Systems-level modeling for CRISPR-based metabolic engineering

Ryan A. L. Cardiff[a,b,c] *, James M. Carothers[a,c],
Jesse G. Zalatan[a,b], Herbert M. Sauro[a,d] *

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

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

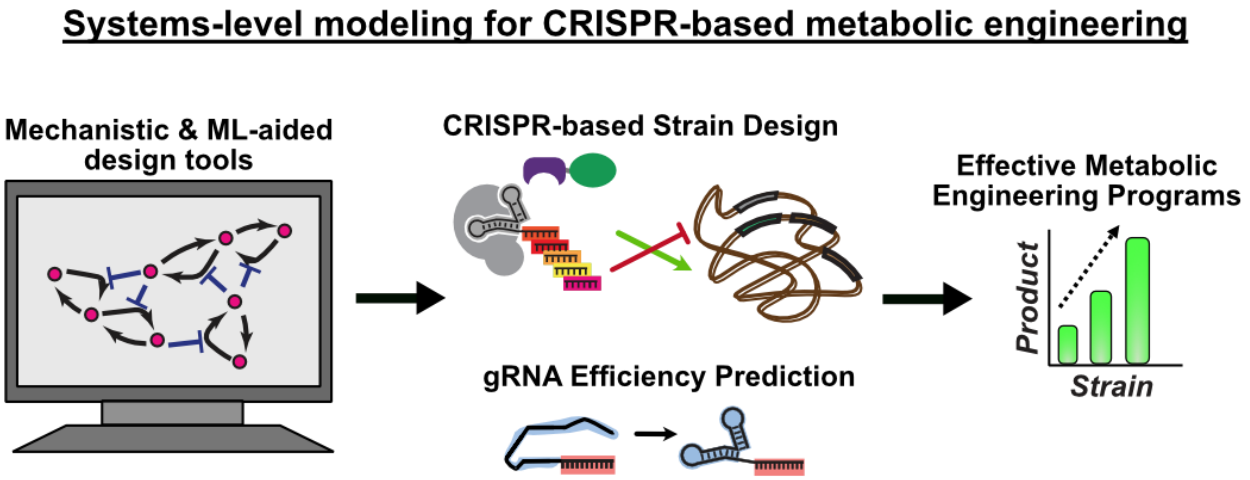
c: Department of Chemical Engineering
University of Washington
Seattle, WA 98195
United States

d: Department of Bioengineering
University of Washington
Seattle, WA 98195
United States

*: Corresponding authors
cardiffr@uw.edu
707-338-3971

hsauro@uw.edu
206-880-8093

Graphical Abstract



Abstract

The CRISPR-Cas system has enabled the development of sophisticated, multi-gene metabolic engineering programs through the use of guide RNA-directed activation or repression of target genes. To optimize biosynthetic pathways in microbial systems, we need improved models to inform design and implementation of transcriptional programs. Recent progress has resulted in new modeling approaches for identifying gene targets and predicting the efficacy of guide RNA targeting. Genome-scale and flux balance models have successfully been applied to identify targets for improving biosynthetic production yields using combinatorial CRISPR-interference (CRISPRi) programs. The advent of new approaches for tunable and dynamic CRISPR activation (CRISPRa) promises to further advance these engineering capabilities. Once appropriate targets are identified, guide RNA prediction models can lead to increased efficacy in gene targeting. Developing improved models and incorporating approaches from machine learning may be able to overcome current limitations and greatly expand the capabilities of CRISPR-Cas9 tools for metabolic engineering.

Keywords: metabolic engineering, CRISPR, genome-scale modeling, gRNA design

I. Introduction

The CRISPR-Cas9 system has enabled sophisticated, multi-gene metabolic engineering programs in a variety of organisms [1,2,3]. Catalytically inactive Cas9 (dCas9) has proven to be a powerful tool for gene regulation due to its ability for programmable RNA-guided DNA binding, allowing it to be used for CRISPR-interference (CRISPRi) or CRISPR-activation (CRISPRa) [4,5]. CRISPRi blocks transcription by recruiting dCas9 or a transcriptional repressor to the promoter or open reading frame (ORF) of a gene of interest, while CRISPRa increases transcription through recruitment of a transcriptional activator to the promoter region. dCas9 can be recruited to a specific DNA sequence through the use of modified guide RNAs (gRNAs) known as single guide RNAs (sgRNAs) or scaffold RNAs (scRNAs) that recognize targets based on Watson-Crick base pairing [6]. scRNAs mediate CRISPRa by acting as sgRNAs with RNA hairpins appended at the 3' end to recruit activator domains to the CRISPR complex [6,7]. Orthogonal gRNAs can be used for multi-gene metabolic engineering platforms capable of targeting an arbitrary set of endogenous and heterologous targets [8].

Metabolic engineering is currently limited by our incomplete understanding of the native gene regulatory networks in the cell and our inability to predictably regulate target gene expression [9]. For effective metabolic engineering programs, models are needed to both identify gene targets and implement effective transcriptional programs. Recently, systems-level modeling has sought to overcome these limitations. The applications of biological modeling for CRISPR-based metabolic engineering typically fall into one of two main categories: (1) using modeling tools to predict favorable gene targets [10–12], and (2) gRNA design to improve on-target efficiency and reduce off-target effects [13–15]. Constraint-based genome-scale modeling is a commonly used approach for suggesting CRISPRi gene knockdown targets predicted to yield improved titers [12,16]. Several recent theoretical and experimental works have shown CRISPRi targeting to be effective for redirecting metabolic flux [1,12,16,17]. Additionally, the development of microbial CRISPRa has enabled activation of targets in the genome that can direct flux towards a pathway of interest [5,18]. Algorithms that can simultaneously recommend targets for up- and down-regulation will greatly advance our control of metabolic flux. Once targets have been

identified, effective gRNAs are needed to implement the desired perturbation. gRNA design tools typically make predictions from biophysical models that determine the free energies of RNA folding and RNA:DNA hybridization, genome-wide screens that rely on empirical data-driven approaches, or some combination of the two methods. In both cases, models have been used to learn rules for sgRNA sequence design to enable highly efficient genome targeting for both metabolic engineering and gene editing applications. In this review, we will discuss the current state and outstanding limitations of CRISPR-based metabolic engineering, as well as how systems level modeling tools will help advance the field.

II. Genome-scale modeling for CRISPR-based metabolic engineering

There is a prevailing question in metabolic engineering as to which genes to perturb and how much to up- or down-regulate their expression to achieve a desired phenotype. Engineers aim to divert flux away from competing pathways and maximize flux through their desired pathway. However, strongly repressing genes involved in essential pathways, such as central carbon metabolism or amino acid biosynthesis, can have deleterious effects on the organism including growth defects, undesired mutations [19,20]. Similarly, overexpression of some pathways can lead to metabolic burden or toxicity [21,22]. Therefore, engineers have turned to model-guided metabolic engineering platforms to provide insight into which genes to target and how strongly to activate or repress them to maximize product yield and accelerate the design-build-test-learn (DBTL) cycle.

Genome-scale metabolic models (GEMs) are mathematical representations of metabolic networks based on experimental data and biochemical assumptions such as steady-state. Flux balance analysis (FBA) is one method to predict the movement of metabolites through a GEM. Algorithms such as FBA make use of a user-defined objective function and other constraints defined by the GEM to provide insight into phenotypic effects of genetic perturbations [23]. FBA is based on constructing the objective function, assuming steady-state, and the use of linear programming to solve for a flux distribution. The flux distribution output fulfills the constraints derived by the GEM and optimizes the user-defined objective function [24,25]. The solver will maximize or minimize the objective function depending on the user definition. The choice of the objective

function has a large effect on the output flux distribution and therefore the closeness of the distribution to measured data [26,27]. In industrial biotechnology, algorithms derived from FBA have long been used to identify genetic interventions that maximize product yield [28–30].

Traditionally, genetic interventions are implemented through strain engineering to create knockouts or heterologously overexpress pathway genes. However, there are recent examples of more complex tools that can recommend optimal enzyme levels [31,32] or targets for dynamic regulation [33,34]. CRISPR systems may be particularly well-suited for implementing recommendations from these tools due to their tunable and dynamic capabilities. Additionally, there is growing attention in the use of the CRISPRa/i system to regulate gene targets identified through flux balance analysis [17,35–37]. Still, there is a gap in our ability to reliably predict the degree to which a CRISPRa/i perturbation will affect gene expression, and the degree to which changes in expression will affect enzyme activity for a given gene [38,39].

Although there are many methods able to recommend multiplexed interventions from GEMs, slow strain engineering cycles often limits the ability to broadly explore these design spaces [28,40]. The use of CRISPR tools to implement perturbations may address this limitation. The demonstrated ability to simultaneously activate and repress multiple genes using CRISPRa/i in microorganisms is a fairly recent development. Lian and colleagues were the first to demonstrate a tri-functional CRISPR system; taking advantage of activation, interference, and deletion in yeast [2]. Their system achieved significantly improved titers of two metabolic engineering targets through combinatorial perturbations of up to 36 genes. Similar tri-functional systems have yet to be demonstrated in other model organisms for biosynthesis. However, other experimental work has shown that combining CRISPRa/i control can help improve titers over a single method of regulation [2,8,41]. There are several notable examples of strain design algorithms that use genome-scale models to recommend perturbations to multiple genes through multiplexed activation, repression, and knockouts [28–30,40,42–44]. These tools are particularly well suited for use with multi-functional CRISPR-based metabolic engineering programs as perturbations can be quickly implemented in combination. However, none of these strain design tools have been validated with CRISPR-based perturbations. Instead, most

strain engineering has been achieved through gene knockouts and heterologous overexpression. The use of multi-functional CRISPR tools could greatly accelerate the prototyping of strains recommended by these algorithms, providing insight into the effects of finely-tuned and combined perturbations.

To take greater advantage of the tunability accessible with CRISPR tools, current strain design tools may need to be adapted. Updating constraint-based genome-scale models that focus solely on deletion or knockdowns to allow activation through a given pathway should be straightforward [28,42]. CRISPR-Cas tools have made it faster and easier to implement multi-functional metabolic engineering programs. Therefore, engineers should be encouraged to use computational tools that match their experimental capabilities.

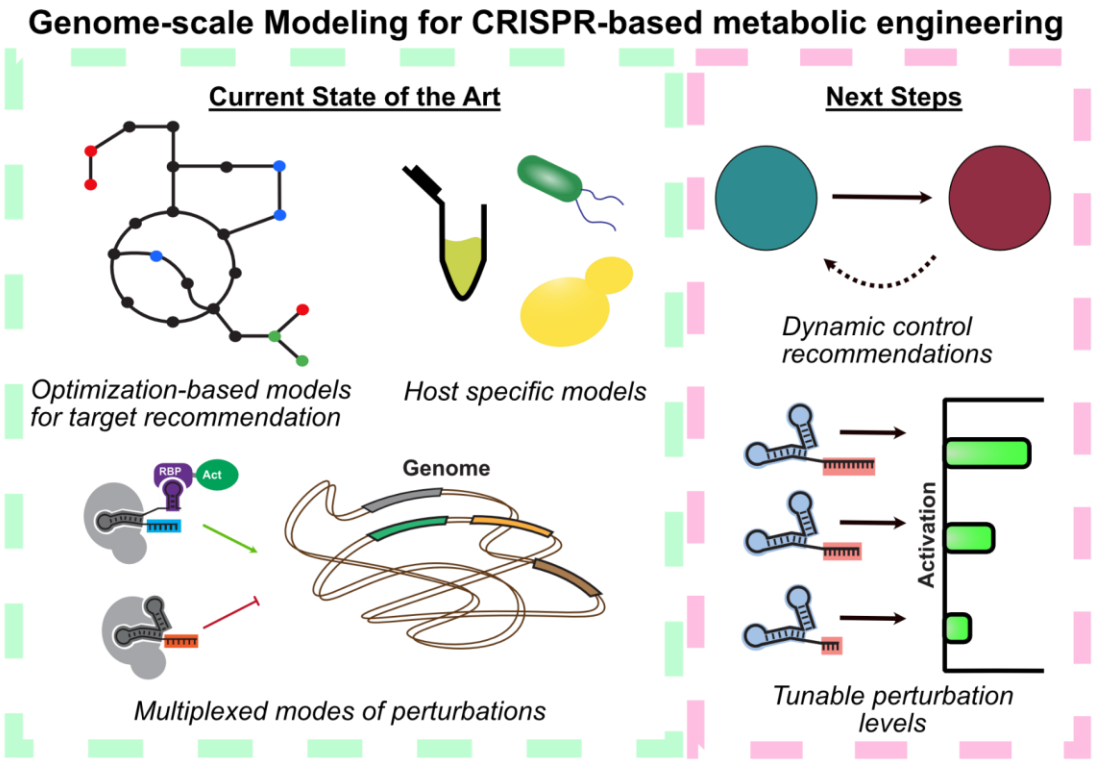


Figure 1. Current tools and opportunities for improvement of genome-scale models for metabolic engineering.

In recent years, the development of CRISPR tools has enabled programmable regulation at genomic targets in a wide array of novel microbes [45]. However, identifying relevant targets for metabolic engineering can be challenging for microbes with industrially useful but understudied metabolisms. To address this issue, computational tools can be used to generate novel metabolic models. Model reconstruction can be achieved using species-specific genomic data and "top-down" or "bottom-up" reconstruction methods, or through a combination of these approaches [46–49]. Typically, "top-down" approaches use a universal template model to assist with gap-filling, while a "bottom-up" approach may rely more on sequence homology and experimental data to derive metabolic network topologies [46–49]. Crucially, these tools minimize the need for manual construction and are broadly applicable to diverse bacterial species. Recent work has demonstrated that the development of genome-scale models can be transformative for realizing the potential for metabolic engineering for non-model organisms with unique, desirable metabolic properties [50]. For example, a previously published GEM and thermodynamic model of the Wood-Ljungdahl pathway in *Clostridium ljungdahlii* provided novel insights on autotrophic metabolism and potential pathways to high-value products generated from syngas. This model then informed the first example of CRISPRi-based metabolic engineering in *Clostridium ljungdahlii* several years later [51–53].

To circumvent the need for a new model of each unique organism's metabolism, several groups have developed machine learning-based (ML) strain design tools. These tools are able to recommend targets to improve bioproduction and accelerate the DBTL cycle without a full mechanistic understanding of the biological system [31,54,55]. Both the Automated Recommendation Tool (ART) and EVOLVE use Bayesian ensemble approaches to make strain recommendations with minimal training instances [31,54]. Bayesian models often outperform other ML algorithms on datasets that are small or noisy, making them a common choice for biological datasets [55]. From initial experimentation, the user provides input parameters such as transcriptome data or promoter strengths, as well as the observed output, e.g. product yield. The model then returns predicted output levels for different input profiles. In a direct comparison of ART and EVOLVE, the tools showed similar overall performances for recommendations of promoter strengths for a five-gene tryptophan overproduction strain [54].

While a meaningful step towards realizing an organism-agnostic model for metabolic engineering, machine learning approaches typically require a great deal of input data to generate a predictive model. Moreover, current ML based methods are generally unable to provide a rationale for their predictions. The difficulty in automated and high-throughput data collection and analysis is a major challenge to the practical use of machine learning methods for synthetic biology. An integrative tool that combines machine learning techniques and genome-scale modeling may offer the predictive power of machine learning algorithms with minimal training data [54]. This could theoretically be achieved by feeding fluxes determined from the literature and GEM simulations as inputs to existing machine learning tools. In addition, by combining ML with mechanistic modeling, better insights can be achieved to understand why a given set of perturbations fails to achieve a desired phenotype.

Cell-free systems (CFS) have recently gained popularity as a means for rapidly prototyping metabolic engineering programs and improving experimental throughput [56,57]. The use of CFS as a metabolic engineering platform offers many advantages over traditional *in vivo* systems, such as reduced carbon diversion towards endogenous metabolism and potential for exotic and toxic chemistries [58–60]. However, models of CFS tend to focus on protein expression and resource competition and lack consideration for the presence of central metabolism and endogenous enzymes in the CFS [61–66]. Recently, Martin et al. constructed a kinetic model of butanol production in a non-steady state CFS [67]. Their model incorporated over 200 reactions, metabolites, and inhibitor-enzyme pairs to accurately capture interactions between heterologous and endogenous metabolic networks and offer unintuitive predictions to improve butanol production. Their model was effective in identifying aldehyde dehydrogenase as the step with the most flux control using minimal training data, and lending insights into the effect of pairwise perturbations on butanol production. While these insights may have been possible without a large ensemble modeling approach, a significant advantage of the ensemble is that the scope of the model may lend itself well for being easily applied to other cell-free metabolic engineering efforts with minimal training data. Still, the process of generating individual knockout strains and cell-free lysates can be time-intensive. To address this limitation, other work from this group has explored the use of multiplexed CRISPR tools in a *S.*

cerevisiae CFS to improve production of 2,3-butanediol [68]. The use of CRISPR tools allowed the authors to rapidly assess combinatorial perturbations in the CFS without the need for extensive engineering of the underlying lysate strain. Further work to integrate these two efforts by screening model-informed combinatorial CRISPR perturbations may greatly accelerate DBTL cycles for cell-free metabolic engineering.

Metabolic engineers have long been interested in leveraging dynamic control to separate growth and production phases, maintain healthy cell growth, and ultimately improve product yields [69,70]. Dynamic regulation can describe simple systems that rely on addition of inducers to initiate gene expression, and autonomous systems that regulate flux based on sensing the internal state of the cell [69,71]. As chemical inducers are often too expensive for large-scale fermentation, there is growing interest in the rational design of self-regulating gene expression programs [19]. For effective dynamic regulation, metabolic engineers may require simultaneous overexpression and repression at several gene targets [70,72]. Engineers must also avoid redundant genetic sequences, such as the repeated use of inducible promoters, as they may be prone to recombination [73]. The programmability and orthogonality of the CRISPR system addresses these challenges, and has enabled significant improvements in the ability of synthetic biologists to dynamically control genes in response to environmental stimuli [74–76].

Despite the advantages of CRISPR tools, it remains difficult to implement regulation at the correct time and intensity during fermentation. For example, dynamic control in response to the presence of a toxic intermediate must be rapid and sensitive to prevent accumulation while maintaining flux through the desired pathway [69,71,77]. To date, most algorithms recommending perturbations have been limited to static control, leaving them unable to provide valuable insight as to when and how to implement dynamic control [78]. A model for dynamic control would be difficult to construct in practice as it may require further parameterization to understand how implementing perturbations at different stages of growth affects metabolism [79]. The integration of mathematical optimization with kinetic models of gene expression will improve our ability to predictably implement dynamic control for metabolic engineering programs using CRISPR-Cas systems [36,80].

III. Model-informed guide RNA design

gRNA design tools typically aim to maximize either gRNA specificity or efficiency. gRNA specificity is important to ensure that a given gRNA does not bind off-target in the genome, causing undesired growth defects or changes to metabolism. Tools for gRNA efficiency also ensure that the designed gRNA will fold properly, based on predicted RNA structure and sequence context. There are several recent reviews discussing gRNA design in great detail with focuses on machine learning and web-based tools [81,82]. In this review, we focus on models with the greatest impact for constructing multi-gene biosynthetic programs. Therefore, special attention is paid to models that give consideration to systems-level design, CRISPRa/i efficiency, gRNA competition, and portability to microorganisms.

There have been many models developed to predict gRNA performance, quantified by off- and on-target effects, using combinations of machine learning [38,82–84], kinetic [85,86], and thermodynamic models [14,85,87]. While early gRNA prediction models have been around for nearly a decade, it was only recently that large-scale models were updated to include functionality for CRISPRa/i targeting in addition to genome editing [88]. For CRISPRa, predicting high-performing guides may be challenging as targetable positions are limited to PAMs in the promoter region. CRISPRi target positions are more flexible, but knockdowns closer to the TSS have been shown to be more effective [89]. Therefore, there is significant interest in computational models that can predict ideal target sites and guide RNA sequences for a gene of interest.

The target sequence has been shown to have large effects on the degree of repression or activation achievable with CRISPRa/i control, in part due to both primary sequence composition and secondary structure of the gRNA [9,90,91]. CRISPRa/i control is highly sensitive to target sequence composition in both prokaryotic and eukaryotic systems. However, models that rely solely on gRNA sequence information are vastly outperformed by those that incorporate sequence context and structural information as well, highlighting the complexity of effective gRNA design [83,92]. gRNA efficiency is highly influenced by sequence context, both with respect to the sequence surrounding the PAM site [87] and the local structure at the target site [93,94]. Farasat & Salis (2016) demonstrated the importance of incorporating sequence context and structural information into gRNA

design by developing an integrative biophysical model of CRISPR-Cas9 activity. Their model incorporates statistical thermodynamics and kinetics with next-gen sequencing data on on- and off-target specificity to build a system-wide understanding of Cas9 on- and off-target effects in the genome [86]. Their mechanistic model includes kinetic representations of gRNA expression and folding, as well as thermodynamic representations of gRNA-Cas9 binding and Cas9-DNA binding. With this model, they unveil novel insights about differences in Cas9 off-target activity across organisms that previous studies focusing on gRNA sequence failed to identify. According to their model, the PAM-proximal seed region is the most important determinant for dCas9 binding affinity. Other kinetic models have found the free-energy change from the formation of the R-loop needed for DNA displacement is crucial for dictating Cas9 efficiency [95]. These results may be important considerations for the design of effective gRNAs, especially for targets with few available PAM sites.

A great deal about the tunability of CRISPR systems has been learned through the development and implementation of gRNA models. The use of truncated and mismatched gRNAs to tune the level of response in CRISPRa/i systems [96], a rule now leveraged often in metabolic engineering applications [97], was described in the CRISPRscan model [98,99]. The use of mismatched sgRNAs for systematic titration of gene expression has further been modeled by Jost et al. [100]. When using truncations or mismatches to tune CRISPR activity, the potential for off-target effects increases as gRNA specificity decreases. These models provide important rules for the design of effective guides.

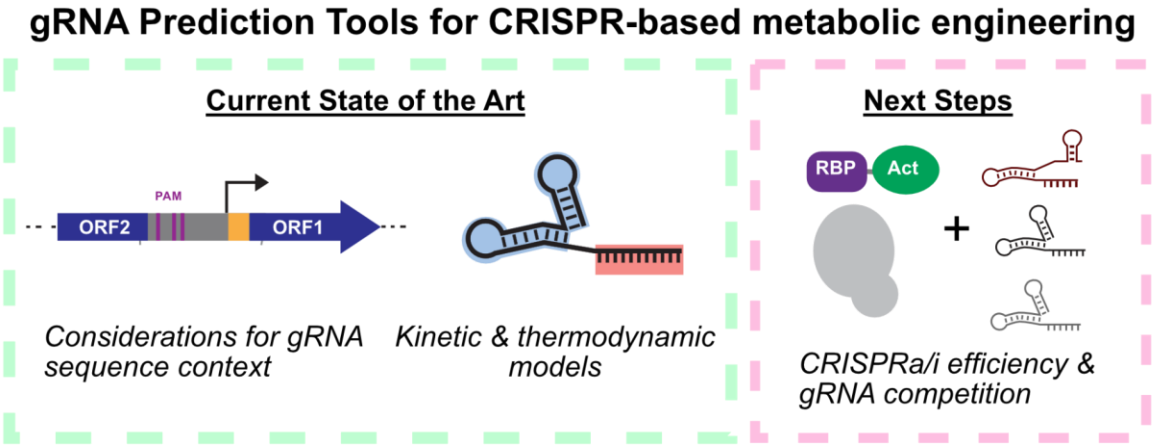


Figure 2. Current considerations and opportunities for improvement of gRNA design tools for metabolic engineering.

Many gRNA models developed for eukaryotic systems are not portable to prokaryotic systems, and vice-versa. Differences in genome organization and rules for effective interference and activation limit the generalizability of these models across branches of life. Wang & Zhang sought to overcome this limitation by developing a deep learning model based on RNA folding parameters, melting temperatures, and experimental data from bacterial cells to accurately predict on- and off-target activities for genome targeting sgRNAs and then use transfer learning to apply their model to eukaryotic systems as well [84]. Transfer learning is a form of machine learning useful to transfer parameters optimized from one dataset to another model with limited data. However, their experimental results confirmed previous findings [101,102] that gRNA activity prediction models developed for eukarya or prokarya are not appropriate for the other, likely due to differences in chromatin structure and genome organization. For microbial metabolic engineering, this results in a limited number of reliable gRNA design tools. Human cell lines are the most popular chassis for the creation of guide prediction models and software. However, these models are generally not transferable to metabolic engineering contexts, which typically require an understanding of CRISPR performance in bacteria or yeast. There is a clear need in the field for a gRNA prediction software suited towards metabolic engineers; that can accurately predict on- and off-target gRNA performance in relevant model organisms.

Beyond predicting gRNA efficiency for a single gene target, several models have investigated the impact of resource competition on the efficacy of CRISPR perturbations. Specifically, the decrease in CRISPR efficiency that occurs as a result of dCas9 resource competition from expressing multiple orthogonal gRNAs is well-characterized [95,103–105]. Clamons & Murray proposed a straightforward model for resource competition in which the expression of competing gRNAs reduces the available pool of CRISPR complexes able to perform CRISPRi at each target site by roughly 1/N, where N is the number of gRNAs. The Clamons & Murray model suggests that CRISPRa is more tolerant to resource competition than CRISPRi, however subsequent work has challenged this claim

[106]. In practice, CRISPRa is likely similarly affected by resource competition as CRISPRi because the number of activator complexes in the cell may be limiting. Recently, Barbier *et al.* showed that differences in gRNA binding affinities between scRNAs and sgRNAs leads to difficulty in predicting the behavior of combined CRISPRa/i programs in bacteria [107]. Through the use of a mathematical model for dCas9 resource competition and experimental validation, they demonstrated that using gRNAs with similar structures reduces differences in gRNA binding affinity and improves the function of multi-gRNA programs. These results suggest that the use of computational tools for predicting gRNA folding and dCas9 binding affinity will be crucial for constructing effective multi-gene CRISPRa/i metabolic engineering programs.

While resource competition in CRISPRa and CRISPRi has been modeled for basic regulatory networks, the effect of dCas9 and activator resource competition on metabolic engineering productivity remains understudied. This is due in part to a lack of integrative models that consider the capabilities of CRISPR systems while making strain recommendations. For example, a model that can account for gRNA competition as the number of genomic targets increases may help engineers prioritize which sets of genes to target. Incorporating a simple gRNA competition model into strain recommendation tools would be straightforward; however, more complex gRNA considerations such as target sequence context may be more difficult to integrate. Moving forward, these integrative models will be an important consideration when building large combinatorial metabolic engineering platforms that aim to simultaneously up- and down- regulate multiple genes, as there is an inherent tradeoff between the number of targets and the degree of perturbation.

IV. Outlook and Future Directions

Significant advances in metabolic engineering have been achieved through the use of model-informed CRISPR transcriptional control programs. Models have enabled the prediction of effective gene targets for perturbation and the rational design of high performing gRNA sequences. Still, there are many limitations to existing tools that, if overcome, would yield substantial benefits to the field of metabolic engineering. Current models are helpful for identifying targets for up- or down-regulation [10–12,16,17,35,36,80,108], but

cannot effectively predict the optimal amount of intervention, as total activation or repression of a gene is rarely the optimal perturbation [19,37,55]. Furthermore, CRISPRa/i tools will enable improved dynamic control over metabolic engineering programs [22]. As dynamic control becomes more prevalent, models able to integrate kinetic predictions with strain optimization will prove increasingly valuable.

CRISPR-Cas systems offer several advantages over other gene regulation tools that make them particularly well-suited for implementing model-driven recommendations. Compared to other tools such as ZF or TALE proteins, CRISPR tools have more ease of programmability, higher modularity and multiplexing capabilities, and lower off-target effects [109,110]. Additionally, over the last several years, CRISPR-Cas tools have been developed in a diverse range of microorganisms, expanding the scope of available chassis for metabolic engineering beyond traditional hosts. While numerous systems are available for implementing genetic interventions, CRISPR-based tools are uniquely able to accelerate engineering and DBTL cycles for strain design informed by genome-scale modeling [109].

Synthetic biologists have recently become interested in how machine learning can advance metabolic engineering goals [111]. The design of effective gRNAs is a well-suited problem for machine learning. The value of machine learning for gRNA design has been demonstrated by the use of a deep learning algorithm to predict gRNA on- and off-target effects for dCas13, a relatively understudied Cas variant [112]. Yu et al. recently developed a machine learning approach using publicly available depletion screens for prediction of CRISPRi guide efficiency that outperformed previous models. In addition, their model highlights the importance of gene expression levels and gene-specific features in making accurate predictions, providing novel insights into the understanding of CRISPRi screens for future work [38]. There are already many existing datasets and large libraries that can easily be screened in a high-throughput fashion if linked to a fitness score [113,114] or biosensor output [115,116].

For strain design, the use of machine learning is less straightforward. Ideally, machine learning-based strain design algorithms would be fit on multi-omics datasets, including transcriptome, proteome, and metabolome data [31,117]. Minimally, metabolomics data would be needed to relate strain genotypes to production phenotypes [97,117].

Alternatively, metabolite-responsive biosensors can be used to generate input data, but this approach is restricted by the space of available biosensors and their respective dynamic ranges [54,118]. Additionally, biosensor outputs do not necessarily correlate strongly with measured metabolite production [54]. In any case, the time, labor, and costs required to collect sufficient perturbation data to create an effective model is limiting [111]. Porting a model to another organism, or the same organism with a novel heterologous pathway, may lead to a decrease in predictive power without the collection of new training data. Alternatively, the use of mechanistic models leads to hypothesis-driven experimentation, which is more cost and time efficient in the near- and long-term. Additionally, there is still interest in understanding the underlying mechanisms behind genotype/phenotype relationships. The use of machine learning hides these relationships, which can introduce challenges when trying to understand how or why machine-learning aided design fails.

The use of constraint-based genome-scale models has been shown to be useful for predicting the phenotypic effect of knockouts. However, such models are less informative when making predictions based on more subtle changes to protein expression, which may be achieved through CRISPR control. Metabolic pathways include gene regulatory networks, and allosteric and feedback control, both of which play critical roles in shaping an organism's phenotype. For example, changing the expression level of a single gene will change metabolite levels which, through feedback mechanisms, will cause changes in protein expression at other sites, resulting in further downstream effects. Such effects cannot be easily captured by current constraint-based models.

Ideally, more detailed kinetic models could be developed to address challenges like feedback effects and regulatory networks, allowing metabolic engineers to make more precise predictions. While kinetic models are difficult to build, there is a growing list of success stories. For example, Martin *et al.* [67] built an in vitro pathway model of butanol production consisting of over 200 reactions, van Niekerk *et al.* [119] developed a detailed and validated model of energy metabolism in *P. falciparum*, and Millard *et al.* [120] built a validated kinetic model of core metabolism in *E. coli*. As kinetic models become larger, there is also growing importance in the reproducibility of these models [121], as a considerable amount of intellectual effort and expense is used to develop them.

Other modeling approaches include building approximate kinetic models based on lin-log kinetics [122,123]. Approximate kinetic models are simpler than full mechanistic models, while also giving engineers direct access to the pathway sensitivities through metabolic control analysis [67,122]. However, a disadvantage of lin-log models is that their predictive power drops if changes in gene expression are too large. Nevertheless, these models may still be able to predict the most promising sites for CRISPR perturbations.

As the available toolbox for metabolic engineering using CRISPR/Cas tools grows, the need for quantitative and predictive models to inform experiments will become increasingly valuable. Integrative models capable of suggesting genetic perturbations, selecting effective gRNAs, and composing selections into large multi-gene CRISPRa/i programs will greatly improve our ability to rapidly survey combinatorial perturbations and accelerate the DBTL cycle.

Acknowledgements

We thank Diego Alba Burbano, Benjamin I. Tickman, Cholpisit (Ice) Kiattisewee, and all members of the Carothers and Zalatan groups for advice, materials, and comments on the manuscript.

Competing interests

J.G.Z. and J.M.C are members of the Wayfinder Biosciences scientific advisory board. R.A.L.C., J.G.Z, and J.M.C. are inventors on patents and/or patent applications filed by the University of Washington that describe CRISPRa/i tools in prokaryotic systems.

Funding information

This work was supported by US National Science Foundation (NSF) Award MCB 2225632 MCB (J.M.C. and J.G.Z.) and US Department of Energy (DOE) Award DE-SC0023091 (J.M.C., J.G.Z., and H.M.S.). 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.

Author Information

Corresponding Authors

Ryan A. L. Cardiff - *Molecular Engineering & Sciences Institute, Center for Synthetic Biology, Department of Chemistry, Department of Chemical Engineering, University of Washington, Seattle, WA 98195, United States,* https://orcid.org/0009-0001-0530-9010, Email: cardiffr@uw.edu

Herbert M. Sauro - *Molecular Engineering & Sciences Institute, Center for Synthetic Biology, Department of Bioengineering, University of Washington, Seattle, WA 98195, United States,* https://orcid.org/0000-0002-3659-6817, Email: hsauro@uw.edu

Authors

Jesse G. Zalatan - *Molecular Engineering & Sciences Institute, Center for Synthetic Biology, and Department of Chemistry, University of Washington, Seattle, WA 98195, United States,* https://orcid.org/0000-0002-1458-0654, Email: zalatan@uw.edu

James M. Carothers - *Molecular Engineering & Sciences Institute, Center for Synthetic Biology, Department of Chemical Engineering, University of Washington, Seattle, WA 98195, United States,* https://orcid.org/0000-0001-6728-7833, Email: jcaroth@uw.edu

Author Contributions

R.A.L.C. and H.M.S. wrote the manuscript with input and supervision by J.M.C. and J.G.Z.

References

(1) Li, S.; Jendresen, C. B.; Grünberger, A.; Ronda, C.; Jensen, S. I.; Noack, S.; Nielsen, A. T. Enhanced Protein and Biochemical Production Using CRISPRi-Based Growth Switches. *Metab. Eng.* **2016**, *38*, 274–284. https://doi.org/10.1016/j.ymben.2016.09.003.

(2) Lian, J.; HamediRad, M.; Hu, S.; Zhao, H. Combinatorial Metabolic Engineering Using an Orthogonal Tri-Functional CRISPR System. *Nat. Commun.* **2017**, *8* (1), 1688. https://doi.org/10.1038/s41467-017-01695-x.

(3) Kiattisewee, C.; Dong, C.; Fontana, J.; Sugianto, W.; Peralta-Yahya, P.; Carothers, J. M.; Zalatan, J. G. Portable Bacterial CRISPR Transcriptional Activation Enables Metabolic Engineering in Pseudomonas Putida. *Metab. Eng.* **2021**, *66*, 283–295. https://doi.org/10.1016/j.ymben.2021.04.002.

(4) Qi, L. S.; Larson, M. H.; Gilbert, L. A.; Doudna, J. A.; Weissman, J. S.; Arkin, A. P.; Lim, W. A. Repurposing CRISPR as an RNA-Guided Platform for Sequence-Specific Control of Gene Expression. *Cell* **2013**, *152* (5), 1173–1183. https://doi.org/10.1016/j.cell.2013.02.022.

(5) Bikard, D.; Jiang, W.; Samai, P.; Hochschild, A.; Zhang, F.; Marraffini, L. A. Programmable Repression and Activation of Bacterial Gene Expression Using an Engineered CRISPR-Cas System. *Nucleic Acids Res.* **2013**, *41* (15), 7429–7437. https://doi.org/10.1093/nar/gkt520.

(6) 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.; Lim, W. A. Engineering Complex Synthetic Transcriptional Programs with CRISPR RNA Scaffolds. *Cell* **2015**, *160* (1), 339–350. https://doi.org/10.1016/j.cell.2014.11.052.

(7) Dong, C.; Fontana, J.; Patel, A.; Carothers, J. M.; Zalatan, J. G. Synthetic CRISPR-Cas Gene Activators for Transcriptional Reprogramming in Bacteria. *Nat. Commun.* **2018**, *9* (1), 2489. https://doi.org/10.1038/s41467-018-04901-6.

(8) McCarty, N. S.; Graham, A. E.; Studená, L.; Ledesma-Amaro, R. Multiplexed CRISPR Technologies for Gene Editing and Transcriptional Regulation. *Nat. Commun.* **2020**, *11* (1), 1281. https://doi.org/10.1038/s41467-020-15053-x.

(9) Fontana, J.; Sparkman-Yager, D.; Zalatan, J. G.; Carothers, J. M. Challenges and Opportunities with CRISPR Activation in Bacteria for Data-Driven Metabolic Engineering. *Curr. Opin. Biotechnol.* **2020**, *64*, 190–198. https://doi.org/10.1016/j.copbio.2020.04.005.

(10) McAnulty, M. J.; Yen, J. Y.; Freedman, B. G.; Senger, R. S. Genome-Scale Modeling Using Flux Ratio Constraints to Enable Metabolic Engineering of Clostridial Metabolism in Silico. *BMC Syst. Biol.* **2012**, *6* (1), 42. https://doi.org/10.1186/1752-0509-6-42.

(11) Cautha, S. C.; Gowen, C. M.; Lussier, F.-X.; Gold, N. D.; Martin, V. J. J.; Mahadevan, R. Model-Driven Design of a Saccharomyces Cerevisiae Platform Strain with Improved Tyrosine Production Capabilities. *IFAC Proc. Vol.* **2013**, *46* (31), 221–226. https://doi.org/10.3182/20131216-3-IN-2044.00066.

(12) Banerjee, D.; Eng, T.; Lau, A. K.; Sasaki, Y.; Wang, B.; Chen, Y.; Prahl, J.-P.; Singan, V. R.; Herbert, R. A.; Liu, Y.; Tanjore, D.; Petzold, C. J.; Keasling, J. D.; Mukhopadhyay, A. Genome-Scale Metabolic Rewiring Improves Titers Rates and Yields of the Non-Native Product Indigoidine at Scale. *Nat. Commun.* **2020**, *11* (1),

5385. https://doi.org/10.1038/s41467-020-19171-4.
(13) Li, W.; Xu, H.; Xiao, T.; Cong, L.; Love, M. I.; Zhang, F.; Irizarry, R. A.; Liu, J. S.; Brown, M.; Liu, X. S. MAGeCK Enables Robust Identification of Essential Genes from Genome-Scale CRISPR/Cas9 Knockout Screens. *Genome Biol.* **2014**, *15* (12), 554. https://doi.org/10.1186/s13059-014-0554-4.
(14) Doench, J. G.; Fusi, N.; Sullender, M.; Hegde, M.; Vaimberg, E. W.; Donovan, K. F.; Smith, I.; Tothova, Z.; Wilen, C.; Orchard, R.; Virgin, H. W.; Listgarten, J.; Root, D. E. Optimized sgRNA Design to Maximize Activity and Minimize Off-Target Effects of CRISPR-Cas9. *Nat. Biotechnol.* **2016**, *34* (2), 184–191. https://doi.org/10.1038/nbt.3437.
(15) Labun, K.; Montague, T. G.; Gagnon, J. A.; Thyme, S. B.; Valen, E. CHOPCHOP v2: A Web Tool for the next Generation of CRISPR Genome Engineering. *Nucleic Acids Res.* **2016**, *44* (W1), W272–W276. https://doi.org/10.1093/nar/gkw398.
(16) Kim, S. K.; Seong, W.; Han, G. H.; Lee, D.-H.; Lee, S.-G. CRISPR Interference-Guided Multiplex Repression of Endogenous Competing Pathway Genes for Redirecting Metabolic Flux in Escherichia Coli. *Microb. Cell Factories* **2017**, *16* (1), 188. https://doi.org/10.1186/s12934-017-0802-x.
(17) Kozaeva, E.; Volkova, S.; Matos, M. R. A.; Mezzina, M. P.; Wulff, T.; Volke, D. C.; Nielsen, L. K.; Nikel, P. I. Model-Guided Dynamic Control of Essential Metabolic Nodes Boosts Acetyl-Coenzyme A–Dependent Bioproduction in Rewired Pseudomonas Putida. *Metab. Eng.* **2021**, *67*, 373–386. https://doi.org/10.1016/j.ymben.2021.07.014.
(18) Chavez, A.; Scheiman, J.; Vora, S.; Pruitt, B. W.; Tuttle, M.; P R Iyer, E.; Lin, S.; Kiani, S.; Guzman, C. D.; Wiegand, D. J.; Ter-Ovanesyan, D.; Braff, J. L.; Davidsohn, N.; Housden, B. E.; Perrimon, N.; Weiss, R.; Aach, J.; Collins, J. J.; Church, G. M. Highly Efficient Cas9-Mediated Transcriptional Programming. *Nat. Methods* **2015**, *12* (4), 326–328. https://doi.org/10.1038/nmeth.3312.
(19) Holtz, W. J.; Keasling, J. D. Engineering Static and Dynamic Control of Synthetic Pathways. *Cell* **2010**, *140* (1), 19–23. https://doi.org/10.1016/j.cell.2009.12.029.
(20) Schober, A. F.; Mathis, A. D.; Ingle, C.; Park, J. O.; Chen, L.; Rabinowitz, J. D.; Junier, I.; Rivoire, O.; Reynolds, K. A. A Two-Enzyme Adaptive Unit within Bacterial Folate Metabolism. *Cell Rep.* **2019**, *27* (11), 3359-3370.e7. https://doi.org/10.1016/j.celrep.2019.05.030.
(21) Zhang, F.; Carothers, J. M.; Keasling, J. D. Design of a Dynamic Sensor-Regulator System for Production of Chemicals and Fuels Derived from Fatty Acids. *Nat. Biotechnol.* **2012**, *30* (4), 354–359. https://doi.org/10.1038/nbt.2149.
(22) Gupta, A.; Reizman, I. M. B.; Reisch, C. R.; Prather, K. L. J. Dynamic Regulation of Metabolic Flux in Engineered Bacteria Using a Pathway-Independent Quorum-Sensing Circuit. *Nat. Biotechnol.* **2017**, *35* (3), 273–279. https://doi.org/10.1038/nbt.3796.
(23) Fang, X.; Lloyd, C. J.; Palsson, B. O. Reconstructing Organisms in Silico: Genome-Scale Models and Their Emerging Applications. *Nat. Rev. Microbiol.* **2020**, *18* (12), 731–743. https://doi.org/10.1038/s41579-020-00440-4.
(24) Orth, J. D.; Thiele, I.; Palsson, B. Ø. What Is Flux Balance Analysis? *Nat. Biotechnol.* **2010**, *28* (3), 245–248. https://doi.org/10.1038/nbt.1614.
(25) Lewis, N. E.; Nagarajan, H.; Palsson, B. O. Constraining the Metabolic Genotype-

Phenotype Relationship Using a Phylogeny of in Silico Methods. *Nat. Rev. Microbiol.* **2012**, *10* (4), 291–305. https://doi.org/10.1038/nrmicro2737.
(26) Schuetz, R.; Kuepfer, L.; Sauer, U. Systematic Evaluation of Objective Functions for Predicting Intracellular Fluxes in Escherichia Coli. *Mol. Syst. Biol.* **2007**, *3* (1), 119. https://doi.org/10.1038/msb4100162.
(27) García Sánchez, C. E.; Torres Sáez, R. G. Comparison and Analysis of Objective Functions in Flux Balance Analysis. *Biotechnol. Prog.* **2014**, *30* (5), 985–991. https://doi.org/10.1002/btpr.1949.
(28) Pharkya, P.; Maranas, C. D. An Optimization Framework for Identifying Reaction Activation/Inhibition or Elimination Candidates for Overproduction in Microbial Systems. *Metab. Eng.* **2006**, *8* (1), 1–13. https://doi.org/10.1016/j.ymben.2005.08.003.
(29) *OptDesign: Identifying Optimum Design Strategies in Strain Engineering for Biochemical Production | ACS Synthetic Biology*. https://pubs.acs.org/doi/10.1021/acssynbio.1c00610 (accessed 2024-03-25).
(30) *Redirector: Designing Cell Factories by Reconstructing the Metabolic Objective | PLOS Computational Biology*. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002882 (accessed 2024-03-25).
(31) Radivojević, T.; Costello, Z.; Workman, K.; Garcia Martin, H. A Machine Learning Automated Recommendation Tool for Synthetic Biology. *Nat. Commun.* **2020**, *11* (1), 4879. https://doi.org/10.1038/s41467-020-18008-4.
(32) Sabzevari, M.; Szedmak, S.; Penttilä, M.; Jouhten, P.; Rousu, J. Strain Design Optimization Using Reinforcement Learning. *PLOS Comput. Biol.* **2022**, *18* (6), e1010177. https://doi.org/10.1371/journal.pcbi.1010177.
(33) van Rosmalen, R. P.; Smith, R. W.; Martins dos Santos, V. A. P.; Fleck, C.; Suarez-Diez, M. Model Reduction of Genome-Scale Metabolic Models as a Basis for Targeted Kinetic Models. *Metab. Eng.* **2021**, *64*, 74–84. https://doi.org/10.1016/j.ymben.2021.01.008.
(34) Venayak, N.; von Kamp, A.; Klamt, S.; Mahadevan, R. MoVE Identifies Metabolic Valves to Switch between Phenotypic States. *Nat. Commun.* **2018**, *9* (1), 5332. https://doi.org/10.1038/s41467-018-07719-4.
(35) *Bidirectional titration of yeast gene expression using a pooled CRISPR guide RNA approach*. PNAS. https://www.pnas.org/doi/abs/10.1073/pnas.2007413117 (accessed 2022-02-28).
(36) Dusad, V.; Thiel, D.; Barahona, M.; Keun, H. C.; Oyarzún, D. A. Opportunities at the Interface of Network Science and Metabolic Modeling. *Front. Bioeng. Biotechnol.* **2021**, *8*, 1570. https://doi.org/10.3389/fbioe.2020.591049.
(37) Liu, Z.; Robson, P.; Cheng, A. Simultaneous Multifunctional Transcriptome Engineering by CRISPR RNA Scaffold. bioRxiv June 22, 2022, p 2022.06.21.497089. https://doi.org/10.1101/2022.06.21.497089.
(38) Yu, Y.; Gawlitt, S.; de Andrade e Sousa, L. B.; Merdivan, E.; Piraud, M.; Beisel, C. L.; Barquist, L. Improved Prediction of Bacterial CRISPRi Guide Efficiency from Depletion Screens through Mixed-Effect Machine Learning and Data Integration. *Genome Biol.* **2024**, *25* (1), 13. https://doi.org/10.1186/s13059-023-03153-y.
(39) Wang, T.; Guan, C.; Guo, J.; Liu, B.; Wu, Y.; Xie, Z.; Zhang, C.; Xing, X.-H. Pooled

CRISPR Interference Screening Enables Genome-Scale Functional Genomics Study in Bacteria with Superior Performance. *Nat. Commun.* **2018**, *9* (1), 2475. https://doi.org/10.1038/s41467-018-04899-x.
(40) Mahadevan, R.; von Kamp, A.; Klamt, S. Genome-Scale Strain Designs Based on Regulatory Minimal Cut Sets. *Bioinformatics* **2015**, *31* (17), 2844–2851. https://doi.org/10.1093/bioinformatics/btv217.
(41) Shaw, W. M.; Studená, L.; Roy, K.; Hapeta, P.; McCarty, N. S.; Graham, A. E.; Ellis, T.; Ledesma-Amaro, R. Inducible Expression of Large gRNA Arrays for Multiplexed CRISPRai Applications. *Nat. Commun.* **2022**, *13* (1), 4984. https://doi.org/10.1038/s41467-022-32603-7.
(42) Shen, F.; Sun, R.; Yao, J.; Li, J.; Liu, Q.; Price, N. D.; Liu, C.; Wang, Z. OptRAM: In-Silico Strain Design via Integrative Regulatory-Metabolic Network Modeling. *PLOS Comput. Biol.* **2019**, *15* (3), e1006835. https://doi.org/10.1371/journal.pcbi.1006835.
(43) Kim, J.; Reed, J. L. OptORF: Optimal Metabolic and Regulatory Perturbations for Metabolic Engineering of Microbial Strains. *BMC Syst. Biol.* **2010**, *4* (1), 53. https://doi.org/10.1186/1752-0509-4-53.
(44) Razaghi-Moghadam, Z.; Nikoloski, Z. GeneReg: A Constraint-Based Approach for Design of Feasible Metabolic Engineering Strategies at the Gene Level. *Bioinformatics* **2021**, *37* (12), 1717–1723. https://doi.org/10.1093/bioinformatics/btaa996.
(45) Call, S. N.; Andrews, L. B. CRISPR-Based Approaches for Gene Regulation in Non-Model Bacteria. *Front. Genome Ed.* **2022**, *4*, 892304. https://doi.org/10.3389/fgeed.2022.892304.
(46) Machado, D.; Andrejev, S.; Tramontano, M.; Patil, K. R. Fast Automated Reconstruction of Genome-Scale Metabolic Models for Microbial Species and Communities. *Nucleic Acids Res.* **2018**, *46* (15), 7542–7553. https://doi.org/10.1093/nar/gky537.
(47) Zimmermann, J.; Kaleta, C.; Waschina, S. Gapseq: Informed Prediction of Bacterial Metabolic Pathways and Reconstruction of Accurate Metabolic Models. *Genome Biol.* **2021**, *22* (1), 81. https://doi.org/10.1186/s13059-021-02295-1.
(48) Mendoza, S. N.; Olivier, B. G.; Molenaar, D.; Teusink, B. A Systematic Assessment of Current Genome-Scale Metabolic Reconstruction Tools. *Genome Biol.* **2019**, *20* (1), 158. https://doi.org/10.1186/s13059-019-1769-1.
(49) Henry, C. S.; DeJongh, M.; Best, A. A.; Frybarger, P. M.; Linsay, B.; Stevens, R. L. High-Throughput Generation, Optimization and Analysis of Genome-Scale Metabolic Models. *Nat. Biotechnol.* **2010**, *28* (9), 977–982. https://doi.org/10.1038/nbt.1672.
(50) Greene, J.; Daniell, J.; Köpke, M.; Broadbelt, L.; Tyo, K. E. J. Kinetic Ensemble Model of Gas Fermenting Clostridium Autoethanogenum for Improved Ethanol Production. *Biochem. Eng. J.* **2019**, *148*, 46–56. https://doi.org/10.1016/j.bej.2019.04.021.
(51) Woolston, B. M.; Emerson, D. F.; Currie, D. H.; Stephanopoulos, G. Rediverting Carbon Flux in Clostridium Ljungdahlii Using CRISPR Interference (CRISPRi). *Metab. Eng.* **2018**, *48*, 243–253. https://doi.org/10.1016/j.ymben.2018.06.006.
(52) Schuchmann, K.; Müller, V. Autotrophy at the Thermodynamic Limit of Life: A

Model for Energy Conservation in Acetogenic Bacteria. *Nat. Rev. Microbiol.* **2014**, *12* (12), 809–821. https://doi.org/10.1038/nrmicro3365.
(53) Nagarajan, H.; Sahin, M.; Nogales, J.; Latif, H.; Lovley, D. R.; Ebrahim, A.; Zengler, K. Characterizing Acetogenic Metabolism Using a Genome-Scale Metabolic Reconstruction of Clostridium Ljungdahlii. *Microb. Cell Factories* **2013**, *12* (1), 118. https://doi.org/10.1186/1475-2859-12-118.
(54) Zhang, J.; Petersen, S. D.; Radivojevic, T.; Ramirez, A.; Pérez-Manríquez, A.; Abeliuk, E.; Sánchez, B. J.; Costello, Z.; Chen, Y.; Fero, M. J.; Martin, H. G.; Nielsen, J.; Keasling, J. D.; Jensen, M. K. Combining Mechanistic and Machine Learning Models for Predictive Engineering and Optimization of Tryptophan Metabolism. *Nat. Commun.* **2020**, *11* (1), 4880. https://doi.org/10.1038/s41467-020-17910-1.
(55) HamediRad, M.; Chao, R.; Weisberg, S.; Lian, J.; Sinha, S.; Zhao, H. Towards a Fully Automated Algorithm Driven Platform for Biosystems Design. *Nat. Commun.* **2019**, *10* (1), 5150. https://doi.org/10.1038/s41467-019-13189-z.
(56) Karim, A. S.; Dudley, Q. M.; Juminaga, A.; Yuan, Y.; Crowe, S. A.; Heggestad, J. T.; Garg, S.; Abdalla, T.; Grubbe, W. S.; Rasor, B. J.; Coar, D. N.; Torculas, M.; Krein, M.; Liew, F. (Eric); Quattlebaum, A.; Jensen, R. O.; Stuart, J. A.; Simpson, S. D.; Köpke, M.; Jewett, M. C. In Vitro Prototyping and Rapid Optimization of Biosynthetic Enzymes for Cell Design. *Nat. Chem. Biol.* **2020**, *16* (8), 912–919. https://doi.org/10.1038/s41589-020-0559-0.
(57) Krüger, A.; Mueller, A. P.; Rybnicky, G. A.; Engle, N. L.; Yang, Z. K.; Tschaplinski, T. J.; Simpson, S. D.; Köpke, M.; Jewett, M. C. Development of a Clostridia-Based Cell-Free System for Prototyping Genetic Parts and Metabolic Pathways. *Metab. Eng.* **2020**, *62*, 95–105. https://doi.org/10.1016/j.ymben.2020.06.004.
(58) Grubbe, W. S.; Rasor, B. J.; Krüger, A.; Jewett, M. C.; Karim, A. S. Cell-Free Styrene Biosynthesis at High Titers. *Metab. Eng.* **2020**, *61*, 89–95. https://doi.org/10.1016/j.ymben.2020.05.009.
(59) Kay, J. E.; Jewett, M. C. A Cell-Free System for Production of 2,3-Butanediol Is Robust to Growth-Toxic Compounds. *Metab. Eng. Commun.* **2020**, *10*, e00114. https://doi.org/10.1016/j.mec.2019.e00114.
(60) Dudley, Q. M.; Karim, A. S.; Jewett, M. C. Cell-Free Metabolic Engineering: Biomanufacturing beyond the Cell. *Biotechnol. J.* **2015**, *10* (1), 69–82. https://doi.org/10.1002/biot.201400330.
(61) Marshall, R.; Noireaux, V. Quantitative Modeling of Transcription and Translation of an All-E. Coli Cell-Free System. *Sci. Rep.* **2019**, *9* (1), 11980. https://doi.org/10.1038/s41598-019-48468-8.
(62) Tickman, B. I.; Burbano, D. A.; Chavali, V. P.; Kiattisewee, C.; Fontana, J.; Khakimzhan, A.; Noireaux, V.; Zalatan, J. G.; Carothers, J. M. Multi-Layer CRISPRa/i Circuits for Dynamic Genetic Programs in Cell-Free and Bacterial Systems. *Cell Syst.* **2021**, *0* (0). https://doi.org/10.1016/j.cels.2021.10.008.
(63) Moore, S. J.; MacDonald, J. T.; Wienecke, S.; Ishwarbhai, A.; Tsipa, A.; Aw, R.; Kylilis, N.; Bell, D. J.; McClymont, D. W.; Jensen, K.; Polizzi, K. M.; Biedendieck, R.; Freemont, P. S. Rapid Acquisition and Model-Based Analysis of Cell-Free Transcription–Translation Reactions from Nonmodel Bacteria. *Proc. Natl. Acad. Sci.* **2018**, *115* (19), E4340–E4349. https://doi.org/10.1073/pnas.1715806115.

(64) Lehr, F.-X.; Hanst, M.; Vogel, M.; Kremer, J.; Göringer, H. U.; Suess, B.; Koeppl, H. Cell-Free Prototyping of AND-Logic Gates Based on Heterogeneous RNA Activators. *ACS Synth. Biol.* **2019**, *8* (9), 2163–2173. https://doi.org/10.1021/acssynbio.9b00238.
(65) Koch, M.; Faulon, J.-L.; Borkowski, O. Models for Cell-Free Synthetic Biology: Make Prototyping Easier, Better, and Faster. *Front. Bioeng. Biotechnol.* **2018**, *6*, 182. https://doi.org/10.3389/fbioe.2018.00182.
(66) Vilkhovoy, M.; Horvath, N.; Shih, C.-H.; Wayman, J. A.; Calhoun, K.; Swartz, J.; Varner, J. D. Sequence Specific Modeling of E. Coli Cell-Free Protein Synthesis. *ACS Synth. Biol.* **2018**, *7* (8), 1844–1857. https://doi.org/10.1021/acssynbio.7b00465.
(67) Martin, J. P.; Rasor, B. J.; DeBonis, J.; Karim, A. S.; Jewett, M. C.; Tyo, K. E. J.; Broadbelt, L. J. A Dynamic Kinetic Model Captures Cell-Free Metabolism for Improved Butanol Production. *Metab. Eng.* **2023**, *76*, 133–145. https://doi.org/10.1016/j.ymben.2023.01.009.
(68) Rasor, B. J.; Yi, X.; Brown, H.; Alper, H. S.; Jewett, M. C. An Integrated in Vivo/in Vitro Framework to Enhance Cell-Free Biosynthesis with Metabolically Rewired Yeast Extracts. *Nat. Commun.* **2021**, *12* (1), 5139. https://doi.org/10.1038/s41467-021-25233-y.
(69) Brockman, I. M.; Prather, K. L. J. Dynamic Metabolic Engineering: New Strategies for Developing Responsive Cell Factories. *Biotechnol. J.* **2015**, *10* (9), 1360–1369. https://doi.org/10.1002/biot.201400422.
(70) Wu, Y.; Chen, T.; Liu, Y.; Lv, X.; Li, J.; Du, G.; Ledesma-Amaro, R.; Liu, L. CRISPRi Allows Optimal Temporal Control of N-Acetylglucosamine Bioproduction by a Dynamic Coordination of Glucose and Xylose Metabolism in Bacillus Subtilis. *Metab. Eng.* **2018**, *49*, 232–241. https://doi.org/10.1016/j.ymben.2018.08.012.
(71) Hartline, C. J.; Schmitz, A. C.; Han, Y.; Zhang, F. Dynamic Control in Metabolic Engineering: Theories, Tools, and Applications. *Metab. Eng.* **2021**, *63*, 126–140. https://doi.org/10.1016/j.ymben.2020.08.015.
(72) Wu, Y.; Chen, T.; Liu, Y.; Tian, R.; Lv, X.; Li, J.; Du, G.; Chen, J.; Ledesma-Amaro, R.; Liu, L. Design of a Programmable Biosensor-CRISPRi Genetic Circuits for Dynamic and Autonomous Dual-Control of Metabolic Flux in Bacillus Subtilis. *Nucleic Acids Res.* **2020**, *48* (2), 996–1009. https://doi.org/10.1093/nar/gkz1123.
(73) Watt, V. M.; Ingles, C. J.; Urdea, M. S.; Rutter, W. J. Homology Requirements for Recombination in Escherichia Coli. *Proc. Natl. Acad. Sci. U. S. A.* **1985**, *82* (14), 4768–4772.
(74) Nihongaki, Y.; Kawano, F.; Nakajima, T.; Sato, M. Photoactivatable CRISPR-Cas9 for Optogenetic Genome Editing. *Nat. Biotechnol.* **2015**, *33* (7), 755–760. https://doi.org/10.1038/nbt.3245.
(75) Cunningham-Bryant, D.; Sun, J.; Fernandez, B.; Zalatan, J. G. CRISPR–Cas-Mediated Chemical Control of Transcriptional Dynamics in Yeast. *ChemBioChem* **2019**, *20* (12), 1519–1523. https://doi.org/10.1002/cbic.201800823.
(76) Kundert, K.; Lucas, J. E.; Watters, K. E.; Fellmann, C.; Ng, A. H.; Heineike, B. M.; Fitzsimmons, C. M.; Oakes, B. L.; Qu, J.; Prasad, N.; Rosenberg, O. S.; Savage, D. F.; El-Samad, H.; Doudna, J. A.; Kortemme, T. Controlling CRISPR-Cas9 with Ligand-Activated and Ligand-Deactivated sgRNAs. *Nat. Commun.* **2019**, *10* (1),

2127. https://doi.org/10.1038/s41467-019-09985-2.
(77) Fontana, J.; Voje, W. E.; Zalatan, J. G.; Carothers, J. M. Prospects for Engineering Dynamic CRISPR–Cas Transcriptional Circuits to Improve Bioproduction. *J. Ind. Microbiol. Biotechnol.* **2018**, *45* (7), 481–490. https://doi.org/10.1007/s10295-018-2039-z.
(78) Lee, J. W.; Na, D.; Park, J. M.; Lee, J.; Choi, S.; Lee, S. Y. Systems Metabolic Engineering of Microorganisms for Natural and Non-Natural Chemicals. *Nat. Chem. Biol.* **2012**, *8* (6), 536–546. https://doi.org/10.1038/nchembio.970.
(79) Maia, P.; Rocha, M.; Rocha, I. In Silico Constraint-Based Strain Optimization Methods: The Quest for Optimal Cell Factories. *Microbiol. Mol. Biol. Rev.* **2015**, *80* (1), 45–67. https://doi.org/10.1128/mmbr.00014-15.
(80) Liu, D.; Mannan, A. A.; Han, Y.; Oyarzún, D. A.; Zhang, F. Dynamic Metabolic Control: Towards Precision Engineering of Metabolism. *J. Ind. Microbiol. Biotechnol.* **2018**, *45* (7), 535–543. https://doi.org/10.1007/s10295-018-2013-9.
(81) Motoche-Monar, C.; Ordoñez, J. E.; Chang, O.; Gonzales-Zubiate, F. A. gRNA Design: How Its Evolution Impacted on CRISPR/Cas9 Systems Refinement. *Biomolecules* **2023**, *13* (12), 1698. https://doi.org/10.3390/biom13121698.
(82) Konstantakos, V.; Nentidis, A.; Krithara, A.; Paliouras, G. CRISPR–Cas9 gRNA Efficiency Prediction: An Overview of Predictive Tools and the Role of Deep Learning. *Nucleic Acids Res.* **2022**, *50* (7), 3616–3637. https://doi.org/10.1093/nar/gkac192.
(83) Yu, Y.; Gawlitt, S.; de Andrade e Sousa, L. B.; Merdivan, E.; Piraud, M.; Beisel, C.; Barquist, L. *Improved Prediction of Bacterial CRISPRi Guide Efficiency through Data Integration and Automated Machine Learning*; preprint; Bioinformatics, 2022. https://doi.org/10.1101/2022.05.27.493707.
(84) Wang, L.; Zhang, J. Prediction of sgRNA On-Target Activity in Bacteria by Deep Learning. *BMC Bioinformatics* **2019**, *20* (1), 517. https://doi.org/10.1186/s12859-019-3151-4.
(85) Eslami-Mossallam, B.; Klein, M.; Smagt, C. V. D.; Sanden, K. V. D.; Jones, S. K.; Hawkins, J. A.; Finkelstein, I. J.; Depken, M. A Kinetic Model Predicts SpCas9 Activity, Improves off-Target Classification, and Reveals the Physical Basis of Targeting Fidelity. *Nat. Commun.* **2022**, *13* (1), 1367. https://doi.org/10.1038/s41467-022-28994-2.
(86) Farasat, I.; Salis, H. M. A Biophysical Model of CRISPR/Cas9 Activity for Rational Design of Genome Editing and Gene Regulation. *PLoS Comput. Biol.* **2016**, *12* (1), e1004724. https://doi.org/10.1371/journal.pcbi.1004724.
(87) Corsi, G. I.; Qu, K.; Alkan, F.; Pan, X.; Luo, Y.; Gorodkin, J. CRISPR/Cas9 gRNA Activity Depends on Free Energy Changes and on the Target PAM Context. *Nat. Commun.* **2022**, *13* (1), 3006. https://doi.org/10.1038/s41467-022-30515-0.
(88) Labun, K.; Montague, T. G.; Krause, M.; Torres Cleuren, Y. N.; Tjeldnes, H.; Valen, E. CHOPCHOP v3: Expanding the CRISPR Web Toolbox beyond Genome Editing. *Nucleic Acids Res.* **2019**, *47* (W1), W171–W174. https://doi.org/10.1093/nar/gkz365.
(89) Gilbert, L. A.; Horlbeck, M. A.; Adamson, B.; Villalta, J. E.; Chen, Y.; Whitehead, E. H.; Guimaraes, C.; Panning, B.; Ploegh, H. L.; Bassik, M. C.; Qi, L. S.; Kampmann, M.; Weissman, J. S. Genome-Scale CRISPR-Mediated Control of Gene

Repression and Activation. *Cell* **2014**, *159* (3), 647–661. https://doi.org/10.1016/j.cell.2014.09.029.
(90) Smith, J. D.; Suresh, S.; Schlecht, U.; Wu, M.; Wagih, O.; Peltz, G.; Davis, R. W.; Steinmetz, L. M.; Parts, L.; St.Onge, R. P. Quantitative CRISPR Interference Screens in Yeast Identify Chemical-Genetic Interactions and New Rules for Guide RNA Design. *Genome Biol.* **2016**, *17* (1), 45. https://doi.org/10.1186/s13059-016-0900-9.
(91) Kocak, D. D.; Josephs, E. A.; Bhandarkar, V.; Adkar, S. S.; Kwon, J. B.; Gersbach, C. A. Increasing the Specificity of CRISPR Systems with Engineered RNA Secondary Structures. *Nat. Biotechnol.* **2019**, *37* (6), 657–666. https://doi.org/10.1038/s41587-019-0095-1.
(92) Haeussler, M.; Schönig, K.; Eckert, H.; Eschstruth, A.; Mianné, J.; Renaud, J.-B.; Schneider-Maunoury, S.; Shkumatava, A.; Teboul, L.; Kent, J.; Joly, J.-S.; Concordet, J.-P. Evaluation of Off-Target and on-Target Scoring Algorithms and Integration into the Guide RNA Selection Tool CRISPOR. *Genome Biol.* **2016**, *17* (1), 148. https://doi.org/10.1186/s13059-016-1012-2.
(93) Chen, Y.; Zeng, S.; Hu, R.; Wang, X.; Huang, W.; Liu, J.; Wang, L.; Liu, G.; Cao, Y.; Zhang, Y. Using Local Chromatin Structure to Improve CRISPR/Cas9 Efficiency in Zebrafish. *PLoS ONE* **2017**, *12* (8), e0182528. https://doi.org/10.1371/journal.pone.0182528.
(94) Jensen, K. T.; Fløe, L.; Petersen, T. S.; Huang, J.; Xu, F.; Bolund, L.; Luo, Y.; Lin, L. Chromatin Accessibility and Guide Sequence Secondary Structure Affect CRISPR-Cas9 Gene Editing Efficiency. *FEBS Lett.* **2017**, *591* (13), 1892–1901. https://doi.org/10.1002/1873-3468.12707.
(95) Zhang, S.; Voigt, C. A. Engineered dCas9 with Reduced Toxicity in Bacteria: Implications for Genetic Circuit Design. *Nucleic Acids Res.* **2018**, *46* (20), 11115–11125. https://doi.org/10.1093/nar/gky884.
(96) Van Hove, B.; De Wannemaeker, L.; Missiaen, I.; Maertens, J.; De Mey, M. Taming CRISPRi: Dynamic Range Tuning through Guide RNA Diversion. *New Biotechnol.* **2023**, *77*, 50–57. https://doi.org/10.1016/j.nbt.2023.07.001.
(97) Fontana, J.; Sparkman-Yager, D.; Faulkner, I.; Cardiff, R.; Kiattisewee, C.; Walls, A.; Primo, T. G.; Kinnunen, P. C.; Martin, H. G.; Zalatan, J. G.; Carothers, J. M. Guide RNA Structure Design Enables Combinatorial CRISPRa Programs for Biosynthetic Profiling. bioRxiv November 17, 2023, p 2023.11.17.567465. https://doi.org/10.1101/2023.11.17.567465.
(98) Moreno-Mateos, M. A.; Vejnar, C. E.; Beaudoin, J.-D.; Fernandez, J. P.; Mis, E. K.; Khokha, M. K.; Giraldez, A. J. CRISPRscan: Designing Highly Efficient sgRNAs for CRISPR-Cas9 Targeting in Vivo. *Nat. Methods* **2015**, *12* (10), 982–988. https://doi.org/10.1038/nmeth.3543.
(99) Ding, W.; Zhang, Y.; Shi, S. Development and Application of CRISPR/Cas in Microbial Biotechnology. *Front. Bioeng. Biotechnol.* **2020**, *8*, 711. https://doi.org/10.3389/fbioe.2020.00711.
(100) Jost, M.; Santos, D. A.; Saunders, R. A.; Horlbeck, M. A.; Hawkins, J. S.; Scaria, S. M.; Norman, T. M.; Hussmann, J. A.; Liem, C. R.; Gross, C. A.; Weissman, J. S. Titrating Gene Expression Using Libraries of Systematically Attenuated CRISPR Guide RNAs. *Nat. Biotechnol.* **2020**, *38* (3), 355–364.

https://doi.org/10.1038/s41587-019-0387-5.
(101) Guo, J.; Wang, T.; Guan, C.; Liu, B.; Luo, C.; Xie, Z.; Zhang, C.; Xing, X.-H. Improved sgRNA Design in Bacteria via Genome-Wide Activity Profiling. *Nucleic Acids Res.* **2018**, *46* (14), 7052–7069. https://doi.org/10.1093/nar/gky572.
(102) Cui, L.; Bikard, D. Consequences of Cas9 Cleavage in the Chromosome of Escherichia Coli. *Nucleic Acids Res.* **2016**, *44* (9), 4243–4251. https://doi.org/10.1093/nar/gkw223.
(103) Huang, H.-H.; Bellato, M.; Qian, Y.; Cárdenas, P.; Pasotti, L.; Magni, P.; Del Vecchio, D. dCas9 Regulator to Neutralize Competition in CRISPRi Circuits. *Nat. Commun.* **2021**, *12* (1), 1692. https://doi.org/10.1038/s41467-021-21772-6.
(104) Chen, P.-Y.; Qian, Y.; Del Vecchio, D. A Model for Resource Competition in CRISPR-Mediated Gene Repression. In *2018 IEEE Conference on Decision and Control (CDC)*; 2018; pp 4333–4338. https://doi.org/10.1109/CDC.2018.8619016.
(105) Clamons, S.; Murray, R. *Modeling Predicts That CRISPR-Based Activators, Unlike CRISPR-Based Repressors, Scale Well with Increasing gRNA Competition and dCas9 Bottlenecking*; preprint; Synthetic Biology, 2019. https://doi.org/10.1101/719278.
(106) Manoj, K.; Del Vecchio, D. Emergent Interactions Due to Resource Competition in CRISPR-Mediated Genetic Activation Circuits. In *2022 IEEE 61st Conference on Decision and Control (CDC)*; IEEE: Cancun, Mexico, 2022; pp 1300–1305. https://doi.org/10.1109/CDC51059.2022.9993376.
(107) Barbier, I.; Kusumawardhani, H.; Chauhan, L.; Harlapur, P. V.; Jolly, M. K.; Schaerli, Y. Synthetic Gene Circuits Combining CRISPR Interference and CRISPR Activation in E. Coli: Importance of Equal Guide RNA Binding Affinities to Avoid Context-Dependent Effects. *ACS Synth. Biol.* **2023**, *12* (10), 3064–3071. https://doi.org/10.1021/acssynbio.3c00375.
(108) Fang, L.; Fan, J.; Luo, S.; Chen, Y.; Wang, C.; Cao, Y.; Song, H. Genome-Scale Target Identification in Escherichia Coli for High-Titer Production of Free Fatty Acids. *Nat. Commun.* **2021**, *12* (1), 4976. https://doi.org/10.1038/s41467-021-25243-w.
(109) Liu, R.; Bassalo, M. C.; Zeitoun, R. I.; Gill, R. T. Genome Scale Engineering Techniques for Metabolic Engineering. *Metab. Eng.* **2015**, *32*, 143–154. https://doi.org/10.1016/j.ymben.2015.09.013.
(110) Gaj, T.; Gersbach, C. A.; Barbas, C. F., III. ZFN, TALEN, and CRISPR/Cas-Based Methods for Genome Engineering. *Trends Biotechnol.* **2013**, *31* (7), 397–405. https://doi.org/10.1016/j.tibtech.2013.04.004.
(111) Antonakoudis, A.; Barbosa, R.; Kotidis, P.; Kontoravdi, C. The Era of Big Data: Genome-Scale Modelling Meets Machine Learning. *Comput. Struct. Biotechnol. J.* **2020**, *18*, 3287–3300. https://doi.org/10.1016/j.csbj.2020.10.011.
(112) Wessels, H.-H.; Stirn, A.; Méndez-Mancilla, A.; Kim, E. J.; Hart, S. K.; Knowles, D. A.; Sanjana, N. E. Prediction of On-Target and off-Target Activity of CRISPR–Cas13d Guide RNAs Using Deep Learning. *Nat. Biotechnol.* **2023**, 1–10. https://doi.org/10.1038/s41587-023-01830-8.
(113) de Bakker, V.; Liu, X.; Bravo, A. M.; Veening, J.-W. CRISPRi-Seq for Genome-Wide Fitness Quantification in Bacteria. *Nat. Protoc.* **2022**, *17* (2), 252–281. https://doi.org/10.1038/s41596-021-00639-6.

(114) McGlincy, N. J.; Meacham, Z. A.; Reynaud, K. K.; Muller, R.; Baum, R.; Ingolia, N. T. A Genome-Scale CRISPR Interference Guide Library Enables Comprehensive Phenotypic Profiling in Yeast. *BMC Genomics* **2021**, *22* (1), 205. https://doi.org/10.1186/s12864-021-07518-0.
(115) Mormino, M.; Lenitz, I.; Siewers, V.; Nygård, Y. Identification of Acetic Acid Sensitive Strains through Biosensor-Based Screening of a Saccharomyces Cerevisiae CRISPRi Library. *Microb. Cell Factories* **2022**, *21* (1), 214. https://doi.org/10.1186/s12934-022-01938-7.
(116) Wang, J.; Li, C.; Jiang, T.; Yan, Y. Biosensor-Assisted Titratable CRISPRi High-Throughput (BATCH) Screening for over-Production Phenotypes. *Metab. Eng.* **2023**, *75*, 58–67. https://doi.org/10.1016/j.ymben.2022.11.004.
(117) St. John, P. C.; Bomble, Y. J. Approaches to Computational Strain Design in the Multiomics Era. *Front. Microbiol.* **2019**, *10*. https://doi.org/10.3389/fmicb.2019.00597.
(118) Kaczmarek, J. A.; Prather, K. L. J. Effective Use of Biosensors for High-Throughput Library Screening for Metabolite Production. *J. Ind. Microbiol. Biotechnol.* **2021**, *48* (9–10), kuab049. https://doi.org/10.1093/jimb/kuab049.
(119) van Niekerk, D. D.; du Toit, F.; Green, K.; Palm, D.; Snoep, J. L. A Detailed Kinetic Model of Glycolysis in Plasmodium Falciparum-Infected Red Blood Cells for Antimalarial Drug Target Identification. *J. Biol. Chem.* **2023**, *299* (9), 105111. https://doi.org/10.1016/j.jbc.2023.105111.
(120) Millard, P.; Smallbone, K.; Mendes, P. Metabolic Regulation Is Sufficient for Global and Robust Coordination of Glucose Uptake, Catabolism, Energy Production and Growth in Escherichia Coli. *PLOS Comput. Biol.* **2017**, *13* (2), e1005396. https://doi.org/10.1371/journal.pcbi.1005396.
(121) Shin, J.; Porubsky, V.; Carothers, J.; Sauro, H. M. Standards, Dissemination, and Best Practices in Systems Biology. *Curr. Opin. Biotechnol.* **2023**, *81*, 102922. https://doi.org/10.1016/j.copbio.2023.102922.
(122) John, P. C. S.; Strutz, J.; Broadbelt, L. J.; Tyo, K. E. J.; Bomble, Y. J. Bayesian Inference of Metabolic Kinetics from Genome-Scale Multiomics Data. *PLOS Comput. Biol.* **2019**, *15* (11), e1007424. https://doi.org/10.1371/journal.pcbi.1007424.
(123) McNaughton, A. D.; Bredeweg, E. L.; Manzer, J.; Zucker, J.; Munoz Munoz, N.; Burnet, M. C.; Nakayasu, E. S.; Pomraning, K. R.; Merkley, E. D.; Dai, Z.; Chrisler, W. B.; Baker, S. E.; St. John, P. C.; Kumar, N. Bayesian Inference for Integrating Yarrowia Lipolytica Multiomics Datasets with Metabolic Modeling. *ACS Synth. Biol.* **2021**, *10* (11), 2968–2981. https://doi.org/10.1021/acssynbio.1c00267.