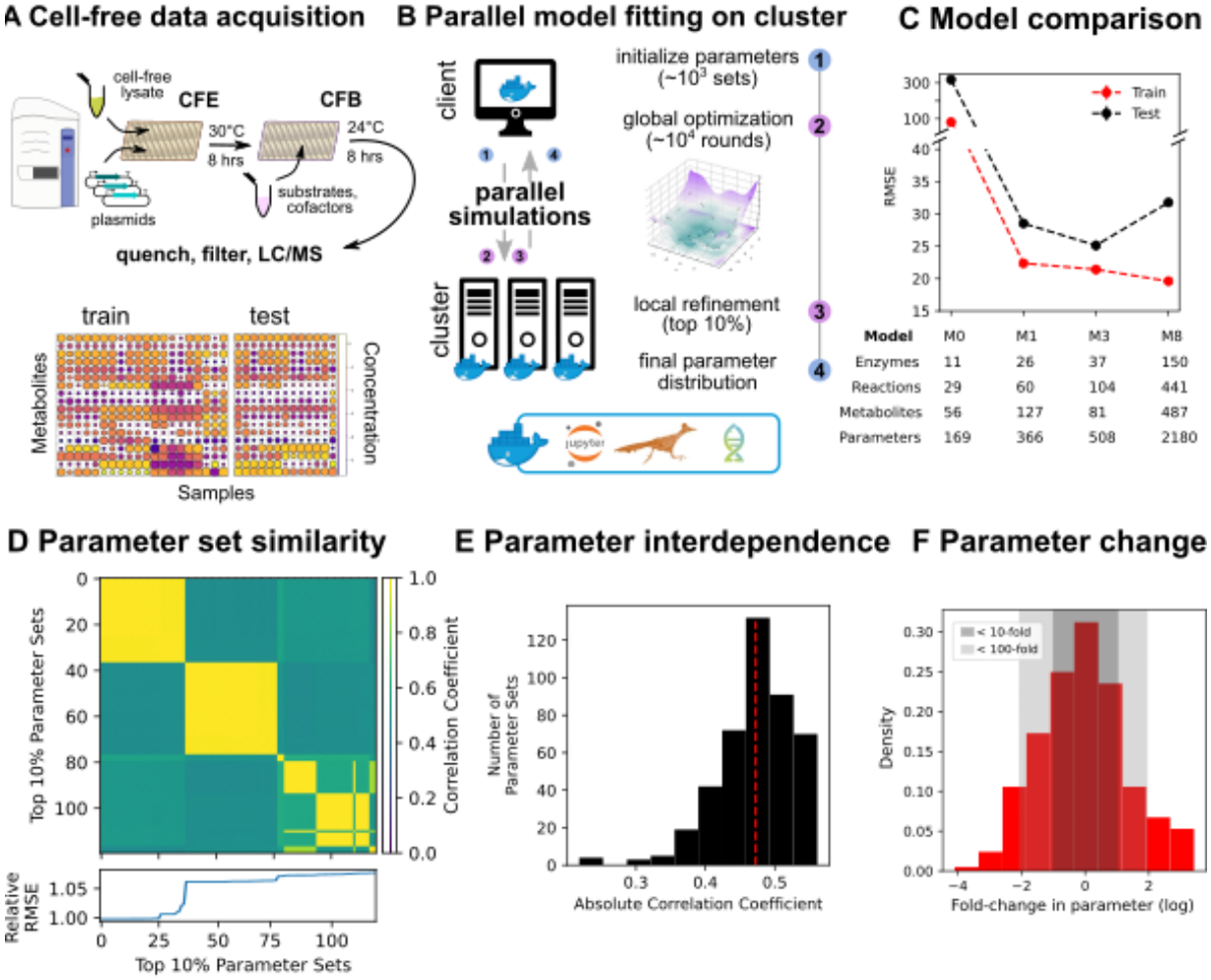
measured LC/MS peaks were below the standard curve, we assumed the metabolite was present in low abundance and assigned a value equal to 1/10 of the limit of detection (LOD) for that metabolite. If a metabolite was not detected in a sample by LC/MS, we considered it to be in very low abundance and filled in the value with 1/100 of the LOD for that metabolite. To account for measurement variability and confidence in missing values, we applied a weighting scheme where each data point was weighted inversely to its relative standard deviation. Data weights were normalized between 1 and 0.3 such that points with lower variance were assigned higher weights, while data points with higher variance were down-weighted. This weighting ensured that more precise measurements had a greater influence on the model's parameter estimates, improving overall model robustness. In total, our dataset contained 594 measurements. To evaluate model performance, we divided the dataset into training and testing subsets. This resulted in a 75-25 train-test condition split. All metabolite measurements and weights are given in Supplementary Tables 1 and 2.

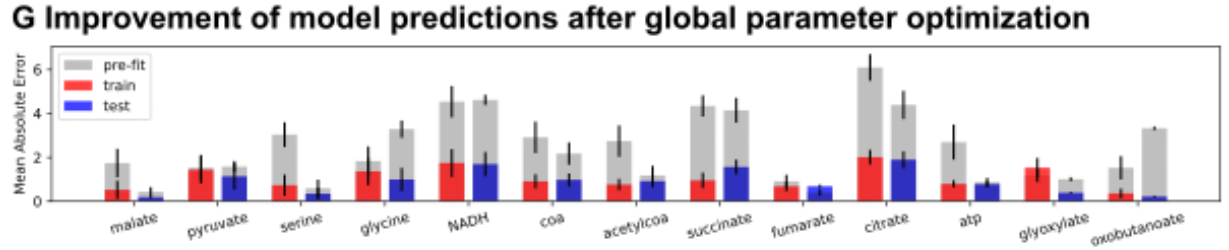
To fit the models, we minimized we minimized the weighted root mean square error (RMSE) for all metabolites across all conditions. When time course data was available, we also minimized the RMSE of the change in concentration over time. To minimize RMSE, we used the canonical PSO algorithm with local velocity update, a high constriction factor of 0.95 and low social and cognitive components of 0.25, and a maximum relative velocity of 0.1. These parameters enhance local exploitation and have been shown to be advantageous for high dimensional problems. In all cases, we used a population size equal to three times the total number of CPUs, and we run each fit with

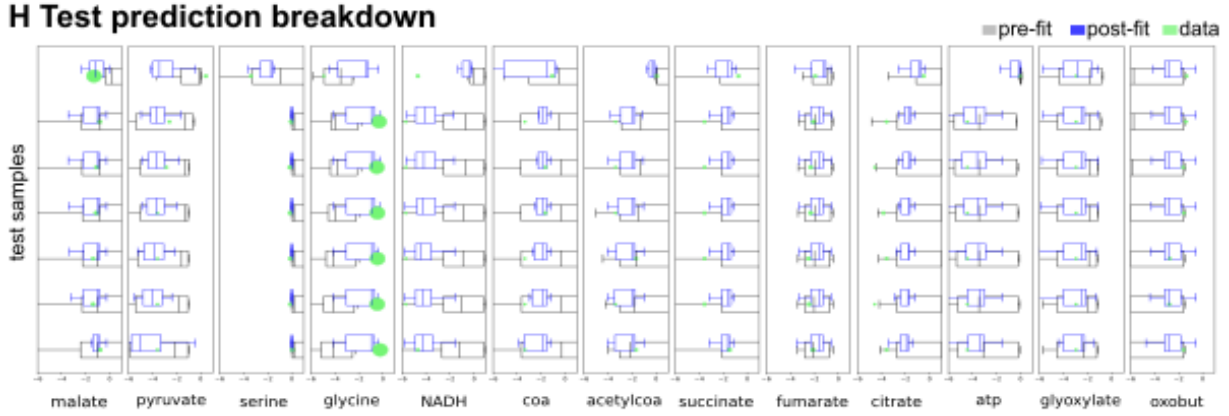
three independent initial populations. Initial parameters were sampled from uniform distributions around the parameter values compiled with FRENDA.

With this method, we compared 4 different CFS models with an increasing number of reactions. For consistency, each model was trained for 10,000 iterations, by which time RMSE was barely changing (Figure SI) . We found that adding complexity through extra parameters the model improved both the train and test RMSE, until a certain point. The largest model, containing over 400 reactions and 2000 parameters, resulted in a smaller train RMSE than the smaller models but a larger test RMSE, indicating that the model was probably overfitting to the train data. Additionally, we compared the models using AIC (Akaike Information Criterion), which balances model fit and complexity by penalizing the number of parameters. We found the third largest model, M3, to have the lowest AIC value (Figure SI), indicating it had the best trade-off between goodness of fit and model complexity.

Figure 2: Fitting large cell-free kinetic models to metabolomics data







A) We used the Echo to assemble cell-free reactions with varying interventions, such as spiked metabolites and enzyme inhibitors. To sample at different timepoints, we had replicate reactions we quenched throughout the length of the reaction. Cell-free reactions were filtered and run through LC/MS for quantification. B) We used parallel computing to distribute the model and data across compute nodes for faster optimization convergence. C) Train and test RMSE score for different models of varying complexity. D) Top: Correlation matrix of the best 10% parameter sets found through global optimization. Bottom: relative RMSE for the best 10% parameter sets. E) Distribution of correlation coefficients across parameters for the top 10% best parameter sets. Red dotted line indicates the average correlation. F) Distribution of fold-change in model parameters in comparison to the initial parameters parsed with FRENDA. Shaded are indicates parameters that have changed less than 10-fold and 100-fold. G) Average metabolite absolute error across train and test datasets, before and after fitting the model to the train test. H) Model metabolite predictions of the test set, before and after fitting the model to the train test.

Analyzing parameter space and model fits

We proceeded to analyze M3 as it was the best-performing model out of the ones tested. Since PSO can converge to multiple solutions, and our system is very underdetermined, we sought to understand the probable parameter space. We first took the top 10% of parameters with the smallest RMSE and calculated the correlation among them (Figure 2D). We found 6 parameters sets with >0.95 correlation, and at least 10 solutions per set. On average, the correlation across these 6 parameter sets was 0.62, suggesting the PSO algorithm was able to find several different solutions within an acceptable RMSE range, or 5% of the best parameter set. We also found an average correlation of 0.47 across parameters (Figure 2E), indicating there is a high degree of parameter interdependence. For this reason, moving forward we decided to model the probable parameter space with a multivariate distribution instead of averaging across the

best parameter set from each replicate. We compared the fit parameters to the original values parsed with FRENDA and found that on average 56% were within 10-fold of the literature or predicted values, and 32% were within 100-fold (Figure 2F), highlighting the need for fitting algorithms.

Unsurprisingly, we found that across the board the model was better at recapitulating the metabolomic data after fitting to the train set. With the exception of pyruvate and glyoxylate, the trained model achieved a lower absolute error on all other metabolites compared to the model using the parameters parsed with FRENDA (Figure 2G). The higher error for pyruvate and glyoxylate may be attributed to their high connectivity within the metabolic network, which introduces greater uncertainty due to missing reactions or regulatory interactions not captured in the current model. Additionally, inconsistencies or noise in the cell-free metabolomics data may contribute to these deviations. To better understand these discrepancies, we investigated whether model error was correlated with the number of reactions a metabolite participates in—used here as a proxy for its network connectivity—or with the magnitude of concentration variation across experimental conditions. Our analysis revealed a weak but noticeable trend: metabolites involved in a larger number of reactions tended to have higher prediction errors (Pearson correlation coefficient of 0.2), suggesting that model performance may be limited by incomplete representation of network connectivity or regulatory complexity. Similarly, metabolites with greater variance across samples also exhibited higher prediction errors, although with a weaker correlation (Pearson correlation coefficient of 0.18). Nevertheless, the test errors were on par with train errors (Figure 2G), and the predictions generally followed the same trends as the data (Figure 2H), which

gave us confidence that the model was able to capture a significant part of the metabolic network activity.

Analyzing Flux Distributions of the Endogenous Lysate

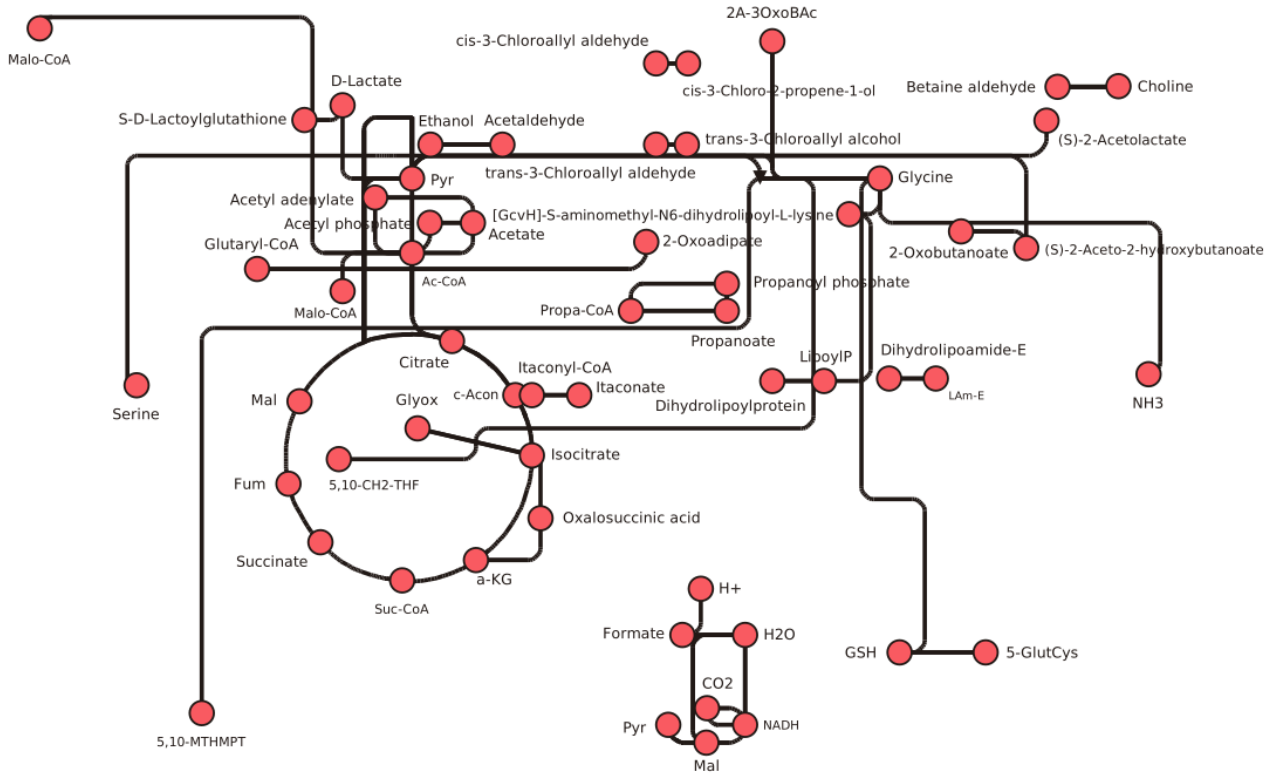
Modeling endogenous reactions in lysate

With a kinetic model that sufficiently captures the experimental responses to both chemical and physical interventions, we next aimed to leverage it to interrogate the system's metabolic behavior. Specifically, we used the model to analyze flux through individual reactions and cumulative flux across biosynthetic modules, identify potential metabolic sinks, and evaluate the system-level effects of perturbations such as lysate dilution and small molecule inhibition.

To evaluate the model holistically, we first visualized the full metabolic network using SBMLNetwork (37), a Python package designed for rendering SBML-formatted models. To maintain consistency with community standards, we adopted coordinate layouts from KEGG's metabolic pathway maps, allowing us to represent all metabolites, reactions, and their associated fluxes in a biologically meaningful and interpretable format (Figure 3). These fluxes reflect system behavior one hour after the onset of biosynthesis. During the first half hour, glycolysis and the TCA cycle exhibit the highest fluxes (Figure 4A), consistent with expectations given that the enzymes involved in these pathways are among the most abundantly expressed in log-phase *E. coli*. This observation underscores the importance of accounting for endogenous metabolism when modeling biosynthetic pathways in cell-free lysates, particularly for pathways that intersect with central carbon metabolism or rely on cofactors such as NADH and ATP. As biosynthesis progresses

beyond four hours, flux through all biosynthetic modules gradually declines. Interestingly, after 2 hours TCA is no longer accounting for ~80-100% of flux, but only 50% while the other half is made up of amino acid metabolism and redox/energy metabolism. However, flux through the TCA and redox/energy metabolism persists the longest, ultimately leading to an accumulation of oxoglutarate and oxaloacetate. These accumulations suggest the presence of metabolic sinks, potentially resulting from imbalanced pathway kinetics, or depletion of essential cofactors. Further analysis of these sinks may reveal opportunities for pathway optimization or targeted interventions to improve yield and pathway robustness.

Figure 3: Metabolic network of the cell-free model



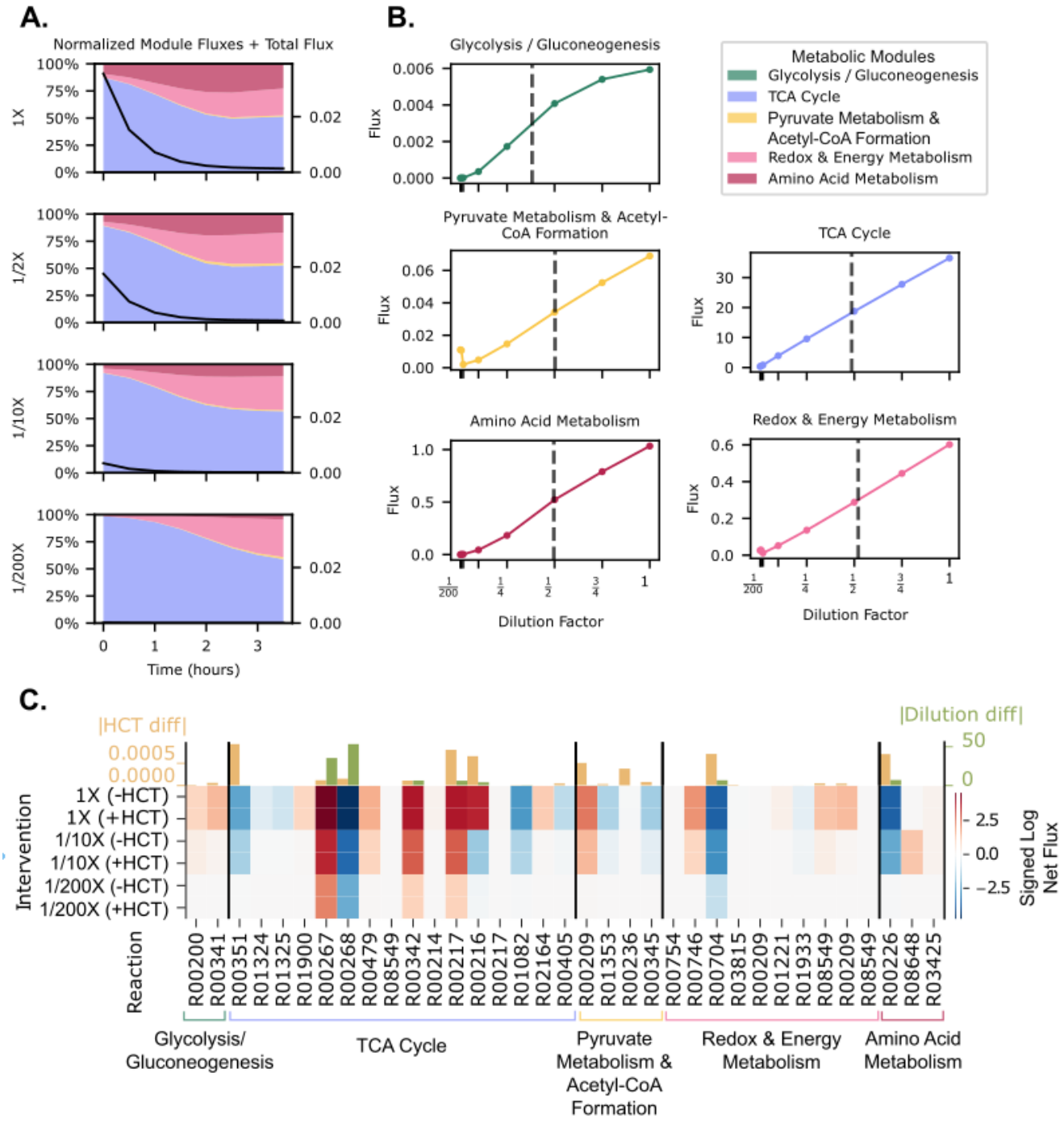
SBMLNetwork representation of the metabolic network of M3 showing all reactions and the most important metabolites. Cofactors have been hidden for simplicity.

One of the key advantages of cell-free lysate systems is the expanded set of control mechanisms available beyond the traditional gene expression regulation used in cell-based bioproduction (4,38). For instance, diluting the lysate to reduce the impact of endogenous metabolism has been shown to improve product yield (7,8). Additionally, small molecule inhibitors can be used to selectively block competing reactions that would otherwise divert pathway intermediates (7). To investigate the effects these interventions actually have on the system, we simulated model M3 with these changes.

To investigate the effects of lysate dilution on pathway dynamics, we simulated progressive dilution conditions using the kinetic model by scaling the concentrations of all endogenous enzymes and metabolites proportionally. This approach mimics experimental protocols where lysate is diluted with buffer to reduce background metabolic activity. Model simulations revealed that dilution led to a decrease in flux through central metabolic pathways such as glycolysis and the TCA cycle, thereby reducing competition for shared cofactors like NADH and ATP (Figure 4A). Interestingly, the model also predicted a threshold beyond which further dilution resulted in diminished overall productivity, likely due to insufficient enzymatic capacity to sustain key reactions. These results highlight a trade-off in lysate dilution strategies and demonstrate how the model can be used to identify optimal dilution levels that balance reduced background metabolism with sufficient catalytic activity. Overall, all modules showed a linear response in response to lysate dilution, with glycolysis being the most resistant to dilution (Figure 4B). Surprisingly, there was a small increase in flux at lower dilution for the reactions in pyruvate metabolism and redox/energy metabolism, which could be due to these

reactions now outcompeting other reactions for their substrate. This insight could be particularly valuable when implementing a biosynthetic pathway, as it enables more precise tuning of the lysate environment to support desired fluxes without unnecessary diversion of resources or loss of catalytic efficiency.

Figure 4: Model-predicted changes in flux due to different reaction conditions



A) Total flux and flux per KEGG metabolic module at different dilution factors. B) Analysis of the change in flux of each metabolic module as a function of dilution factor. C) Bottom:

Changes in flux for all reactions in the model for different intervention combinations. Top: Effect of just adding HCT or diluting the reaction.

To investigate the network-level effects of small molecule inhibition, we simulated the addition of hydroxycitrate (HCT) (39), a known inhibitor of citrate synthase, a key enzyme in the TCA cycle. Although HCT specifically targets citrate synthase, the model revealed that its inhibition leads to broader metabolic consequences. Blocking citrate synthase reduces the production of citrate from oxaloacetate and acetyl-CoA, thereby disrupting the TCA cycle and downstream energy production. This causes an accumulation of upstream metabolites, including acetyl-CoA, which may be diverted into alternative pathways, such as fatty acid biosynthesis. Additionally, the inhibition of citrate synthase reduces NADH production, affecting the redox balance and overall ATP synthesis. These changes ripple through central metabolism, altering flux through other interconnected pathways and impacting cofactor availability. Specifically, we observed the fluxes of reactions catalyzed by malate dehydrogenase, malic enzyme, pyruvate dehydrogenase, lactate dehydrogenase, and acetolactate synthase to be significantly impacted with the addition of HCT. This simulation highlights how a targeted perturbation, even when limited to a single enzyme, can induce significant systemic shifts, emphasizing the need for systems-level modeling to understand and predict the broader metabolic consequences of enzyme inhibition in bioproduction systems.

Discussion

We have developed a comprehensive workflow for building, fitting, and analyzing kinetic metabolic models in cell-free systems. Recognizing the complex nature of such models, we focused on streamlining each stage of the process—from data integration to parameter optimization and systems-level analysis. To support model construction, we developed FRENDA, a tool that parses proteomics and metabolomics data alongside public biochemical databases to automatically assemble a list of enzymatic reactions within the network. To efficiently fit model parameters, we implemented a scalable framework for parallel computation. We further addressed the challenge of solution multiplicity by analyzing parameter interdependence and modeling the high-probability parameter space using a multivariate distribution, rather than relying on single best-fit solutions. Overall, this work offers a generalizable workflow for constructing and fitting large kinetic models, with particular relevance to cell-free systems but broadly applicable to any metabolic network.

We applied this workflow to explore how interventions—such as lysate dilution small-molecule inhibition—impact flux distribution of an *E. coli* cell-free system. These case studies showed that even simple perturbations can lead to complex, system-wide effects, reinforcing the need for detailed kinetic models to understand and predict system behavior. Our investigation of hydroxycitrate (HCT) inhibition of citrate synthase revealed how a targeted perturbation can propagate through the system, disrupting the TCA cycle and leading to downstream shifts in cofactor and metabolite fluxes. These insights emphasize the interconnectedness of metabolic networks and illustrate the importance of systems-level approaches in understanding and optimizing bioproduction processes.

One limitation of our current modeling approach is that inhibition is only explicitly considered for HCT, even though many *E. coli* enzymes are known to be regulated by a variety of inhibitors. Programmatically incorporating inhibition information remains a challenge due to the limited availability of quantitative inhibition data across databases. To begin addressing this, we extended FRENDA to parse enzyme-inhibitor relationships from BioCyc and predict inhibition constants (Ki) using the machine learning tool KM Prediction. On average, with this process we identified approximately five potential inhibitors per enzyme. Incorporating this information into the model, using the modular rate law to account for both competitive and non-competitive inhibition, would have significantly increased model complexity—adding roughly 45% more parameters per model. Future work should aim to constrain these inhibition parameters using literature-derived values, and to include enzyme activators, which also play a critical role in metabolic regulation. Additionally, while the modular rate law offers flexibility and thermodynamic consistency, alternative rate laws may provide more accurate representations for some enzymatic reactions.

Despite the advantages of population-based algorithms for fitting complex, non-linear kinetic models, their computational cost remains a significant limitation. These algorithms require evaluating large populations of parameter sets over many generations, which can be particularly demanding when simulating stiff or high-dimensional ODE systems. Although parallelization can significantly reduce runtime, fitting large models still often requires access to high-performance computing resources. This becomes particularly crucial when increasing the amount of training data, as training time increases linearly with respect to the number of samples. To further improve efficiency, alternative

approaches such as adjoint sensitivity analysis or backpropagation through differential equation solvers could be employed (40–42). These gradient-based methods can compute sensitivities with respect to all parameters in a single backward pass, offering significant speedups for models with many parameters and enabling more scalable and data-efficient optimization. Additionally, model reduction techniques—such as lumping of fast reactions, quasi-steady-state approximations, or pathway compression (43,44)—can significantly reduce the dimensionality and stiffness of the system, lowering computational demands without sacrificing key dynamic behaviors.

Our workflow can readily be used for prospecting different biosynthetic pathways, as FRENDA can parse for reactions and kinetic parameters given just the pathway enzyme names or EC numbers. This enables rapid *in silico* evaluation of candidate pathways by simulating how the introduction of a synthetic pathway may interact with the endogenous lysate metabolism, identifying potential bottlenecks, and predicting flux distributions under different conditions. Coupling this capability with design of experiments strategies or with machine learning-based tools such as the Automated Recommendation Tool (ART) (45), would allow researchers to systematically explore combinatorial design spaces, prioritizing experimental conditions that are most likely to improve biosynthetic pathway performance. This data- and model-driven approach holds the potential to transform metabolic engineering, enabling more rapid and cost-effective development of cell-free biosynthesis platforms.

As cell-free system engineering continues to advance, there is growing interest in generating lysates from a variety of organisms beyond *E. coli*, including non-model bacteria with distinct metabolisms. These alternative lysates can offer unique biochemical

capabilities—such as native tolerance to inhibitory compounds, unusual cofactor usage, or novel biosynthetic routes—that expand the scope of possible bioproduction strategies. However, the diversity and limited characterization of these systems present a significant challenge for rational design. The work by Moore et al. demonstrated that combining high-throughput experimentation with model-based inference enables the rapid characterization of transcription–translation dynamics in nonmodel lysates (46). Our framework builds on this concept by enabling fast, automated construction of organism-specific kinetic models from basic proteomic information, making it a powerful tool for exploring and engineering cell-free systems derived from diverse microbial sources. As the field moves toward bespoke lysates tailored for specific products or pathways, such modeling approaches will be crucial for predicting system behavior, identifying design constraints, and accelerating the development of next-generation biomanufacturing platforms.

Overall, this work highlights the power of integrating automated kinetic modeling with experimental data to systematically explore and optimize complex metabolic networks. By streamlining the reconstruction and parameterization of kinetic models, our framework enables predictive simulations that can guide targeted interventions and rational design strategies. As cell-free platforms continue to diversify and expand in application, such computational tools will play an increasingly central role in navigating the design space, reducing experimental burden, and accelerating the development of efficient, scalable bioproduction processes.

Acknowledgments

We thank members of the Carothers and Sauro groups for advice, materials, and comments on the manuscript.

Author Contributions

M.K., D.A.B., and J.M.C. designed the research, M.K., D.A.B., and R.A.LC. performed experiments, M.K. and D.A.B. developed computational pipelines, M.K., D.A.B., and R.A.LC. analyzed the data. M.K., D.A.B and J.M.C. wrote the manuscript with input from all of the authors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding

The information, data, or work presented herein was funded in part by the Advanced Research Project Agency-Energy (ARPA-E), U.S. Department of Energy under Award Number DE-AR0001514 and by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences Energy Frontier Research Centers program under Award Number DE-SC0023091. This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or

usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of the authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Data availability

All data and code used in this study is given in https://github.com/

Methods and Materials

Cell-free protein synthesis reactions

The cell-free system was acquired from Arbor Biosciences (myTXTL). The cell-free system used for an experiment was thawed on ice and pooled into a 1.5 ml Eppendorf tube, vortexed, and spun-down using a mini benchtop centrifuge to ensure homogeneity across samples. For reactions containing three or fewer genes, reactions were assembled on ice from the CFE, purified DNA, and necessary cofactors. The CFE was pipette mixed and added to each PCR tube in 7.5 µL for a final volume of 10 µL. These PCR tubes were incubated overnight at 30C. For reactions involving more than three genes, plasmids and cofactors were mixed with an acoustic liquid handler robot (Echo Labcyte 525) into Labcyte 384-well destination plates (001–14555). The 384-well plates were then incubated at 30 °C overnight.

Cell-free bioproduction reactions

Cell-free bioproduction reactions were mixed in 25 µL containing 2.5 µL of CFE-expressed enzymes. CFE-expressed enzymes were diluted in 10 mM Tris pH 8 prior to adding if they were diluted beyond 10-fold in the final reaction. Bioproduction reactions were done in 50 mM HEPES pH 8. For reactions containing three or fewer enzymes, reactions were assembled by hand from the CFE-expressed enzymes and necessary substrates and cofactors. Detailed information about sample preparation can be found in Methods S3. For reactions involving more than three enzymes, enzymes and chemicals were mixed with an acoustic liquid handler robot (Echo Labcyte 525) into 96-well V-bottom plates (Costar, Cat. 3363). The plates were sealed with a foil adhesive (Thermo,

AB0626) and the reactions were run for 4–8 h at room temperature. In general, reactions containing three or fewer enzymes were run for 4 h, while larger reactions were run for 8 h. This is due largely to the slow reaction progress of the pyc enzyme, which led to us allowing longer reactions when using pyc as an intermediate step. At the end of the reaction, samples were quenched with 2 vol of acetonitrile to denature proteins. The entire sample volume was then filtered using 96-well 0.2 µm Supor membrane plates (Cytiva: 8019) and spun for 15 min at max speed. Samples were then either analyzed via LC/MS or stored at −20C until ready for analysis.

Metabolite quantification with LC/MS

Samples were analyzed via Agilent 6530 LC/Q-TOF in negative mode using a BEH Amide 50 mm column (Waters, 186004800). Standard curves were prepared by spiking known amounts of metabolites into diluted CFE and HEPES. For LC/MS, the aqueous phase was LC/MS grade water and the organic phase was 95/5 acetonitrile/water with 10 mM ammonium acetate and .04% v/v ammonium hydroxide. The % aqueous/organic gradient was run as follows: hold at 5/95 for 2.5 minutes, move to 33.5/66.5 over 5 minutes, 40/60 over 1 minute, hold 40/60 for 1 minute, then return to 5/95 over 1 minute. The flow rate was held at 0.5 mL/min.

References

1. Lim HJ, Kim DM. Cell-Free Metabolic Engineering: Recent Developments and Future Prospects. Methods Protoc [Internet]. 2019 Apr 30 [cited 2021 Feb 4];2(2). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6632161/

2. Dudley QM, Karim AS, Jewett MC. Cell-Free Metabolic Engineering: Biomanufacturing beyond the cell. Biotechnol J. 2015 Jan;10(1):69–82.

3. Rasor BJ, Vögeli B, Landwehr GM, Bogart JW, Karim AS, Jewett MC. Toward sustainable, cell-free biomanufacturing. Curr Opin Biotechnol. 2021 Jun;69:136–44.

4. Claassens NJ, Burgener S, Vögeli B, Erb TJ, Bar-Even A. A critical comparison of cellular and cell-free bioproduction systems. Curr Opin Biotechnol. 2019 Dec 1;60:221–9.

5. Wagner L, Jules M, Borkowski O. What remains from living cells in bacterial lysate-based cell-free systems. Comput Struct Biotechnol J. 2023;21:3173–82.

6. Dinglasan JLN, Doktycz MJ. Rewiring cell-free metabolic flux in *E. coli* lysates using a block—push—pull approach. Synth Biol. 2023 May 10;8(1):ysad007.

7. Cardiff RAL, Chowdhury S, Sugianto W, Tickman BI, Burbano DA, Meyer PA, et al. Carbon-conserving bioproduction of malate in an E. coli-based cell-free system. Metab Eng. 2025 Sep;91:59–76.

8. Chowdhury S, Westenberg R, Wennerholm K, Cardiff RAL, Beliaev AS, Noireaux V, et al. Carbon Negative Synthesis of Amino Acids Using a Cell-Free-Based Biocatalyst. ACS Synth Biol. 2024 Dec 20;13(12):3961–75.

9. Martin JP, Rasor BJ, DeBonis J, Karim AS, Jewett MC, Tyo KEJ, et al. A dynamic kinetic model captures cell-free metabolism for improved butanol production. Metab Eng. 2023 Mar;76:133–45.

10. Huber N, Alcalá-Orozco EA, Rexer T, Reichl U, Klamt S. Model-based optimization of cell-free enzyme cascades exemplified for the production of GDP-fucose. Metab Eng. 2024 Jan;81:10–25.

11. Van Sluijs B, Zhou T, Helwig B, Baltussen MG, Nelissen FHT, Heus HA, et al. Iterative design of training data to control intricate enzymatic reaction networks. Nat Commun. 2024 Feb 21;15(1):1602.

12. Yasemi M, Jolicoeur M. Modelling Cell Metabolism: A Review on Constraint-Based Steady-State and Kinetic Approaches. Processes. 2021 Feb 9;9(2):322.

13. Fang X, Lloyd CJ, Palsson BO. Reconstructing organisms in silico: genome-scale models and their emerging applications. Nat Rev Microbiol. 2020 Dec;18(12):731–43.

14. Antoniewicz MR. A guide to metabolic flux analysis in metabolic engineering: Methods, tools and applications. Metab Eng. 2021 Jan;63:2–12.

15. Orth JD, Thiele I, Palsson BØ. What is flux balance analysis? Nat Biotechnol. 2010 Mar;28(3):245–8.

16. Chang A, Jeske L, Ulbrich S, Hofmann J, Koblitz J, Schomburg I, et al. BRENDA, the ELIXIR core data resource in 2021: new developments and updates. Nucleic Acids Res. 2021 Jan 8;49(D1):D498–508.

17. Li F, Yuan L, Lu H, Li G, Chen Y, Engqvist MKM, et al. Deep learning-based kcat prediction enables improved enzyme-constrained model reconstruction. Nat Catal. 2022 Jun 16;5(8):662–72.

18. Kroll A, Engqvist MKM, Heckmann D, Lercher MJ. Deep learning allows genome-scale prediction of Michaelis constants from structural features. Locasale JW, editor. PLOS Biol. 2021 Oct 19;19(10):e3001402.

19. Smith LP, Bergmann FT, Chandran D, Sauro HM. Antimony: a modular model definition language. Bioinformatics. 2009 Sep 15;25(18):2452–4.

20. Hucka M, Finney A, Sauro HM, Bolouri H, Doyle JC, Kitano H, et al. The systems biology markup language (SBML): a medium forrepresentation and exchange of biochemical network models. Bioinformatics. 2003 Mar 1;19(4):524–31.

21. Bornstein BJ, Keating SM, Jouraku A, Hucka M. LibSBML: an API Library for SBML. Bioinformatics. 2008 Mar 15;24(6):880–1.

22. Somogyi ET, Bouteiller JM, Glazier JA, König M, Medley JK, Swat MH, et al. libRoadRunner: a high performance SBML simulation and analysis library. Bioinformatics. 2015 Oct 15;31(20):3315–21.

23. Welsh C, Xu J, Smith L, König M, Choi K, Sauro HM. libRoadRunner 2.0: a high performance SBML simulation and analysis library. Martelli PL, editor. Bioinformatics. 2023 Jan 1;39(1):btac770.

24. Biscani F, Izzo D. A parallel global multiobjective framework for optimization: pagmo. J Open Source Softw. 2020 Sep 13;5(53):2338.

25. Kanehisa M, Sato Y, Kawashima M, Furumichi M, Tanabe M. KEGG as a reference resource for gene and protein annotation. Nucleic Acids Res. 2016 Jan 4;44(D1):D457–62.

26. The UniProt Consortium, Bateman A, Martin MJ, Orchard S, Magrane M, Adesina A, et al. UniProt: the Universal Protein Knowledgebase in 2025. Nucleic Acids Res. 2025 Jan 6;53(D1):D609–17.

27. Weininger D. SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. J Chem Inf Comput Sci. 1988 Feb 1;28(1):31–6.

28. Kim S, Chen J, Cheng T, Gindulyte A, He J, He S, et al. PubChem 2025 update. Nucleic Acids Res. 2025 Jan 6;53(D1):D1516–25.

29. Beber ME, Gollub MG, Mozaffari D, Shebek KM, Flamholz AI, Milo R, et al. eQuilibrator 3.0: a database solution for thermodynamic constant estimation. Nucleic Acids Res. 2022 Jan 7;50(D1):D603–9.

30. Liebermeister W, Uhlendorf J, Klipp E. Modular rate laws for enzymatic reactions: thermodynamics, elasticities and implementation. Bioinformatics. 2010 Jun 15;26(12):1528–34.

31. Garenne D, Beisel CL, Noireaux V. Characterization of the all- *E. COLI* transcription-translation system myTXTL by mass spectrometry. Rapid Commun Mass Spectrom. 2019 Jun 15;33(11):1036–48.

32. Garamella J, Marshall R, Rustad M, Noireaux V. The All E. coli TX-TL Toolbox 2.0: A Platform for Cell-Free Synthetic Biology. ACS Synth Biol. 2016 Apr 15;5(4):344–55.

33. Nayak J, Swapnarekha H, Naik B, Dhiman G, Vimal S. 25 Years of Particle Swarm Optimization: Flourishing Voyage of Two Decades. Arch Comput Methods Eng. 2023 Apr;30(3):1663–725.

34. Medley JK, Asifullah S, Hellerstein J, Sauro HM. Accelerated Biochemical Kinetic Model Fitting via the Asynchronous, Generalized Island Method. bioRxiv. 2019 Oct 6;660522.

35. Moles CG, Mendes P, Banga JR. Parameter Estimation in Biochemical Pathways: A Comparison of Global Optimization Methods. Genome Res. 2003 Nov;13(11):2467–74.

36. Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings of ICNN'95 - International Conference on Neural Networks [Internet]. Perth, WA, Australia: IEEE; 1995 [cited 2025 Jun 2]. p. 1942–8. Available from: http://ieeexplore.ieee.org/document/488968/

37. Heydarabadipour A, Smith L, Hellerstein JL, Sauro HM. SBMLNetwork: a framework for standards-based visualization of biochemical models [Internet]. 2025 [cited 2025 Jun 2]. Available from: http://biorxiv.org/lookup/doi/10.1101/2025.05.09.653024

38. Karim AS, Heggestad JT, Crowe SA, Jewett MC. Controlling cell-free metabolism through physiochemical perturbations. Metab Eng. 2018 Jan;45:86–94.

39. Schlachter CR, Klapper V, Radford T, Chruszcz M. Comparative studies of *Aspergillus fumigatus* 2-methylcitrate synthase and human citrate synthase. Biol Chem. 2019 Dec 18;400(12):1567–81.

40. Cao Y, Li S, Petzold L. Adjoint sensitivity analysis for differential-algebraic equations: algorithms and software. J Comput Appl Math. 2002 Dec;149(1):171–91.

41. Chen RTQ, Rubanova Y, Bettencourt J, Duvenaud D. Neural Ordinary Differential Equations [Internet]. arXiv; 2018 [cited 2025 Jun 2]. Available from: https://arxiv.org/abs/1806.07366

42. Fröhlich F, Kaltenbacher B, Theis FJ, Hasenauer J. Scalable Parameter Estimation for Genome-Scale Biochemical Reaction Networks. Stelling J, editor. PLOS Comput Biol. 2017 Jan 23;13(1):e1005331.

43. Snowden TJ, Van Der Graaf PH, Tindall MJ. Methods of Model Reduction for Large-Scale Biological Systems: A Survey of Current Methods and Trends. Bull Math Biol. 2017 Jul;79(7):1449–86.

44. Radulescu O, Gorban AN, Zinovyev A, Noel V. Reduction of dynamical biochemical reactions networks in computational biology. Front Genet [Internet]. 2012 [cited 2025 Jun 2];3. Available from: http://journal.frontiersin.org/article/10.3389/fgene.2012.00131/abstract

45. Radivojević T, Costello Z, Workman K, Garcia Martin H. A machine learning Automated Recommendation Tool for synthetic biology. Nat Commun. 2020 Sep 25;11(1):4879.

46. Moore SJ, MacDonald JT, Wienecke S, Ishwarbhai A, Tsipa A, Aw R, et al. Rapid acquisition and model-based analysis of cell-free transcription–translation reactions from nonmodel bacteria. Proc Natl Acad Sci. 2018 May 8;115(19):E4340–9.

662

Conclusions and Future Directions:

This work advances the synthetic biology toolkit by developing modular, programmable CRISPRa/i circuits and complementary computational frameworks for engineering prokaryotic systems. By integrating programmable, multi-layer genetic circuits with data-driven kinetic models, we demonstrate a comprehensive platform for rationally engineering prokaryotic hosts toward complex phenotypes and biosynthetic functions. These systems are foundational to enabling dynamic, multi-gene control and predictive metabolic engineering in industrially and environmentally relevant bacterial chassis.

We established strategies for building and layering CRISPRa/i circuits capable of temporal control, signal integration, and logic-gated gene expression. Using orthogonal promoter-guide RNA pairs and improved activatable promoters, we constructed regulatory motifs to interface with diverse input signals, including light, small molecules, and peptides. This work demonstrates the feasibility of constructing multi-layer genetic circuits that can execute temporally resolved, conditional gene expression programs in both cell-free and cellular contexts. Building upon the foundation of tunable CRISPRa/i regulation, future work will focus on constructing higher-order transcriptional programs that implement dynamic signal processing, memory, and feedback control. Specifically, we aim to design and validate circuit motifs such as incoherent feedforward loops (IFFLs) for pulse generation and noise filtering, multi-stage cascades for temporal gating, and recombinase-based memory switches for long-term state encoding. These motifs are integral to the development of living biosensors and actuators capable of responding to fluctuating environmental cues with robust and programmable outputs.

Moving forward, a critical priority is the extension of these tools beyond *Escherichia coli* to more metabolically and physiologically diverse organisms. Other bacterial hosts such as *Pseudomonas putida,* a solvent-tolerant and metabolically versatile species, *Bacillus subtilis*, a spore-forming bacteria with a long history of industrial use, and purple non-sulfur bacteria (PNSB), which exhibit photoheterotrophic and nitrogen-fixing capabilities, offer unique advantages for carbon conversion and sustainable bioproduction. However, porting genetic circuits across species requires addressing host-specific variation in transcriptional regulation, promoter activity, and CRISPRa/i machinery function. To facilitate cross-species portability, we propose systematic characterization of synthetic promoters and CRISPRa/i components across different bacterial species using high-throughput reporter assays and RNA-seq, and optimizing Cas protein expression and stability. These data will inform the development of promoter libraries with tunable dynamic ranges and predictable input–output behaviors across phylogenetically distinct backgrounds.

On the computational side, we developed automated workflows for assembling kinetic models of cell-free lysate metabolism using quantitative proteomic and metabolomic datasets. By coupling reaction network construction with parameter estimation pipelines and curated kinetic data, we enable the simulation of dynamic metabolic responses to system perturbations. Future directions include the application of these modeling tools to heterologous pathway prospecting, enabling the identification and evaluation of candidate biosynthetic routes based on pathway thermodynamics, enzyme kinetics, and host network compatibility. Importantly, this modeling strategy is inherently extensible to lysates derived from other prokaryotic hosts. By applying the same workflow across

diverse lysate systems, it is possible to construct species-specific kinetic models that reflect native cofactor availability, endogenous enzymatic backgrounds, and energy metabolism. This capability opens the door to comparative evaluation of heterologous biosynthetic pathways across lysates, providing an orthogonal strategy for host selection and metabolic pathway screening in silico, prior to the implementation of pathways in vivo. Such predictive comparisons may reveal host-specific constraints or advantages that would otherwise require extensive strain engineering to uncover. If expanded to model bacteria metabolism, these models complement static flux-based approaches by capturing transient flux distributions and enzyme capacity constraints—critical features for modeling inducible and dynamically regulated pathways. Together, these tools will support a rational, data-driven approach to pathway engineering, strain optimization, and host selection across a broad range of prokaryotic platforms.

Overall, the integration of programmable CRISPRa/i regulation with kinetic metabolic modeling represents a step toward a unified, quantitative design-build framework for prokaryotic systems. By enabling the co-optimization of regulatory logic and metabolic flux, these tools lay the groundwork for the development of robust, responsive, and high-performing microbial platforms for bioproduction and adaptive environmental sensing.

Acknowledgements

The PhD journey has been a long one, but fortunately not a lone one. I am deeply grateful to the many people who have supported, challenged, and inspired me throughout this process. This work was only possible thanks to:

My advisor, James Carothers, who has provided guidance and freedom to pursue difficult and interesting problems.

My committee, Neda Bagheri, Jorge Marchand, and Nitin Baliga, whose tough questions and thoughtful feedback have continually pushed me to refine my work and to consider its broader impact within the scientific community.

My unofficial advisors, Jesse Zalatan and Herbert Sauro, who have consistently offered invaluable insight and detailed, constructive feedback.

My mentors, Ben Tickman, Ian Faulkner, and Ice Kiattisewee, who have literally taught me everything inside and outside the lab, teaching me not just how to do science, but how to think critically, troubleshoot creatively, all while enjoying the mountains, bread, and karaoke.

My closest collaborators, Ryan Cardiff and Maggie Cook, who have been the other half of my brain and incredible partners throughout this journey. You've made Mondays super exciting.

The Wayfinder duo, Jason Fontana and David Sparkman-Yager, for teaching me the importance of asking the right questions and crafting a compelling scientific narrative.

The cell-free gang, Widianti Sugianto and Yejun Kim, for their meticulous attention to detail and relentless pursuit of clarity in the chaos of biochemistry.

The bacterial engineering crew, Ava Karanjia, Kira Olander, Brian Darst, and Tommy Primo, for their creativity and dedication in constantly pushing the boundaries of what our tools can do. Best froworkers ever.

The computational team, Janis Shin and Adel Heydarabadipour, for their patience and willingness to talk math and debug code.

The whole BANANA-SPLIT, Jake Brandner, Stephen Fedak, Yujia Huang, and Quoc Tran, for your perseverance, curiosity, laughter, and the joy of uncovering unexpected biology together.

The PNSB team, Michael Guzman, Amanda Robert, and Jackson Comes, for being my gateway drug to microbiology and an excellent support network in uncharted territory.

Our exceptional undergraduates, Stella Anastasakis, Sarah Grube, Kieran Heiberg, and Semira Beraki, whose enthusiasm and dedication has been refreshing.

Our neighboring lab mates, Hinako Kawabe, Chris Yin, and Cassandra Maranas, for the late tea parties and coffee chats.

And finally, to all my friends and family, thank you for cheering me on, keeping me grounded, and reminding me there's life outside the lab.

762 Diego Alba Burbano