A Perspective on Synthetic Biology in Drug Discovery and Development - Current Impact and Future Opportunities

Downloaded from: https://research.chalmers.se, 2026-11-10 06:23 UTC

Citation for the original published paper (version of record):
David, F., Davis, A., Gossing, M. et al (2021). A Perspective on Synthetic Biology in Drug Discovery and Development - Current Impact and
Future Opportunities. SLAS Discovery, 26(5): 581-603.
http://dx.doi.org/10.1177/24725552211000669

N.B. When citing this work, cite the original published paper.

research.chalmers.se offers the possibility of retrieving research publications produced at Chalmers University of Technology. It covers all kind of research output: articles, dissertations, conference papers, reports etc. since 2004. research.chalmers.se is administrated and maintained by Chalmers Library

(article starts on next page)

SLAS Discovery
2021, Vol. 26(5) 581–603
© Society for Laboratory Automation and Screening 2021
DOI: 10.1177/24725552211000669
journals.sagepub.com/home/jbx
SAGE

A Perspective on Synthetic Biology in Drug Discovery and Development—Current Impact and Future Opportunities

Florian David[1], Andrew M. Davis[2], Michael Gossing[3], Martin A. Hayes[3], Elvira Romero[3], Louis H. Scott[3], and Mark J. Wigglesworth[4]

Abstract

The global impact of synthetic biology has been accelerating, because of the plummeting cost of DNA synthesis, advances in genetic engineering, growing understanding of genome organization, and explosion in data science. However, much of the discipline's application in the pharmaceutical industry remains enigmatic. In this review, we highlight recent examples of the impact of synthetic biology on target validation, assay development, hit finding, lead optimization, and chemical synthesis, through to the development of cellular therapeutics. We also highlight the availability of tools and technologies driving the discipline. Synthetic biology is certainly impacting all stages of drug discovery and development, and the recognition of the discipline's contribution can further enhance the opportunities for the drug discovery and development value chain.

Keywords

biofoundry, chassis organism, CRISPR-Cas9, drug discovery, late-stage functionalization, synthetic biology

Introduction

Synthetic biology is the design and construction of new biological entities such as enzymes, circuits, modules, or systems, or the redesign of existing biological systems through reprogramming of genetic information for useful purposes. It has been developing since the 1960s,[1] but has expanded rapidly in the last 10 years, not least because of the plummeting cost of DNA synthesis, advances in genetic engineering, and a better understanding of genome organization and data science. Yet much of what surrounds this enigmatic and exciting field has been part of molecular biology for as long as we can remember. What distinguishes synthetic biology from traditional molecular and cellular biology is the focus on the design and construction of core components that can be modeled, understood, and optimized to meet specific performance criteria, and the assembly of these smaller parts and devices into larger integrated systems to solve specific problems.[2] So what makes this synthetic biology different, and what makes the time right for this field to promise so much to the pharmaceutical industry? In this review, we aim to capture some of the most tangible and promising areas for synthetic biology for drug discovery and development. We define how this blossoming field has already impacted the pharmaceutical industry and describe how it can further enhance the drug discovery value chain.

Speed, cost, and quality have long been the fundamental elements that control the productivity of the pharmaceutical industry. With a focus on improving the efficiency of drug discovery, all three are critical elements that have been incrementally improved by applying tools from synthetic biology. However, game-changing acceleration of drug discovery requires radically different approaches to the classical laboratory-based design-make-test cycles that underpin the iterative design of compounds. While the application of artificial intelligence and automated chemistry has strong support within the pharmaceutical

[1]Department of Biology and Biological Engineering, Division of Systems and Synthetic Biology, Chalmers University of Technology, Gothenburg, Sweden
[2]Discovery Sciences, Biopharmaceutical R&D, AstraZeneca, Cambridge, UK
[3]Discovery Sciences, Biopharmaceuticals R&D, AstraZeneca, Gothenburg, Sweden
[4]Discovery Sciences, Biopharmaceuticals R&D, AstraZeneca, Alderley Park, UK

Received Sept 28, 2020, and in revised form Jan 4, 2021. Accepted for publication Jan 29, 2021.

Corresponding Author:
Andrew M. Davis, Discovery Sciences, Biopharmaceuticals R&D, AstraZeneca, 310 Milton Road, Milton, Cambridge CB4 0WG United Kingdom.
Email: Andy.Davis@astrazeneca.com

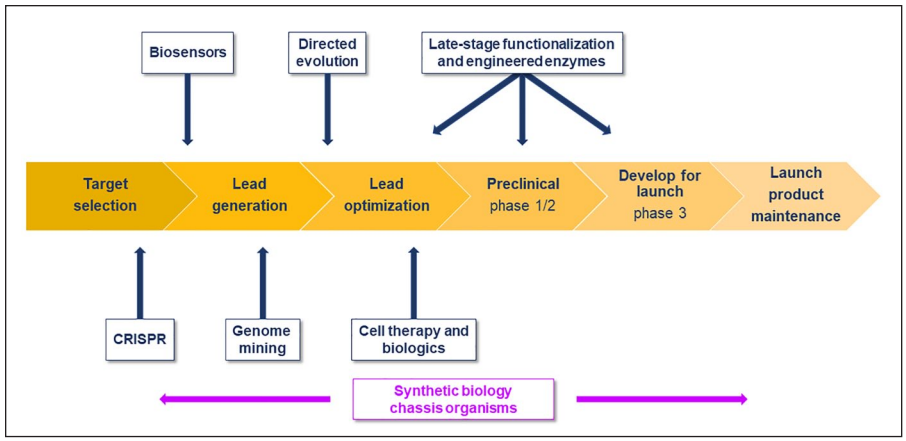

Figure 1. Synthetic biology contributions aligned to the drug discovery process.

industry,[3,4] the potential application of synthetic biology beyond tool-based advances is relatively unappreciated and without specific acknowledgment. It is not just in the pharmaceutical industry that its impact has been underappreciated. A recent *Economist* article highlighted that the money made from organisms that had been genetically engineered underpinned 2% of the U.S. gross domestic product in 2017 ($388B) and is growing, from three sectors, pharmaceuticals ($137B), crops ($104B), and, less visible but even more lucrative, industrial biotechnology ($147B).[5] There is a rapid growth in the new synthetic biology industry, with start-up companies receiving ~$6.1B investment since 2015.[6] The investment is accelerating, with *Forbes* reporting that synthetic biology start-ups received circa $3B in just the first 6 months of 2020.[7]

There are numerous reviews and journal articles that reference the potential of synthetic biology to impact on drug discovery and produce new medicines.[8,9] However, there are relatively few examples where they have been able to specifically describe how synthetic biology has already changed drug discovery. It is the authors' opinion that many pharmaceutical scientists would not be able to identify a single synthetic biology influence on novel medicines, even though the directed evolution of therapeutic monoclonal antibodies has changed the face of modern drug therapy. In contrast to the minimal recognition within the pharmaceutical industry, perhaps two of the most exciting fields for synthetic biology, cell therapy and genetic reprogramming, have been combined within immuno-oncology to produce one of the most novel medical approaches, using chimeric antigen receptor T cells (CAR-T cells).[10]

In this review, we align some of the exciting applications of synthetic biology to the drug discovery value chain (**Fig. 1**), beginning with the toolkit of synthetic biology and the design-make-test-learn cycle, akin to the design-make-test-analyze cycle of medicinal chemistry. In each section, we also highlight the most disruptive innovative pieces of work, which, while a subjective choice, can at least serve as a next step for readers who wish to delve a little deeper into one or the other of the sections.

The Design-Build-Test-Learn Cycle in Synthetic Biology

In recent years, we have seen major advances in the development of robust synthetic biology chassis organisms, the design of biological circuits, and high-throughput screening (HTS) technologies beginning to speed up modern drug discovery cycles. Substantial drops in DNA sequencing and synthesis costs have led synthetic biology to make major steps forward, for example, enabling the generation of DNA-encoded compound libraries,[11] genome mining for novel biosynthetic gene clusters (BGCs) for natural product discovery,[12] and advancing research toward the construction of programmable living therapeutics.[13]

Synthetic biology chassis organisms like *Escherichia coli* and *Saccharomyces cerevisiae* play a key role in this field, as they are very well characterized, facilitate robust applications, and are cellularly reprogrammable. Modular design principles are important to easily exchange parts and to enable rapid generation and characterization of genetic devices and systems. In addition, an orthogonal mode of action is essential to separate biosynthetic functions from the chassis cell's own intrinsic regulatory networks whose behavior is difficult to anticipate and might create background noise due to the environmental stimuli or cell cycle progression.

The concept of design-build-test-learn (DBTL) cycles is at the heart of synthetic biology applications. It enables efficient screening and optimization for desired functions of biosynthetic devices and systems of interest, ranging from proof-of-concept studies to advanced drug discovery screens and genetic circuit design in living therapeutics.

The design step relies on well-characterized biological parts and computer-aided approaches. These parts comprise means to specifically control gene expression or translation, including promoters,[14] terminators,[15] variation in codon usage,[16] ribosome binding sites (RBSs),[17] ribozymes,[18] and

protein degradation tags.[19] We see an increasing number of synthetic promoters developed, allowing for orthogonal function and robust fine-tuning of gene expression.[20] Strategies employed include the combination of a core promoter region with binding sites for heterologous/hybrid transcription factors[21] or Cas9 coupled to transcriptional regulators.[22,23] Also, inducible control of expression through external stimuli is often essential to have a direct control input to the system.[24,25] The mentioned tools are used to fine-tune and characterize the behavior of genetic devices, including activators, repressors, reporters, and environment or drug/metabolite-responsive biosensors. Computational tools like AutoBioCAD,[26] Cello, and the Synthetic Biology Open Language (SBOL)[27] were developed to streamline this design step and facilitate laboratory automation and public reporting of data.[28]

In the build phase, efficient DNA assembly tools, such as the Gibson-based assembly,[29] ligase cycling reaction,[30] BioBrick assembly (BASIC),[31] uracil-specific excision reagent (USER),[32,33] and Golden Gate assembly,[34,35] are used to ease and speed up the combinatorial assembly of parts and construction of combinatorial libraries.

The test phase relies on high-throughput analytics employing parallel cultivation platforms in microwell format[36] or microfluidic droplets[37] and the monitoring of reporter outputs like fluorescent proteins or liquid chromatography–mass spectrometry (LC-MS) analytics for compound detection. To further increase throughput, we might see methods like automated laser-assisted rapid evaporative ionization MS (LA-REIMS) used more in the future. For example, in a recent approach it was successfully used to perform rapid MS direct from agar plate yeast colonies without sample preparation or extraction.[38]

In the learn step, various in silico modeling approaches are employed for simulation and optimization. This includes the usage of differential equations and machine learning algorithms as, for instance, recently shown for the rational tuning of G-protein-coupled receptor (GPCR) signaling,[39] engineering of synthetic gene networks in yeast,[40] or model-guided design of artificial yeast promoters.[41] Design of experiment (DoE) tools explore the multidimensional experimental space allowing for the modeling and optimization, for example lately employed to optimize the dynamic and operational range of a metabolite-responsive biosensor.[42]

A main future driver within this cycle is the increasing application of automation and software development, integrating the different steps of the DBTL cycle in an industrial-like pipeline and enabling the characterization of thousands of parts and biosynthetic designs.[43] The concept was termed biofoundry (BioFAB) and has recently led to the establishment of a Global Biofoundry Alliance[44] coordinating activities in the field worldwide.

Synthetic Biology Chassis Organisms

Synthetic biology chassis organisms are key to rapid engineering and testing cycles to speed up drug discovery and optimize production. Detailed knowledge about the genome, function of proteins, and predictability and modeling of metabolism enable targeted genetic modifications to create, for example, platform strains for the production of certain molecule classes, construction of compound libraries, and implementation of biology circuits to screen for new drug leads. Synthetic biology tools as described above are well established, making the implemented systems easier to study and tune more efficiently. In the following, we summarize the role of the two main synthetic biology chassis and their role in the drug discovery pipeline.

The Prokaryotic Synthetic Biology Chassis: Escherichia coli

The gram-negative bacterium *E. coli* is the most studied in detail prokaryotic model organism. With its 4.6 Mbp genome, 4288 protein-coding genes, and 85% of the operons with proven function,[45] it is very well characterized and used in many applications within synthetic biology and natural product production. *E. coli* was successfully genetically engineered to produce complex natural products, including terpenoids, polyketides, phenylpropanoids, and alkaloids, as recently reviewed elsewhere.[46] To name a few, production was enabled for complex molecules like the antibiotic valinomycin,[47,48] anticancer drug taxol precursor taxadiene,[49] and morphine precursor reticuline[50] (**Fig. 2**).

E. coli cells are an integral part of several high-throughput drug screening platforms. The industrially most relevant is the 2018 Nobel Prize-winning phage display technology used to identify and mature specific antigen binding domains for antibody engineering.[51] Libraries of antibody variable domains are expressed and presented on *E. coli*-produced bacteriophages that are used for panning rounds to identify superior antigen binders. In a different method, named phage-assisted continuous evolution (PACE),[52] the phage replication is dependent on the production of effective protein binders enabling the continuous in vivo evolution and selection of potent protein inhibitors. In another approach, *E. coli* was shown to be a suitable host to express cyclic peptide libraries and screen these for specific protein–protein interaction (PPI) inhibitors in vivo.[53,54]

Synthetic biology-engineered *E. coli* are being developed as cell therapies directly, as will be described in the Cell Therapy and Biologics section. Besides many examples in the preclinical stage,[55] we already see the first therapeutic bacteria that have entered phase I and II clinical trials for various indications, including hyperammonemia,[56,57] phenylketonuria,[58,59] and oral mucositis.[60] Despite holding

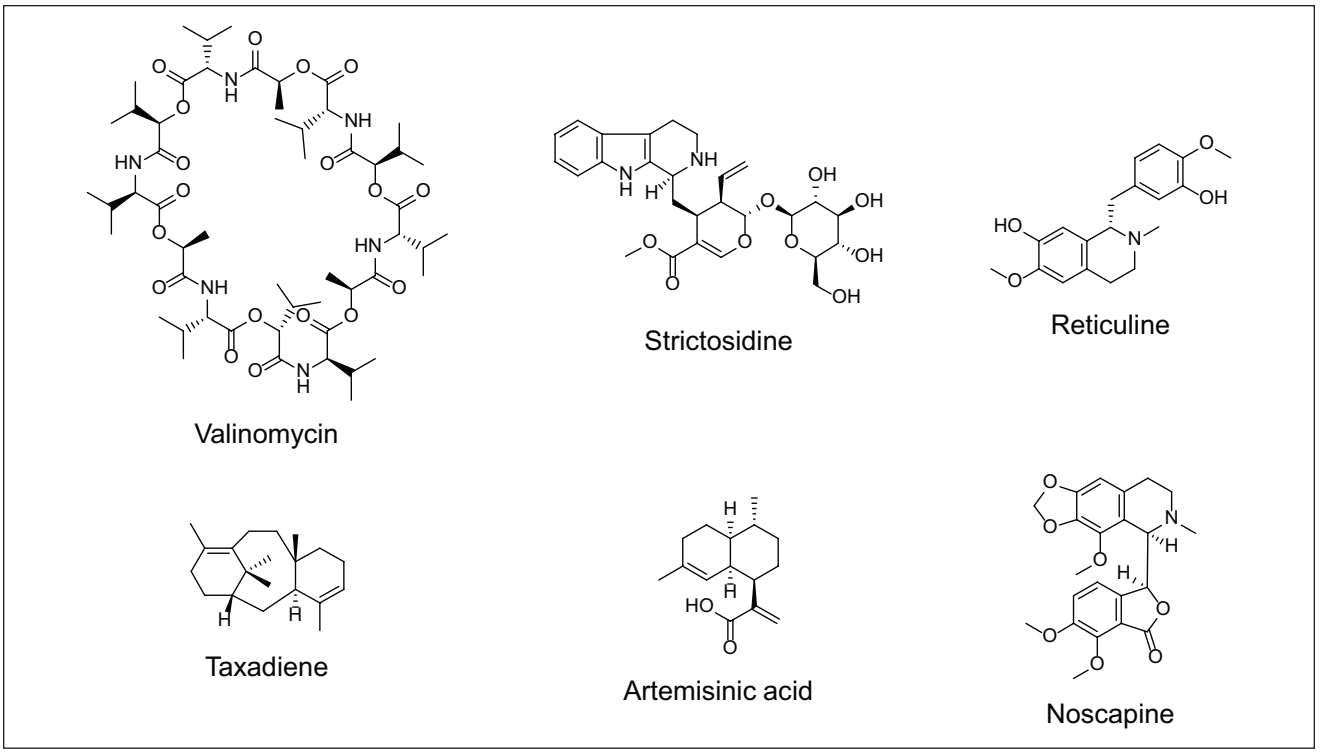

Figure 2. Structures of natural products synthesized in engineered *E. coli* or *S. cerevisiae*.

great promise for many future applications, safety, regulatory, and public concerns have to be addressed before these new types of therapeutics become commercially available.

Although used in many studies and with a proven track record for drug screening approaches, *E. coli* has certain limitations when it comes to advanced chemical modifications of products like glycosylation of antibodies and monooxygenations of natural products. Also, looking at its potential for commercial drug production, *E. coli*-based bioprocesses have the disadvantage of being prone to phage contamination and come with increased purification costs due to the pyrogenic nature of its endotoxic membrane structures.

The Eukaryotic Synthetic Biology Chassis: Saccharomyces cerevisiae

Baker's yeast *S. cerevisiae* is a very well-characterized eukaryotic model organism. It harbors 16 chromosomes, comprising a genome of 12 Mbp, with 6275 genes coding for 6049 proteins, where 90% have a characterized function.[61] *S. cerevisiae* is used for the commercial production of several drugs, including insulin,[62] vaccines,[63] and the antimalaria drug artemisinin.[64] Besides, *S. cerevisiae* is an established disease model organism to study, for example, aging,[65] neurodegenerative diseases,[65] drug toxicity,[66] and cancer.[67] For example, yeast single-gene deletion libraries, comprising more than 4000 strains,[68] were used to find synthetic lethal interactions using large-scale primary interaction screens, revealing insights into new combinatorial cancer therapies.[69]

A prime synthetic biology undertaking is the ongoing Yeast 2.0 project aiming to resynthesize its entire genome. So far, seven chromosomes[70] have been synthesized, including specific modifications allowing for so-called genome scrambling (Synthetic Chromosome Recombination and Modification by LoxP-mediated Evolution [SCRaMbLE]). Through induction of a heterologous Cre recombinase, the genome can be cut and reassembled to easily create a large genotypic diversity. These can be screened for advanced phenotypes, for example, enabling better drug production,[71] as recently shown for heterologous production of penicillin.[72]

In the field of drug screening, *S. cerevisiae* has successfully been used to find new antibody variants employing the yeast surface display technology. Developed in 1997 by Boder and Wittrup,[73] it has become an attractive eukaryotic alternative to the above-mentioned phage display. Yeast has the advantage to be able to efficiently secrete and display proteins and can be engineered to conduct posttranslational modifications like glycosylation similar to mammalian systems.[74–77] In contrast to bacterial hosts like *E. coli*, *S. cerevisiae* has intracellular organelle systems, including extended membrane structures like the endoplasmic reticulum and Golgi apparatus. These are important to functionally express certain enzymes, like cytochrome P450 monooxygenases, which are essential for the biosynthesis of many natural products of high structural diversity. A prominent example is the production of the antimalaria compound artemisinic acid (**Fig. 2**), where it was essential to have a functional, efficiently expressed P450 enzyme present to boost

production by 25-fold to 25 g/L compared with the system used in *E. coli*.[78,79]

Recent impressive examples demonstrating the power of yeast as a synthetic biology chassis are the production of various plant-derived complex natural products.[80,81] These include monoterpene indole alkaloids such as strictosidine, a precursor of the potent chemotherapeutic alkaloid vinblastine,[82,83] and opioids such as noscapine[84,85] and cannabinoids[86] (**Fig. 2**). Advanced metabolic engineering approaches and implementation of up to 50 different genes from bacteria, mammals, and plants were necessary to establish production in these cases.

Besides the above-described classic synthetic biology chassis, other nonconventional chassis organisms can be more beneficial for the production of other compound classes.[87] This includes, for instance, *Bacillus subtilis*,[88] used due to its high secretion capacities for the production of enzymes and certain nutraceuticals; lactic acid bacteria for their application as living therapeutics;[89] the yeast *Komagataella phaffii* (formerly *Pichia pastoris*) for therapeutic protein production;[90] and *Streptomyces*, employed for its ability to produce high amounts of natural products, including polyketides, nonribosomal peptides (NRPs), and terpenes.[91]

Choosing the right chassis organism is dependent on various parameters, including its final application, the compound class in focus, the importance and availability of synthetic biology tools, screening libraries, or even considerations toward the final industrial production of the compounds of interest.

Disruptive Science

The assembly methods[30–35] have accelerated the assembly of parts to increase the speed of the DBTL cycle and, together with automation and software development, as described in Chao et al.,[43] are set to be a main driver for future growth of synthetic biology.

Target Validation

One of the biggest reasons why drug discovery projects fail remains the identification and validation of targets.[92] Paradoxically, where targets are validated in the clinic following phenotypic or target-based approaches, second-generation medicines often build on this success.[93] Furthermore, where careful consideration is taken to validate targets and follow rigorous scientific principles, as exemplified by AstraZeneca 5R's framework, success rates can be dramatically improved.[94] Hence, discovering novel targets and validating them is a vital part of establishing a successful pipeline and ensuring future medical advances. It therefore comes as no surprise that there has been significant interest and investment in tools that can help us achieve this.

The use of CRISPR-Cas9[95,96] tools fits perfectly into this arena. With the ability to exquisitely modify genetic sequences, this technology has allowed the generation of several new and important components of target validation that are poised to offer significant advantages in selecting the right targets.

At a very simple level, CRISPR-Cas9 has been used successfully in pharmaceutical research for several years to generate precise cell models replacing historic molecular biology approaches and introducing the ability to modulate gene function by ablation, downregulation, upregulation, or mutation, across cell types. Where previous technologies based on primer-driven mutagenesis, PCR, and subcloning or technologies such as siRNA screening could achieve many of these, they were laborious, were time-consuming, or could not be completed with the same level of confidence in the resulting cell population. However, in terms of target validation, CRISPR-Cas9 is poised to have a significant impact, at least in part due to parallel advances in complex cell systems. Although challenges remain in the cell types available and the cost of these techniques, the use of primary cells in these settings is growing, and the ability to sustain and differentiate stem cells presents an opportunity to modulate and explore disease in a way that previously would not have been possible. For example, recent reports include use of CRISPR-Cas9 in primary T cells to identify factors involved in HIV infection and pathogenesis,[97] while Martufi et al.[98] and Borestrom et al.[99] have used primary human fibroblasts and stem cell-derived kidney cells to probe genes involved in disease. This means that the cells that are now being used in validation experiments are closer to patient cells than ever before; hence, when interesting associations with disease are made, there is a higher confidence that these observations will translate into real effects in the clinic, if this target can be modulated with a drug.

Furthermore, CRISPR-Cas9 offers the ability to produce genome-wide screening sets that can ablate a gene function, which present very powerful tools that can be used to rapidly examine the role these genes play in different cell models. Pooled CRISPR-Cas9 screens whereby a population of cells are modified with a library of CRISPRs and then placed under a selection/growth condition, followed by next-generation sequencing (NGS) to determine which genes are enriched (resistant) or absent from a population (detrimental), have been used extensively over the last few years to implicate target validation.[100–103] With arrayed CRISPR libraries, individual genes are targeted (usually with several different guide RNAs targeting the same gene) in isolated cell populations. Hence, it is easy to establish the direct effect or lack of effect each gene has within the array, providing higher confidence in the data output, but at a cost of requiring more complex platforms to run these at scale. Pooled screens also require fewer cells compared with array-based CRISPR and siRNA screening.

Shifrut et al.[104] nicely exemplify the combination of pooled CRISPR screening with primary human T cells, overcoming the issue of low lentiviral transduction rates for vectors carrying Cas9 in primary cells by introducing a new

approach using single-guide RNA lentiviral infection with Cas9 protein electroporation (SLICE). Raising successful targeting to more than 80% of cells with a simplified protocol allowed Shifrut et al. to conduct large-scale genome-wide screening with increased confidence that the genes within the pool would be ablated in their experiment. Further, they were able to set defined conditions to identify genes that regulate T-cell proliferation in response to T-cell receptor (TCR) stimulation, in replicate screens and at differing doses of TCR stimulation, further increasing confidence in the targets identified. Finally, they combined their study with single-cell RNAseq to map pathways regulated by these genes and were able to repeat the screening in the presence of adenosine to test for targets that can allow the T cells to escape immunosuppression within adenosine-releasing cancers. These types of study have long been the promise of CRISPR as a synthetic biology tool, but this study is one that enables a paradigm shift in the ability to work with these types of complex cells with high confidence in the quality and completeness of the data generated.

To date, array-based screens have largely been used for smaller, more directed gene modulation,[105–108] and at the time of writing, there were approximately five times the number of references for pooled screens compared with arrayed CRISPR in PubMed.

Yet, the greatest opportunity for these tools is to come. Without exception, large pharmaceutical companies are now investing in new genomic initiatives. Hailing a genomic era of discovery, patient NGS data and artificial intelligence are poised to identify many new targets across mainstream and rare diseases.[109] These data, when combined with powerful CRISPR arrays, will facilitate rapid confirmation of target association[110,111] in cell, patient, and animal studies, and provide the foundations for new and improved medicines.

The ability to exquisitely control genes and connect them together to form functional circuits with desired outputs is not a new phenomenon. Classical molecular biology tools have been used historically to create impressive biological circuits in bacteria,[112,113] yeast,[114] and mammalian[115] systems. The flexibility and fidelity of CRISPR-Cas9 means that these same things can be done in fractions of time and form more complex architectures.[116] Some may argue that CRISPR is not a product of synthetic biology, but rather a molecular biology tool. However, the high level of design, the use of standardized components, and the precise execution and engineering involved in the design of CRISPR tools certainly fit the definition provided in the introduction, and CRISPR is included by the European Union as a synbio tool.[117,118] Through coupling a Cas9 with a designed RNA guide, CRISPR is a synthetic biology circuit all by itself. PRIME editors, consisting of a catalytically impaired Cas9 endonuclease fused to an engineered reverse transcriptase enzyme, and a prime editing guide RNA (pegRNA) offer a bespoke search-and-replace system for specific insertion, deletions and base-to-base swaps, with drug-inducible Cas9 allowing temporal control of gene editing for in vitro and in vivo applications at any desired point in the genome.[119]

Disruptive Science

CRISPR-Cas9 [95,96] is revolutionizing the way we do everything, from target validation to animal model development to the way we think about new therapeutics. It is not a surprise that the 2020 Nobel Prize for Chemistry was awarded to Emmanuelle Charpentier and Jennifer Doudna for the discovery of CRISPR-Cas9 genetic scissors.

Assay Development: Biosensors and Genetic Selections in Screening

Using synthetic biology to reprogram cells in an orthogonal and predictable manner offers many solutions to the drug discovery process. Previous efforts have already seen the creation of genetic circuits that can follow digital logic, are dynamic, or can even mimic electronic devices.[120–122] Such circuits provide screenable or selectable outputs for diverse inputs, such as small molecules or metabolites, and can be incorporated into the drug discovery process as bioassays and biosensors for phenotypic screening. As more biological parts are characterized, the complexity of synthetic cellular systems will increase, and with that, so will their computational power. The potential translation of this power into the discovery of new and better drugs is promising, with examples and prospects outlined in this section.

Plasma membrane receptors, in particular GPCRs, play a key role in regulating the physiology of virtually every cell.[123] Given their fundamental importance, it is unsurprising that they are the largest family of proteins targeted by prescription drugs.[124] However, discovering hits against receptors remains challenging, as it is difficult to translate drug binding to a selectable or screenable output. Synthetic biology offers some solutions to this problem. Early, crude biosensor efforts took advantage of GPCR signaling cascades that increased cellular Ca^{2+} after drug binding, as this could be detected by the Ca^{2+}-dependent fluorescence of recombinant aequorin.[125] Nevertheless, coupling receptors to a selectable signal transduction pathway, be it endogenous or synthetic, is a better and more direct way to assay receptor–ligand interaction. Barnea et al. were able to artificially rewire mammalian GPCRs, receptor tyrosine kinases, and steroid hormone receptor signaling by tethering transcription factors to receptors by a protease cleavable linker.[126] Ligand binding to the receptor recruits a signaling protein fused to a protease that frees the transcription factor, thus activating reporter genes. The authors were able to successfully identify a ligand for the orphan receptor GPR1. Other research has seen mammalian GPCRs transplanted

into the yeast *S. cerevisiae* to create artificial signal transduction pathways that can then be used as a powerful tool to assess function or identify ligands, or as a pharmaceutical screening assay.[127,128] Rewired receptors in yeast were used to make a bioassay for odorant screening,[129] and to identify peptide agonists against the chemokine receptor CXCR4 important in HIV infection.[130] Recently, Shaw et al. engineered a tunable refactored GPCR signaling pathway in yeast, making a significant advance in how future predictable artificial signal transduction pathways will be applied to drug discovery.[131]

While the extracellular location of receptors makes them easily accessible as drug targets, many causes of disease are from aberrant PPIs that occur within the cell.[132] Synthetic biology approaches have been designed to select or screen for these interactions and monitor their disruption using chimeric proteins. For example, the protein–fragment complementation assay (PCA) functions where two proteins are each fused to a fragment of a third reporter protein; if interaction between the proteins of interest (POIs) occurs, the reporter protein is reconstituted and becomes active.[133] This approach has mapped the effects of small molecules on signal transduction pathways and both their on- and off-target effects,[134,135] as well as identifying agonists of the glucocorticoid receptor.[136] An adaption of the PCA is the Förster resonance energy transfer (FRET) assay, which was used to screen small-molecule libraries for antiviral activity against poliovirus[137] or hepatitis C.[138] The bacterial reverse two-hybrid system (RTHS) is another option to assay PPI. Here, a POI is fused with the DNA binding domain of an obligate dimeric repressor.[139] When the POI fusions dimerize, a functional repressor is formed and binds operator sites engineered to prevent the expression of downstream reporter genes. Inhibitors of PPI can be screened by selecting for reporter gene activation. This strategy was used to find cyclic peptide inhibitors of HIF-1 heterodimerization, AICAR transformylase homodimerization, the HIV Gag-TSG101 interaction, and CtBP involved in cancer.[140–143]

Naturally evolved transcriptional repressors are often promiscuous allosteric regulators of bacterial physiology.[144] They are generally modular, with a DNA binding and a sensing domain; they are ordinarily used by bacteria to respond to other small molecules, secondary metabolites, the environment, and the cell cycle, but they are also functionally vital to sense antibiotics and regulate self-resistance.[145] Their modularity and ability to detect medicinally relevant molecules make them important biological tools that have been reprogrammed into synthetic transcription factors to regulate artificial eukaryotic genetic circuits. For example, a genetic switch for the antibiotic streptogramin was constructed in mammalian cells,[146] which led to the discovery of noncytotoxic antibiotics.[147] Furthermore, switches against tetracycline antibiotics[148] have been incorporated with a bacterial resistance mechanism TetX in yeast, thereby providing a genetic circuit to select not only for new tetracyclines but also for those that can evade antibiotic resistance.[149] A rewired *Mycobacterium tuberculosis* transcriptional repressor was used in human cells to screen for nontoxic drugs that increase the bacterium's sensitivity to ethionamide.[150] Not only are bacterial repressors targets for drug discovery, but also their reappropriation to control gene expression allowed the construction of an inducible genetic circuit in mammalian cells for the HTS of cytotoxic drugs that preferentially targeted proliferation competent cells that mimic cancer.[151] The future potential to discover new small molecules by transcriptional repressor reprograming is immense, considering that more than 200,000 tetracycline family regulators have been identified, while the majority of their ligands have not been.[152] Furthermore, our ability to redesign synthetic transcription factors against new ligands exponentially increases this trove of possibilities for the use of genetic circuits in drug discovery.[153]

Another interesting avenue for drug discovery is monitoring whether pharmaceuticals modulate the intracellular concentration of metabolites or small molecules. Following this premise, a naturally occurring riboswitch was adapted to control the fluoride-dependent expression of β-galactosidase in bacteria; this allowed the screen of a small-molecule library for fluoride toxicity agonists.[154] Considering synthetic aptamers that bind with high specificity and selectivity to a wide range of targets can be made and isolated,[155] the concept of sensing intracellular metabolites, or of controlling mammalian gene expression by any small inducer molecule, could be widely applicable.[156,157]

Disruptive Science

Barnea et al.[126] developed an assay where GPCRs can be reprogrammed to activate a reporter construct upon ligand binding. This technology provides a quantifiable measure of ligand interaction with a specific GPCR, useful in drug screens and the identification of ligands for orphan receptors.

Weber et al.'s[150] study creates an unnatural genetic circuit in mammalian cells using parts responsible for the persistence of tuberculosis. By combining parts from the host and the disease, this assay can screen drugs against tuberculosis for specificity, bioavailability, and cytotoxicity at the same time.

Hit Generation

While natural products have historically been a highly successful source of drugs,[158,159] the modern drug discovery process has become focused upon the HTS of highly curated compound libraries in the search for leadlike[160] molecules as an easier source of equity for optimization.[161] Current screening practice encompasses the screening of thousands of fragments with x-ray-guided optimization to discover hit

quality leads (micromolar activity), the HTS of 1–2 million compound libraries, and the recent development of DNA-encoded libraries (DELs) to screen libraries of billions of compounds. In such screens, the aforementioned synthetically engineered bioassays are implicit to the hit discovery process. In many ways, the ability to screen very large lead-like and druglike small-molecule libraries to find hits amenable to medicinal chemistry optimization has somewhat eclipsed natural products as drug modality. Historically, therapeutic natural products have been identified from biological samples collected from across the globe and HTS. Natural product drugs continue to be a valuable source of new drugs.[162] However, the screening of natural product extracts no longer drives the drug discovery engines of major pharmaceutical companies, even though many invested heavily in this area over the past 20 years and work continues in this area.[163] The modern fragment screening/HTS/DEL hit discovery paradigm is a very efficient process compared with the complexity of natural product chemical synthesis. Natural products can only compete if either they are already good enough to be the candidate drug or the candidate drug lies a short semisynthetic journey away. However, data science and synthetic biology offer new opportunities for natural products. Natural extract library screening can be replaced by the bioinformatic searching of BGCs to find putative natural products that could be synthesized. Now, DNA sequencing can find BGCs coding for pharmaceutical success in natural product drugs literally under our feet. Charlop-Powers et al. found gene clusters for 11 therapeutically important natural product families historically identified from across the globe, in the soils of Central Park, New York.[164] Less than 1% of the reads were aligned to known BGCs, which suggested a large reservoir of untapped biodiversity in the urban environment.

We are beginning to discover thousands of new BGCs identified through metagenome analysis. These clusters range from several kilobases to >100 kb in length and code for enzymes enabling complex small-molecule synthesis of large diversity. This covers various classes of secondary metabolites, including polyketides, NRPs, terpenes, and ribosomally synthesized and posttranslationally modified peptides (RiPPs).[165] Though the discovery of novel BGCs is progressing rapidly, only a fraction of these BGCs have been characterized and tested for their potential to produce novel drug leads.

Most of the BGCs are derived from nonculturable microorganisms, which makes the heterologous expression in established hosts necessary. As mentioned above, well-characterized synthetic biology chassis organisms have been established, which have a plethora of synthetic biology tools available to optimize and establish production. Gene synthesis enables the expression of codon-optimized versions of the BGC adapted to each host strain, combining these with well-characterized genetic parts to control transcription and translation. This refactoring not only has the advantage to remove regulatory elements, enabling controlled production, but also allows for the generation of combinatorial libraries. For example, it is possible to substitute enzymes with homologous variants to create derivatives or completely random enzyme clusters that generate novel chemical diversity.[166,167] Another important point is the precise control of gene expression of the BGC members. For many BGCs, it was shown that gene expression is necessary to be fine-tuned to achieve optimal production.[168–170] Using genetic BioBricks as described above, many gene cluster variants can be generated reusing a set of regulatory parts. Libraries can be screened for setups with optimized expression levels and increased titers. In addition, we see more platform strains becoming available, providing sufficient precursor and co-factor supply, easing the discovery and production for specific product classes of interest.[171–173]

Synthetic genetic circuits can be used to establish switches to specifically turn on BGC-derived pathway genes, allowing the de-coupling of growth from production.[174] For example, feedforward regulation has been described for many biosynthetic pathways of natural products,[175] allowing for dynamic regulation and the adaption to precursor availability and activating product export. Positive and negative feedback loops can be used to specifically allocate cellular resources to secondary metabolism.[176]

The effort required to express the diversity of natural products encoded in these cryptic BGCs and test them in HTS to identify function in an unguided fashion would be enormous. However, a quirk of genome organization may offer clues to the purpose of the cryptic natural products that can be mined bioinformatically.

Along with the enzyme-coding genes, many BGCs contain other genes not involved in product synthesis, including transcription factors and genes that could be self-protective. These self-protective genes encode drug efflux transporters, detoxifying enzymes, and sometimes even resistant versions of the protein targeted by the BGC.[177] Many BGCs contain more than one of these self-protecting genes. These resistance genes offer a clue as to the function of the encoded natural product.

A search of the genomes of 86 *Salinispora* bacteria for protein-coding genes for lipid transport and metabolism associated with BGCs identified a 22 kb polyketide synthase–nonribosomal peptide synthetase (PKS-NRPS) hybrid gene cluster that, when heterologously expressed, produced thiolactomycin analogs previously shown to be fatty acid synthase II inhibitors.[178] An examination of an uncharacterized *Aspergillus nidulans* BGC identified a putative gene (*inpE*) with no obvious role in natural product synthesis.[179] InpE has homology to the b6 subunit of the

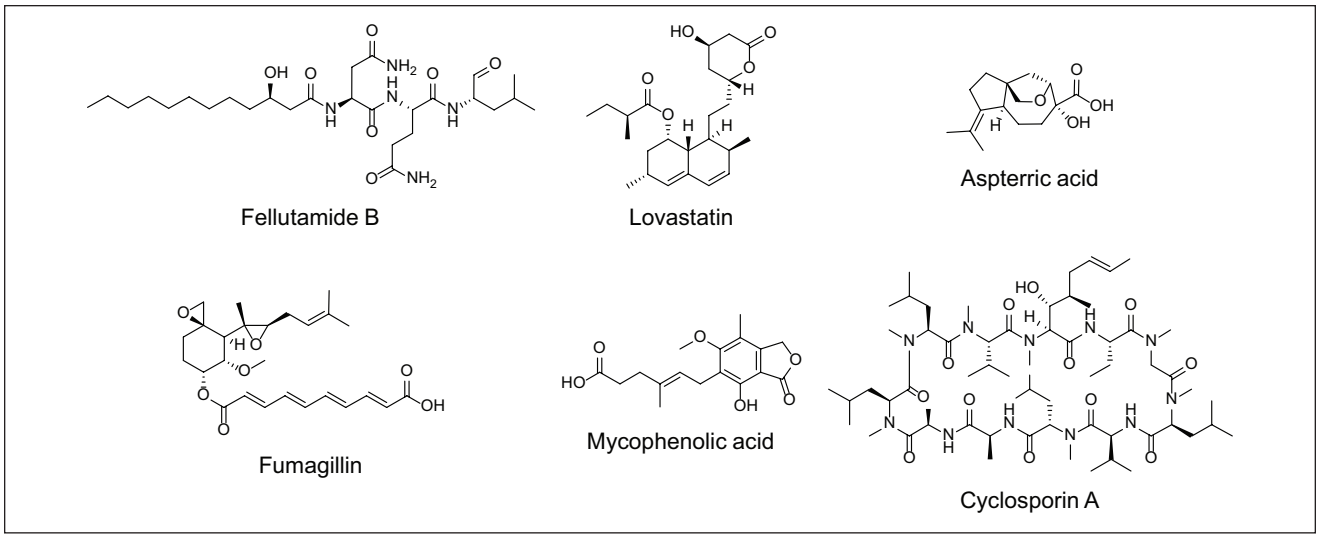

Figure 3. Examples of natural products with in-cluster resistance genes. As suggested by Keller,[177] the identification of other genes in BGCs offers a tactic to reveal the targets of natural products encoded by BGCs, and focus on useful clusters to characterize.

proteasome, suggesting that the product of the *inp* BGC might be a proteasome inhibitor. Through use of a serial promoter exchange approach to sequentially replace six promoters in genes of the cluster with a regulatable promoter, the cluster was successfully expressed and the product of the BGC was identified as fellutamide B (**Fig. 3**). By deleting *inpE* in *A. nidulans* and activating the expression of fellutamide, Hsu-Hua et al.[179] were able to demonstrate that *inpE* is required for resistance of the internally produced fellutamide B.

The fungal metabolite lovastatin blocks cholesterol synthesis by inhibiting 3-hydroxy-3-methylglutaryl-coenzyme A (HMG-CoA) reductase, an enzyme required for the synthesis of fungal ergosterol. To prevent toxicity in the host strain, the *Aspergillus terreus* lovastatin gene cluster encodes a HMG-CoA reductase proposed to be resistant to lovastatin,[180] a feature common in related statin BGCs.[181]

Fumagillin inhibits methionine aminopeptidase 2, required for the removal of N-terminal methionine residues from nascent proteins. The BGC for fumagillin in *Aspergillus fumigatus* contains an additional "in-cluster" methionine aminopeptidase 2 gene, maintained in analogous BGCs in other fungi.[182]

Dihydroxyacid dehydratase (DHAD) is a key enzyme in the production of branched chain amino acids (BCAAs) in plants but not present in animals, making it an excellent target for new herbicide discovery. Although no natural products were known to target DHAD, Yan et al.[183] hypothesized that a fungal natural product might exist, as BCAA biosynthesis is required by plants. So, they went hunting for fungal BGCs containing an additional copy of the DHAD homolog. They identified four-well conserved genes across multiple fungal genomes, encoding for a sesquiterpene cyclase homolog (*astA*), two cytochrome P450 genes (*astB* and *astC*), and a homolog of DHAD (*astD*). Heterologous expression of *astA*, *astB*, and *astC* in *S. cerevisiae* produced the tricyclic aspterric acid, a previously identified natural product with a potent activity against *Arabidopsis thaliana*, through an unknown mode of action. The group showed that aspterric acid was a 0.5 μM inhibitor of *A. thaliana* DHAD, but did not inhibit AstD DHAD even up to its solubility limit of 8 mM. Other identified BGCs containing resistant copies of the target gene include echinocandin, which targets β-1,3-D-glucan synthase;[184] mycophenolic acid, targeting inosine 5′ monophosphate dehydrogenase;[185] and cyclosporin A in the tolypocladium genome, which contains the gene for its molecular target cyclophilin (**Fig. 3**).[186]

Disruptive Science

Smanski et al.[174] provide a good review of the area of resistance genes in filamentous fungi and highlights how identification in BGCs with collocated resistance genes can reveal the targets of natural products encoded by BGCs, and other references[184–187] illustrate how this information can be used as a drug discovery tactic.

Lead Optimization

Directed Evolution and Synthetic Biology

Hit identification and lead generation are followed by lead optimization, a process in which an initial lead compound is subjected to iterative rounds of (chemical) modifications

and characterization to give insight into its structure–activity relationship and metabolic stability. These iterative rounds of modifications are a trait that lead optimization shares with directed evolution.[187] In directed evolution, a genetically encoded molecule (nucleic acids or proteins/peptides) is subjected to iterative rounds of modification/mutagenesis, followed by the selection/screening of a user-defined goal, while ensuring the inheritance of the advantageous characteristic of the evolved molecule (genotype–phenotype linkage). For example, linear peptides targeting G-protein subunit α[188] and cyclic peptides inhibiting HIV protease[189] have evolved by directed evolution.

To further increase the chemical space accessible for lead optimization, one can take a look at natural products and their biosynthesis for inspiration. NRPs are synthesized by NRPSs, and the peptides often contain noncoded amino acids that might be further modified by, for example, *N*-methylation. However, directed evolution of NRPSs presents additional challenges for lead optimization. The correlation between the enzyme primary structure and the structure of the produced peptide is often elusive, as the sequence of the NRP is not genetically encoded. Attempts to diversify natural products by domain swapping or mutation of NRPSs often result in greatly reduced enzymatic activity accompanied by low product titers. Heterologous expression of NRPSs further limits diversity due to the absence of noncognate substrates in the heterologous host. Despite these shortcomings, some successful cases have been reported. Directed evolution was used to restore the functionality of a heterologously expressed AdmK-CytC1 hybrid.[190] This hybrid was also capable of producing new derivatives of andrimid, a peptide antibiotic that inhibits prokaryotic acetyl-CoA carboxylase. Additional derivative compounds of andrimid have been obtained by directed evolution of AdmK and expression in the native host *Pantoea agglomerans*.[191] Recent advances in the design and engineering of NRPs might simplify the diversification of natural peptide products and the generation of novel NRPs in the future.[192]

Unlike NRPs, RiPPs are genetically encoded peptides that are synthesized on ribosomes and further modified by RiPP biosynthetic enzymes. The precursor peptides contain a conserved leader peptide and a variable core peptide, and the latter is posttranslationally modified and proteolytically processed to the mature form. The discovery of novel post-translational modifications, previously thought unique to NRPs, such as *N*-methylation,[193] largely increases the chemical space accessible by RiPPs. Lanthipeptides are a class of macrocyclic RiPP that contain characteristic lanthionine moieties. Phage display and yeast display techniques have been adapted to allow for screening of in vitro-generated libraries of lanthipeptides,[194] enabling directed evolution of this modality in the future. The accessible chemical space can be increased further by using synthetic biology to create nonnatural hybrid RiPPs that contain moieties from different classes of RiPP.[195]

The advantage of genetically encoded molecules is the ease of library generation by means of in vitro mutagenesis, for example, by error-prone PCR (epPCR). However, in vitro-generated libraries that can be screened in vivo often have size limitations due to the transformation efficiency of the host organism, limiting the solution space that can be probed in practice. A simple octapeptide composed of amino acids encoded by standard genetic code already has 2.56×10^{10} possible combinations, and the DEL would have to be even larger than this to account for codon degeneracy. Additionally, the screening of in vitro libraries requires the isolation and identification of beneficial mutations after each iterative round, which is laborious and time-consuming. Continuous directed evolution avoids these limitations by performing all steps of directed evolution—library generation/mutagenesis, selection/screening, and inheritance—in vivo, allowing for a continuous process.

Recent years have seen a surge in method development allowing for continuous directed evolution in a variety of host organisms, such as PACE,[52] in vivo continuous evolution (ICE),[196] CRISPR-AID,[197] orthogonal replication (OrthoRep),[198] and EvolvR.[199] Both classes of the previously described peptides, NRPs and RiPPs, are promising lead compounds for drug development.[200]

Disruptive Science

With the growing number of continuous directed evolution techniques and the improving toolbox provided by synthetic biology, these techniques, as described in several references,[52,196–200] have the potential to revolutionize the drug development process, from hit discovery all the way to lead optimization.

Late-Stage C–H Functionalization of Drug Leads Using Engineered Enzymes

During early drug discovery, bioactive molecules can be structurally modified by selective additions, deletions, and/or replacement of specific atom(s). This process, referred to as late-stage functionalization (LSF) or molecular editing, is generally faster and more cost-effective than de novo synthesis in generating libraries of drug lead analogs. [201,202] Even a subtle structural change can have dramatic effects on the properties of a drug, as exemplified by the addition of a single methyl group drastically increasing kinase inhibitor selectivity[203,204] or the important role played by a halogen-bonding interaction.[205] Thus, LSF has great potential to accelerate the discovery of drug lead derivatives exhibiting optimized activity, safety, and/or drug metabolism/pharmacokinetic (DMPK) profile. To rapidly build

large libraries of molecules derived from a single druglike scaffold, unactivated C–H bonds have high value as points of diversification (i.e., C–H bonds are replaced with carbon–heteroatom or carbon–carbon bonds). Over the last two decades, new methods in transition metal, photoredox, and metallophotoredox catalysis and other chemistry fields have proven to be powerful tools for LSF in drug discovery laboratories.[202] However, predicting and controlling regioselectivity still remains a challenge in structurally complex molecules.[206] Therefore, additional methodologies that achieve a high degree of C–H selectivity are highly desirable to further expand into novel chemical space in a fast, cheap, and sustainable manner. In the last few years, many impressive advances have been made in the field of synthetic biology, which have great potential for filling the space left by synthetic chemistry methodologies.

To create nonnatural biological systems for new applications, synthetic biology borrows and combines tools from different disciplines, including molecular biology, protein engineering, metabolic engineering, and bioinformatics. The directed evolution of enzymes has contributed enormously to propelling synthetic biology advances forward. The relevance of this methodology is illustrated by the 2018 Nobel Prize in Chemistry awarded to Prof. Frances Arnold for her outstanding contributions to directed evolution.[207,208] In addition to the directed evolution of enzymes, rational and semirational design have also proven to be a very powerful strategy when either x-ray crystal structural data or high-quality homology models, combined with mechanistic data, are available. The generation of huge mutagenesis libraries is now both time- and cost-effective using well-established methods, for example, epPCR, DNA shuffling, single-site saturation mutagenesis, and combinatorial active-site saturation test (CAST).[209] Furthermore, high-throughput enzyme expression using heterologous hosts is efficient in many cases. One bottleneck in directed evolution can be HTS using agar plates or microtiter plates for large ($>10^{12}$) mutant libraries. Nevertheless, the optimization of ultra-HTS technologies relying on flow cytometry and chip-based microfluidic screening for their use in directed evolution campaigns will likely open new doors in the near future.[210] Importantly, machine learning is attracting increasing attention in prediction and decision-making in the field of enzyme engineering.[211] As a result of all these efforts, engineered robust and selective enzymes currently have enormous potential to expedite diverse projects in the pharmaceutical industry, including drug discovery.

Currently, synthetic biology's toolbox contains fascinating biocatalysts that exhibit improved stability, activity, substrate scope, and/or selectivity engineered to fulfill industrial demands. These enzymes can provide access to drug lead analogs that would be challenging, inefficient, or unsafe to prepare by nonenzymatic chemical reactions. **Figure 4** shows relevant examples of heme-, flavin-, *S*-adenosyl-L-methionine (SAM)-, and iron/2-oxoglutarate (Fe^{II}/2OG)-dependent enzymes that have potential in late-stage C–H functionalization of druglike molecules. Besides selectively transforming synthetic compounds under mild aqueous conditions, enzymes have unique capabilities to functionalize natural products (e.g., bacterial and fungal secondary metabolites). Natural products are an important source of scaffolds for drug discovery, together with the more recent alternatives described above. Indeed, 41% of small-molecule anticancer drugs approved in 1981–2019 are natural products or their derivatives.[212,213] Various studies show the successful implementation of engineered cytochromes P450 variants in the late-stage diversification of natural products such as artemisinin,[214] parthenolide,[215] β-cembrenediol,[216–218] tylactone-based macrolide antibiotics,[219] cyperenoic acid,[220] and nigelladine A (**Fig. 4**).[221]

In addition, LSF approaches can rapidly synthesize human drug metabolites. Drug development processes should include toxicity assessments of human drug metabolites that are generated at a level higher than 10% of the parent drug exposure, following the metabolites in safety testing (MIST) guidance reported by the U.S. Food and Drug Administration (FDA).[228] P450 enzymes are the key players in phase I drug metabolism in human liver, turning over the majority of commercially available drugs.[229,230] The pharmaceutical industry therefore needs efficient methods to synthesize products resulting from the action of human P450s on new drug candidates. Microbial P450s are often preferred for the synthesis of putative human drug metabolites, mainly due to the expense of purified recombinant human P450 or hepatic microsomes.[231] Various studies show that mutagenesis can be used to tune the substrate scope, selectivity, and/or activity of microbial P450 enzymes, which often results in improved mimics of the desired human P450 isoform. For example, enzyme engineering was used to obtain P450 variants that produce major human metabolites of omeprazole,[232] diclofenac,[233] naproxen, and propranolol (**Fig. 4A**).[234] Engineered unspecific peroxygenases (UPOs) are an attractive alternative to P450 for synthesizing human drug metabolites.[235,236] Oxyfunctionalization chemistry catalyzed by UPO is comparable to that performed by P450s.[237] However, UPO reactions only require hydrogen peroxide as a co-substrate, in contrast to P450 catalysis, which requires a redox partner(s)/domain, NAD(P)H, and dioxygen. Remarkably, despite the complex catalysis of wild-type P450s, engineered P450 enzymes have been used to catalyze even nonnatural chemistries (**Fig. 4A**),[238] for example, cyclopropanations,[239] which is of current interest in pharmaceutical development.

The presence of halogen atoms (F, Cl, Br, I) can improve the potency and pharmacokinetic profile of many drugs.[240] Indeed, around 30% of pharmaceuticals on the market today contain a halogen. Flavin-dependent halogenases are attractive catalysts for the replacement of C–H bonds with

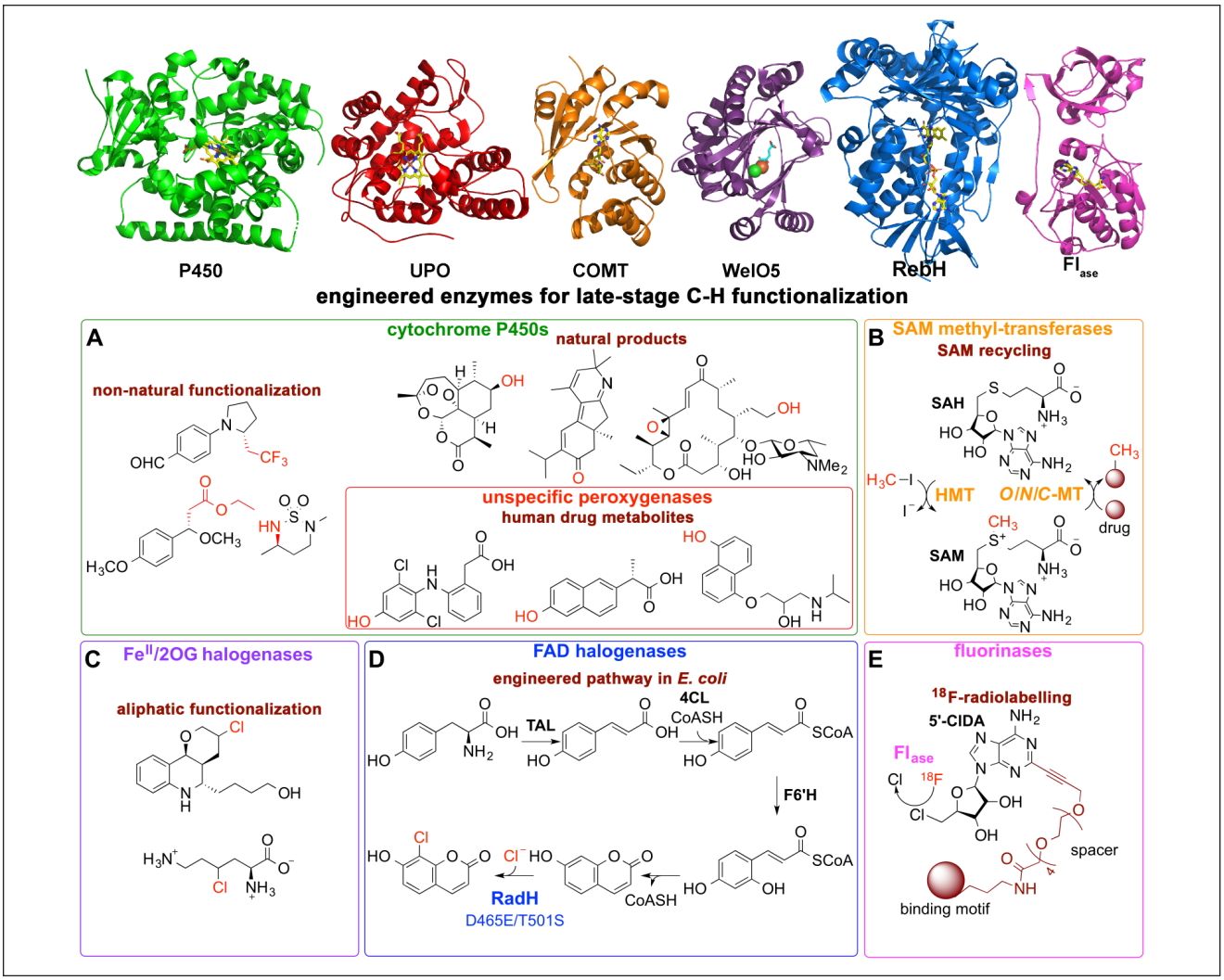

Figure 4. Enzymes harboring great potential for late-stage C–H functionalization of drug leads. Enzyme, reaction, and pathway engineering can provide industrially relevant compounds as the examples shown here. (**A**) P450 variants produce (i) natural products 7-hydroxy-artemisinin,[222] nigelladine A,[221] and juvenimicin A4[219] (left to right); (ii) human drug metabolites derived from diclofenac,[233] naproxen,[233] and propranolol[234] (left to right) (these drugs are also UPO substrates[223,235,236]); and (iii) C–H fluoroalkylation,[224] alkylation,[225] and amination[226] products. (**B**) Conversion of *S*-adenosyl homocysteine (SAH) into SAM is catalyzed by halide methyltransferase (HMT). *O*-, *N*-, or *C*-specific methyltransferases use SAM regenerating SAH.[254] (**C**) Martinelline-derived fragment (top) and lysine (bottom), chlorinated by WelO5* variants[246] and BesD,[227] respectively. (**D**) Engineered enzyme cascade including tyrosine ammonia-lyase (TAL), 4-coumaryl-CoA ligase (4CL), feruloyl CoA 6′-hydroxylase (F6′H), and flavin-dependent halogenase (RadH).[244] (**E**) Fluorinase (Fl_{ase}) and improved variants replace Cl with ^{18}F in 5′-chloro-5′-deoxyadenosine (5′-ClDA), even when the C2 position of the adenine ring presents a long moiety (brown) attached to a bioactive molecule (e.g., peptide with affinity for cancer cells).[249,251] Crystal structures of P450, UPO, catechol- *O*- methyltransferase (COMT), WelO5, RebH, and Fl_{ase} correspond to PDB IDs 1JPZ, 2YOR, 1VID, 5IQS, 2OAL, and 1RQP, respectively.

C–halogen bonds in drug leads due to their high stereo- and regioselectivity.[241] Unfortunately, wild-type flavin-dependent halogenases generally suffer from narrow substrate scope, low catalytic efficiency, and poor stability under industrial conditions. Enzyme, reaction, and pathway engineers, using synthetic biology tools, have made significant progress in increasing the potential of various halogenases for their implementation in lead optimization. For example, variants of the flavin-dependent halogenase RebH exhibiting a significantly higher thermostability than the wild-type enzyme were discovered by only three rounds of directed evolution.[242] In a subsequent study, the substrate scope of one of these thermostabilized variants was expanded using random mutagenesis and a substrate walking approach.[243] Halogenases have been successfully incorporated into biosynthetic pathways to catalyze late-stage halogenations of

druglike scaffolds in an appropriate host. For example, a double mutant of the flavin-dependent halogenase RadH was successfully used in a metabolic engineering effort, directed toward the synthesis of a chlorinated coumarin from glucose in *E. coli* (**Fig. 4D**).[244] While flavin-dependent halogenases act on electron-rich substrates, Fe^{II}/2OG-dependent halogenases catalyze the halogenation of unactivated aliphatic sp^3 carbon centers (**Fig. 4C**). So far, only a few carrier protein-independent Fe^{II}/2OG halogenases have been discovered (e.g., WelO5, AmbO5, and BesD), but engineered variants are already available.[245–247] They seem to exhibit great potential for C–H functionalization in the pharmaceutical industry. Another interesting halogenase type is an adenosyl-fluoride synthase, often referred to as fluorinase, which catalyzes the conversion of fluoride ion and SAM to 5′-fluoro-5′-deoxyadenosine and L-methionine. Fluorinase and its engineered variants have a demonstrated applicability in the late-stage ^{18}F-radiolabeling of bioactive molecules used in positron emission tomography (**Fig. 4E**).[248–251] During drug development, stable and radioisotope labeling of drug leads is often carried out to observe interactions with specific targets and elucidate their DMPK properties.[252,253] As in enzyme-catalyzed late-stage halogenation, there is a great interest in selective installation of methyl groups in complex drug scaffolds using SAM-dependent methyltransferases and their variants. Adding a single methyl group to a drug lead may result in an extraordinary boost in binding affinity, often referred to as the "magic methyl" effect in the medicinal chemistry community.[204] The lack of an efficient SAM recycling system is considered the major bottleneck for the implementation of methyltransferases in industry. However, a breakthrough of pivotal importance in SAM recycling has been recently reported (**Fig. 4B**).[254]

Many of the biocatalysts described here and in previous works[255,256] look ready to be incorporated into screening kits for use in drug discovery laboratories. The challenges to be overcome before they are more widely used in industry have been identified for each enzyme class, and these are being tackled by numerous research groups. The dream, which is now closer to being realized, is to routinely perform and analyze thousands of enzyme-catalyzed C–H functionalization reactions at the microgram scale in a few days. Importantly, the rapid establishment of structure–activity relationships using these technologies will be translated into improved drug discovery timelines and ultimately speedier delivery of novel medicines to patients.

Disruptive Science

O'Hagan and Deng[248] provide an innovative SAM recycling system involving only one regenerating enzyme, which enormously facilitates the implementation of SAM-dependent methyltransferases in biocatalysis. The authors discovered that *S*-adenosyl homocysteine, a by-product in the methyltransferase reactions, can be converted by a halide methyltransferase into SAM using methyl iodide as the methyl source.

Butler et al.[232] recently reviewed natural enzymes that have acquired the ability to catalyze nonnatural chemical reactions after modifying in vitro their amino acid composition, co-factor, available reagents, environment, or combinations thereof. This work covers a rapidly growing research area of considerable interest to organic synthesis.

Sun et al.[250] provide a comprehensive review of the current role of biocatalysis in drug development, as well as the challenges to overcome in this field and the most promising areas under exploration. Interestingly, both academic and industrial viewpoints are provided.

Cell Therapy and Biologics

The panacea of synthetic biology of the future may be the ability to control how patient cells behave in vivo. The ability to remove, modify, and replace these cells is already a reality in oncology treatments where the body's own immune system can be reprogrammed to hunt and destroy cancerous cells, through CAR-T-cell therapy.[257] Synthetic biology can further enhance the special and temporal aspect of CAR-T cells. "Kill" switches, peptide-specific switchable CAR-T cells (sCARs) can improve reversible tumor-specific activation. In an inducible CAR (iCAR) system, two incomplete CAR molecules that can heterodimerize in the presence of a small molecule can function as a switch to activate the T-cell response.[258] CAR-T cells can also be engineered with AND-gate logic, where activation of one CAR receptor activates the expression of a second CAR receptor for a second antigen, further increasing the specificity of targeting.

Yet, there are other ways in which cells can be utilized to deliver medicines more efficiently than systemic treatments. The generation of sensor actuator circuits presents a real possibility that cells can be generated to replace or repair defective cellular processes. There are a series of animal studies that suggest that cellular medicines are on the verge of being able to translate to much needed improvements in therapy. For example, Shao et al. demonstrated that it is possible in mice studies to generate sentinel cells that can reduce feeding and manage diabetic-glucose peaks in response to light signals,[259] or electrogenetic engineered electrosensitive B cells that can release insulin in vivo upon wireless electrical stimulation.[260] Additionally, Chowdhury et al. have engineered an *E. coli* strain to specifically lyse within the tumor microenvironment.[261] On lysis, they released an encoded nanobody antagonist of CD47 (CD47nb), an antiphagocytic receptor that is commonly overexpressed in several human cancer types. They showed delivery of CD47 nanobody by tumor-colonizing bacteria

increased activation of tumor-infiltrating T cells, stimulated rapid tumor regression, and prevented metastasis, leading to long-term survival in a syngeneic tumor model in mice.

The use of bacterial systems is certainly not a new concept, with anecdotal evidence of bacterial infections curing cancers throughout civilized history and numerous clinical studies from the early 20th century that showed partial efficacy, albeit with some significant side effects.[262] Bacteria are unsurprisingly good at negotiating the complex biology of mammalian systems; with an evolutionary need to avoid and hide from immune systems, they are able to specifically hunt out the necrotic centers of solid tumors and hide deep within them. However, these approaches were soon surpassed and largely forgotten by the developments made in small-molecule drug research that have built the modern pharmaceutical industry. However, with the developments made in synthetic biology, it now seems that bacteria have a similar potential, with researchers now examining how human peptides can be used to arm bacteria with "weapons of mass destruction" that are desirable, with the potential to specifically target and destroy tumor masses.[263] Of course, not all bacteria are pathogenic, and there are other opportunities with the microbiomes of our digestive tract and skin systems that are very important to the healthy balance of human life. *E. coli* and other bacterial hosts have been engineered to diagnose diseases and produce and deliver therapeutics in situ. So-called living therapeutics were successfully engineered to target diseases like inflammatory bowel disease,[264] diabetes mellitus,[265] and cancer.[266] Advanced engineering of the bacterial hosts is necessary, including sensing inputs, controlling gene expression, building memory, producing and delivering active compounds, and genetic switches for biocontainment.[267] For example, in a recent study, bacterial biosensors that trigger a differential response to the healthy or diseased mammalian gut were identified, enabling the future design of specific diagnostic and therapeutic biosynthetic circuits.[268]

It is an exciting prospect that these bacteria can be given medically beneficial properties to target chronic skin and metabolism disorders directly, while there also remains a possibility to use these as routes of administration for small molecules targeting other medical conditions.

Legislation on cell therapies is changing in the United States at least, and we are seeing a number of late-stage clinical studies that should pave the way for the future expansion of this field.

Disruptive Science

Yeo et al.[251] describe the first clinical efficacy of CAR-T cells infused in three patients with advanced chronic lymphocytic leukemia that targeted CD19 and contained a costimulatory domain from CD137 and the TCR ζ chain. It heralded a new era of cancer therapy. This therapy was later developed in the clinic and taken to market by Novartis, and in 2017 Tisagenlecleucel became the first FDA-approved medicine that included a gene therapy step.

Merging Workflows

Projects merging machine learning, medicinal chemistry, and synthetic biology into a common workflow have the potential to push the boundaries of modern drug discovery. We see more examples of applying machine learning approaches in synthetic biology to make genetic modules and designs more predictable. This includes, for instance, the prediction of promoter designs[269] and automated tools recommending engineering strategies to improve the microbial chassis performance.[270] We have seen the first synthetic biology approaches creating compound libraries to explore new chemistries, and the combination of these with intracellular selection regimes for hit discovery.[271] Also, deep learning methods are increasingly applied in modern drug discovery.[272] Machine learning approaches are further developed to better predict structure–activity relationships, for example, by applying recurrent neural networks using molecular descriptors as inputs.[273] However, projects combining all three disciplines of synthetic biology, medical chemistry, and machine learning are just beginning to emerge. A recent example is from the field of protein drugs, where deep learning was used to optimize therapeutic antibodies through exploring the high-dimensional protein sequence space.[274] Screening and deep sequencing of relatively small libraries (10^4) were used to train deep neural networks that accurately predicted antigen binding based on antibody sequence and allowed efficient exploration of a large in silico library of ~10^8 variants.

Summary and Outlook

Synthetic biology has impacts throughout the drug discovery value chain and into the development phases and production of pharmaceuticals. The plummeting cost of synthetic DNA, the increasing detailed understanding of the genome, its organization, gene regulation, and the availability of chassis organisms underpin the successes so far and the future potential.

The CRISPR field is exploding with profound impacts on target identification and validation. The development of assays, which underpin the modern drug discovery process, owes much to synthetic biology through engineering of biological circuits.

Genome mining enabled by combination of the availability of genomic sequences, data science, and synthetic biology may catalyze a resurgence in natural product drug discovery. This may be particularly so for "difficult-to-drug" targets such as PPIs and phosphatases, where the modern hit discovery engines of HTS, DEL, and fragment-based lead generation

may fail, but where nature has found a solution. Combinatorial biosynthesis, where components of BGCs are permuted, may allow further biosynthetic diversification without resorting to difficult chemical synthesis. Artificial intelligence and machine learning may be able to augment the refactoring of BGCs to increase the predictability of forming further novel molecules.

Directed evolution coupling genetically encoded libraries, mutation, and selection pressure has revolutionized the development of therapeutic antibodies. Coupling to biological circuits allows in vivo directed evolution for drug discovery and the evolution of new proteins with novel function, such as enzymes to catalyze new transformations for chemical diversification in drug discovery, and "green chemistry" for bulk drug production. While engineered enzymes are already used in individual steps in bulk drug production, whole pathways can be engineered for biosynthetic production. While the successes have been achieved with stepwise directed evolution, continuous directed evolution can enable many more generations to be explored, allowing a deeper search through structure–activity space.

Synthetic biology is enabling new advances in cell therapy, already in oncology with CAR-T, and providing exciting opportunities in cell reprogramming, precise genome editing to correct genetic defects, and reengineering for cell and tissue regeneration. But further, cells can themselves be engineered to sense their environment and respond to treat acute and chronic diseases.

Finally, the merging of workflows between modern technologies such as artificial intelligence and machine learning with synthetic biology and chemistry is emerging to further push the boundaries of drug discovery.

Synthetic biology certainly is impacting all stages of drug discovery and development, and the recognition of the discipline's contribution can further enhance the opportunities for impact on the drug discovery and development value chain.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: M. G, E. R, and L. H. S are funded through the AstraZeneca Postdoctoral Programme.

ORCID iD

Mark J. Wigglesworth https://orcid.org/0000-0002-6421-3414

References

1. Cameron, D. E.; Bashor, C. J.; Collins, J. J. A Brief History of Synthetic Biology. *Nat. Rev. Microbiol*. **2014**, *12*, 381–390.
2. Engineering Biology Research Consortium. What Is Synthetic/Engineering Biology? https://ebrc.org/what-is-synbio/ (accessed Feb 20, 2021).
3. Yang, X.; Wang, Y.; Byrne, R.; et al. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. *Chem. Rev*. **2019**, *11*, 10520–10594.
4. Smalley, E. AI-Powered Drug Discovery Captures Pharma Interest. *Nat. Biotechnol*. **2017**, *3*, 604–605.
5. Synthetic Biology—A Whole New World. *Economist Technical Quarterly*, April *6*, **2019**.
6. Freemont, P. S. Synthetic Biology Industry: Data-Driven Design Is Creating New Opportunities in Biotechnology. *Emerg. Top. Life Sci*. **2019**, *3*, 651–657.
7. Cumbers, J. Synthetic Biology Startups Raised $2 Billion in the First Half of 2020. *Forbes*, Sept 9, **2020**. https://www.forbes.com/sites/johncumbers/2020/09/09/synthetic-biology-startups-raised-30-billion-in-the-first-half-of-2020/?sh=67da95201265 (accessed Feb 20, 2021).
8. Weber, W.; Fussenegger, M. The Impact of Synthetic Biology on Drug Discovery. *Drug Discov. Today* **2009**, *14*, 956–963.
9. Trosset, J. Y.; Carbonell, P. Synthetic Biology for Pharmaceutical Drug Discovery. *Drug Design Dev. Ther*. **2015**, *9*, 6285–302.
10. Porter, D. L.; Levine, B. L.; Kalos, M.; et al. Chimeric Antigen Receptor-Modified T Cells in Chronic Lymphoid Leukemia. *New Engl. J. Med*. **2011**, *365*, 725–733.
11. Franzini, R. M.; Cassie, R. Chemical Space of DNA-Encoded Libraries: Miniperspective. *J. Med. Chem*. **2016**, *59*, 6629–6644.
12. Cimermancic, P.; Medema, M. H.; Claesen, J.; et al. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* **2014**, *158*, 412–421.
13. Riglar, D. T.; Silver, P. A. Engineering Bacteria for Diagnostic and Therapeutic Applications. *Nat. Rev. Microbiol*. **2018**, *16*, 214.
14. Leeat, K.; Zackay, O.; Lotan-Pompan, M.; et al. Promoters Maintain Their Relative Activity Levels under Different Growth Conditions. *Mol. Syst. Biol*. **2013**, *9*, 1.
15. Curran, K. A.; Arim, A. S.; Gupta, A.; et al. Use of Expression-Enhancing Terminators in *Saccharomyces cerevisiae* to Increase mRNA Half-Life and Improve Gene Expression Control for Metabolic Engineering Applications. *Metab. Eng*. **2013**, *19*, 88–97.
16. Novoa, E. M.; Ribas de Pouplana, L. Speeding with Control: Codon Usage, tRNAs, and Ribosomes. *Trends Genet*. **2012**, *28*, 574–581.
17. Salis, H. M.; Mirsky, E.; Voigt, C. A. Automated Design of Synthetic Ribosome Binding Sites to Control Protein Expression. *Nat. Biotechnol*. **2009**, *27*, 946.
18. Park, S. V.; Yang, J.-S.; Jo, H.; et al. Catalytic RNA, Ribozyme, and Its Applications in Synthetic Biology. *Biotechnol. Adv*. **2019**, *37*, 107452.
19. Houser, J. R.; Ford, E.; Chatterjea, S. M.; et al. An Improved Short-Lived Fluorescent Protein Transcriptional Reporter for *Saccharomyces cerevisiae*. *Yeast* **2012**, *29*, 519–530.
20. Rantasalo, J. K.; Penttila, M.; Jantti, J.; et al. Synthetic Toolkit for Complex Genetic Circuit Engineering in *Saccharomyces cerevisiae*. *ACS Synth. Biol*. **2018**, *7*, 1573–1587.

21. Rantasalo, A.; Czeizler, E.; Virtanen, R.; et al. Synthetic Transcription Amplifier System for Orthogonal Control of Gene Expression in *Saccharomyces cerevisiae*. *PLoS One* **2016**, *11*, e0148320.
22. Farzadfard, F.; Perli, S. D. Lu. T. K. Tunable and Multifunctional Eukaryotic Transcription Factors Based on CRISPR/Cas. *ACS Synth. Biol*. **2013**, *2*, 604–613.
23. Jensen, M. K. Design Principles for Nuclease-Deficient CRISPR-Based Transcriptional Regulators. *FEMS Yeast Res*. **2018**, *18*, foy039.
24. Urlinger, S.; Baron, U.; Thellmann, M.; et al. Exploring the Sequence Space for Tetracycline-Dependent Transcriptional Activators: Novel Mutations Yield Expanded Range and Sensitivity. *Proc. Natl. Acad. Sci. U.S.A*. **2000**, *97*, 7963–7968.
25. McIsaac, R. S.; Gibney, P. A.; Chandran, S. S.; et al. Synthetic Biology Tools for Programming Gene Expression without Nutritional Perturbations in *Saccharomyces cerevisiae*. *Nucleic Acids Res*. **2014**, *42*, e48.
26. Rodrigo, G.; Jaramillo, A. AutoBioCAD: Full Biodesign Automation of Genetic Circuits. *ACS Synth. Biol*. **2013**, *2*, 230–236.
27. Bartley, B.; Beal, J.; Clancy, K.; et al. Synthetic Biology Open Language (SBOL) Version 2.0.0. *J. Integr. Bioinform*. **2015**, *12*, 272.
28. Decoene, T.; De Paepe, B.; Maertens, J.; et al. Standardization in Synthetic Biology: An Engineering Discipline Coming of Age. *Crit. Rev. Biotechnol*. **2018**, *38*, 647–656.
29. Casini, A.; MacDonald, J. T.; De Jonghe, J.; et al. One-Pot DNA Construction for Synthetic Biology: The Modular Overlap-Directed Assembly with Linkers (MODAL) Strategy. *Nucleic Acids Res*. **2014**, *42*, e799.
30. De Kok, S.; Stanton, L. H.; Slaby, T.; et al. Rapid and Reliable DNA Assembly via Ligase Cycling Reaction. *ACS Synth. Biol*. **2014**, *3*, 97–106101.
31. Storch, M.; Casini, A.; Mackrow, B.; et al. BASIC: A New Biopart Assembly Standard for Idempotent Cloning Provides Accurate, Single-Tier DNAassembly for Synthetic Biology. *ACS Synth. Biol*. **2015**, *4*, 781–78798.
32. Jensen, N. B.; Strucko, T.; Kildegaard, K. R.; et al. EasyClone: Method for Iterative Chromosomal Integration of Multiple Genes in *Saccharomyces cerevisiae*. *FEMS Yeast Res*. **2014**, *14*, 238–248.
33. Lund, A. M.; Kildegaard, H. F.; Petersen, M. B. K.; et al. A Versatile System for USER Cloning-Based Assembly of Expression Vectors for Mammalian Cell Engineering. *PLoS One* **2014**, *9*, e96693104.
34. Potapov, V.; Ong, J. L.; Kucera, R. B.; et al. Comprehensive Profiling of Four Base Overhang Ligation Fidelity by T4 DNA Ligase and Application to DNA Assembly. *ACS Synth. Biol*. **2018**, *7*, 2665–2674100.
35. Lee, M. E.; DeLoache, W. C.; Cervantes, B.; et al. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synth. Biol*. **2015**, *4*, 975–986.
36. Leavell, M. D.; Singh, A. H.; Kaufmann-Malaga, B. B. High-Throughput Screening for Improved Microbial Cell Factories, Perspective and Promise. *Curr. Opin. Biotechnol*. **2020**, *62*, 22–28.
37. Gach, P. C.; Iwai, K.; Kim, P. W.; et al. Droplet Microfluidics for Synthetic Biology. *Lab Chip* **2017**, *17*, 3388–3400.
38. Gowers, G.-O. F.; Cameron, S. J. S.; Perdones-Montero, A.; et al. Off-Colony Screening of Biosynthetic Libraries by Rapid Laser-Enabled Mass Spectrometry. *ACS Synth. Biol*. **2019**, *8*, 2566–2575.
39. Shaw, W. M.; Yamauchi, H.; Mead, J.; et al. Engineering a Model Cell for Rational Tuning of GPCR Signaling. *Cell* **2019**, *177*, 782–796.
40. Ellis, T.; Wang, X.; Collins, J. J. Diversity-Based, Model-Guided Construction of Synthetic Gene Networks with Predicted Functions. *Nat. Biotechnol*. **2009**, *27*, 465–471.
41. Kotopka, B. J.; Smolke, C. D. Model-Driven Generation of Artificial Yeast Promoters. *Nat. Commun*. **2020**, *11*, **2113**.
42. Berepiki, A.; Kent, R.; Machado, L. F. M.; et al. Development of High-Performance Whole Cell Biosensors Aided by Statistical Modelling. *ACS Synth. Biol*. **2020**, *9*, 576–589.
43. Chao, R.; Mishra, S.; Si, T.; et al. Engineering Biological Systems Using Automated Biofoundries. *Metab. Eng*. **2017**, *42*, 98–108.
44. Hillson, N.; Caddick, M.; Cai, Y.; et al. Building a Global Alliance of Biofoundries. *Nat. Commun*. **2019** *10*, **2040**.
45. Gupta, S.; Pandey, S. The Structural and Functional Analysis of *Escherichia coli* Genome. In *Food Molecular Microbiology*. CRC Press: Boca Raton, FL, 2019; pp 141–162.
46. Nakagawa, A.; Matsumura, E.; Koyanagi, T.; et al. Total Biosynthesis of Opiates by Stepwise Fermentation Using Engineered *Escherichia coli*. *Nat. Commun*. **2016**, *7*, 10390.
47. Yang, D.; Park, S. Y.; Park, Y. S.; et al. Metabolic Engineering of *Escherichia coli* for Natural Product Biosynthesis. *Trends Biotechnol*. **2020**, *38*, 745–765.
48. Jaitzig, J.; Li, J.; Suessmuth, R. D.; et al. Reconstituted Biosynthesis of the Nonribosomal Macrolactone Antibiotic Valinomycin in *Escherichia coli*. *ACS Synth. Biol*. **2014**, *3*, 432–438.
49. Ajikumar, P. K.; Xiao, W. H.; Tyo, K. E. J.; et al. Stephanopoulos Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*. *Science* **2010**, *330*, 70–74.
50. Nakagawa, A.; Matsumura, E.; Koyanagi, T.; et al. Total Biosynthesis of Opiates by Stepwise Fermentation Using Engineered *Escherichia coli*. *Nat. Commun*. **2016**, *7*, 10390.
51. McCafferty, J.; Griffiths, A. D.; Winter, G.; et al. Phage Antibodies: Filamentous Phage Displaying Antibody Variable Domains. *Nature* **1990**, *348*, 552–554.
52. Esvelt, K. M.; Carlson, J. C.; Liu, D. R. A System for the Continuous Directed Evolution of Biomolecules. *Nature* **2011**, *472*, 499–503.
53. Yang, X.; Lennard, K. R.; He, C.; et al. A Lanthipeptide Library Used to Identify a Protein–Protein Interaction Inhibitor. *Nat. Chem. Biol*. **2018**, *14*, 375–380.
54. Male, A. L.; Forafonov, F.; Cuda, F.; et al. Targeting *Bacillus anthracis* Toxicity with a Genetically Selected Inhibitor of the PA/CMG2 Protein-Protein Interaction. *Sci. Rep*. **2017**, *7*, 3104.
55. Riglar, D. T.; Silver, P. A. Engineering Bacteria for Diagnostic and Therapeutic Applications. *Nat. Rev. Microbiol*. **2018**, *16*, 214.
56. Kurtz, C. B.; Millet, Y. A.; Puurunen, M. K.; et al. An Engineered *E. coli* Nissle Improves Hyperammonemia and Survival in Mice and Shows Dose-Dependent Exposure in Healthy Humans. *Sci. Transl. Med*. **2019**, *11*, eaau7975.

57. Safety, Tolerability and Pharmacodynamics of SYNB1020. NCT03447730. https://clinicaltrials.gov (accessed Feb 21, 2021).
58. Isabella, V. M.; Ha, B. N.; Castillo, M. J.; et al. Development of a Synthetic Live Bacterial Therapeutic for the Human Metabolic Disease Phenylketonuria. *Nat. Biotechnol*. **2018**, *36*, 857–864.
59. Safety and Tolerability of SYNB1618 in Healthy Adult Volunteers and Adult Subjects with Phenylketonuria (PKU). NCT03516487. https://clinicaltrials.gov (accessed Feb 21, 2021).
60. Efficacy, Safety and Tolerability of AG013 in Oral Mucositis Compared to Placebo When Administered Three Times per Day. NCT03234465. https://clinicaltrials.gov (accessed Feb 21, 2021).
61. UniProt. www.uniprot.org (accessed Feb 21, 2021).
62. Baeshen, N. A.; Baeshen, M. N.; Sheikh, A.; et al. Cell Factories for Insulin Production. *Microb. Cell Fact*. **2014**, *13*, 141–149.
63. Bill, R. M. Recombinant Protein Subunit Vaccine Synthesis in Microbes: A Role for Yeast? *J. Pharm. Pharmacol*. **2015**, *67*, 319–328.
64. Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; et al. High-Level Semi-Synthetic Production of the Potent Antimalarial Artemisinin. *Nature* **2013**, *496*, 528–532.
65. Khurana, V.; Lindquist, S. Modelling Neurodegeneration in *Saccharomyces cerevisiae*: Why Cook with Baker's Yeast? *Nat. Rev. Neurosci*. **2010**, *11*, 436–449.
66. Menacho-Marquez, M.; Murguia, J. R. Yeast on Drugs: *Saccharomyces cerevisiae* as a Tool for Anticancer Drug Research. *Clin. Transl. Oncol*. **2007**, *9*, 221–228.
67. Ferreira, R.; Limeta, A.; Nielsen, J. Tackling Cancer with Yeast-Based Technologies. *Trends Biotechnol*. **2019**, *37*, 592–603.
68. Giaever, G.; Chu, A. M.; Ni, L.; et al. Functional Profiling of the *Saccharomyces cerevisiae* Genome. *Nature* **2002**, *418*, 387–391.
69. Srivas, R.; Shen, J. P.; Yang, C. C.; et al. A Network of Conserved Synthetic Lethal Interactions for Exploration of Precision Cancer Therapy. *Mol. Cell* **2016**, *63*, 514–525.
70. Pretorius, I. S.; Boeke, J. D. Yeast 2.0—Connecting the Dots in the Construction of the World's First Functional Synthetic Eukaryotic Genome. *FEMS Yeast Res*. **2018**, *18*, foy032.
71. Luo, Z.; Wang, L.; Wang, Y.; et al. Identifying and Characterizing SCRaMbLEd Synthetic Yeast Using ReSCuES. *Nat. Commun*. **2018**, *9*, **1930**.
72. Blount, B. A.; Gowers, G-O. F.; Ho, J. C. H.; et al. Rapid Host Strain Improvement by In Vivo Rearrangement of a Synthetic Yeast Chromosome. *Nat. Commun*. **2018**, *9*, **1932**.
73. Boder, E. T.; Wittrup, K. D. Yeast Surface Display for Screening Combinatorial Polypeptide Libraries. *Nat. Biotechnol*. **1997**, *15*, 553–557.
74. Spadiut, O.; Capone, S.; Krainer, F.; et al. Microbials for the Production of Monoclonal Antibodies and Antibody Fragments. *Trends Biotechnol*. **2014**, *32*, 54–60.
75. Doerner, A.; Rhiel, L.; Zielonka, S.; et al. Therapeutic Antibody Engineering by High Efficiency Cell Screening. *FEBS Lett*. **2014**, *588*, 278–287.
76. Tomimoto, K.; Fujita, Y.; Iwaki, T.; et al. Protease-Deficient *Saccharomyces cerevisiae* Strains for the Synthesis of Human-Compatible Glycoproteins. *Biosci. Biotechnol. Biochem*. **2013**, *77*, 2461–2466.
77. De Wachter, C.; Van Landuyt, L.; Callewaert, N. Engineering of Yeast Glycoprotein Expression. In *Advances in Biochemical Engineering/Biotechnology*. Springer: Berlin, 2018; pp 1–43.
78. Tsuruta, H.; Paddon, C. J.; Eng, D.; et al. High-Level Production of Amorpha-4,11-Diene, a Precursor of the Antimalarial Agent Artemisinin, in *Escherichia coli*. *PLoS One* **2009**, *4*, e4489.
79. Paddon, C. J.; Keasling, J. D. Semi-Synthetic Artemisinin: A Model for the Use of Synthetic Biology in Pharmaceutical Development. *Nat. Rev. Microbiol*. **2014**, *12*, 355–367.
80. Cravens, A.; Payne, J.; Smolke, C. D. Synthetic Biology Strategies for Microbial Biosynthesis of Plant Natural Products. *Nat. Commun*. **2019**, *10*, **2142**.
81. Carqueijeiro, I.; Langley, C.; Grzech, D.; et al. Beyond the Semi-Synthetic Artemisinin: Metabolic Engineering of Plant-Derived Anti-Cancer Drugs. *Curr. Opin. Biotechnol*. **2020**, *65*, 17–24.
82. Brown, S.; Clastre, M.; Courdavault, V.; et al. De Novo Production of the Plant-Derived Alkaloid Strictosidine in Yeast. *Proc. Natl Acad. Sci. U.S.A*. **2015**, *112*, 3205–3210.
83. Qu, Y.; Easson, M. L. A. E.; Froese, J.; et al. Completion of the Seven-Step Pathway from Tabersonine to the Anticancer Drug Precursor Vindoline and Its Assembly in Yeast. *Proc. Natl Acad. Sci. U.S.A*. **2015**, *112.19* 6224–6229.
84. Li, Y.; Li, S.; Thodey, K.; et al. Complete Biosynthesis of Noscapine and Halogenated Alkaloids in Yeast. *Proc. Natl Acad. Sci. U.S.A*. **2018**, *115*, E3922–E3931.
85. Galanie, S.; Thodey, K.; Trenchard, I. J.; et al. Complete Biosynthesis of Opioids in Yeast. *Science* **2015**, *349*, 1095–1100.
86. Luo, X.; Reiter, M. A.; d'Espaux, L.; et al. Complete Biosynthesis of Cannabinoids and Their Unnatural Analogues in Yeast. *Nature* **2019**, *567*, 123–126.
87. Xu, X.; Liu, Y.; Du, G.; et al. Microbial Chassis Development for Natural Product Biosynthesis. *Trends Biotechnol*. **2020**, *38*, 779–796.
88. Liu, Y.; Long, L. J.; Du, G.; et al. Synthetic Biology Toolbox and Chassis Development in *Bacillus subtilis*. *Trends Biotechnol*. **2019**, *37*, 548–562.
89. van Tilburg, A. Y.; Cao, H.; van der Meulen, S. B.; et al. Metabolic Engineering and Synthetic Biology Employing *Lactococcus lactis* and *Bacillus subtilis* Cell Factories. *Curr. Opin. Biotechnol*. **2019**, *59*, 1–7.
90. Yang, Z.; Zisheng, Z. Engineering Strategies for Enhanced Production of Protein and Bio-Products in *Pichia pastoris*: A Review. *Biotechnol. Adv*. **2018**, *36*, 182–195.
91. Liu, R.; Zixin, D.; Tiangang, L. Streptomyces Species: Ideal Chassis for Natural Product Discovery and Overproduction. *Metab. Eng*. **2018**, *50*, 74–84.
92. Cook, D.; Brown, D.; Alexander, R; et al. Lessons Learned from the Fate of AstraZeneca's Drug Pipeline: A Five-Dimensional Framework. *Nat. Rev. Drug Discov*. **2014**, *133*, 419–431.
93. Swinney, D. C.; Anthony, J. How Were New Medicines Discovered? *Nat. Rev. Drug Discov*. **2011**, *10*, 507–519.

94. Morgan, P.; Brown, D. G.; Lennard, S.; et al. Impact of a Five-Dimensional Framework on R&D Productivity at AstraZeneca. *Nat. Rev. Drug Discov*. **2018**, *17*, 167–181.
95. Jinek, M.; Chylinski, K.; Fonfara, I.; et al. A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* **2012**, *337*, 816–821.
96. Cong, L.; Ran, F. A.; Cox, D.; et al. Multiplex Genome Engineering Using CRISPR/Cas Systems. *Science* **2013**, *339*, 819–823.
97. Hultquist, J. F.; Hiatt, J.; Schumann, K.; et al. CRISPR-Cas9 Genome Engineering of Primary CD4(+) T Cells for the Interrogation of HIV-Host Factor Interactions. *Nat. Protoc*. **2019**, *14*, 1–27.
98. Martufi, M.; Good, R. B.; Rapiteanu, R.; et al. Single-Step, High-Efficiency CRISPR-Cas9 Genome Editing in Primary Human Disease-Derived Fibroblasts. *CRISPR J*. **2019**, *2*, 31–40.
99. Borestrom, C.; Jonebring, A.; Guo, J.; et al. A CRISP(e) R View on Kidney Organoids Allows Generation of an Induced Pluripotent Stem Cell-Derived Kidney Model for Drug Discovery. *Kidney Int*. **2018**, *94*, 1099–1110.
100. Henriksson, J.; Chen, X.; Gomes, T.; et al. Genome-Wide CRISPR Screens in T Helper Cells Reveal Pervasive Crosstalk between Activation and Differentiation. *Cell* **2019**, *176*, 882–896.e18.
101. Behan, F. M.; Iorio, F.; Picco, G.; et al. Prioritization of Cancer Therapeutic Targets Using CRISPR-Cas9 Screens. *Nature* **2019**, *568*, 511–516.
102. Wroblewska, A.; Dhainaut, M.; Ben-Zvi, B.; et al. Protein Barcodes Enable High-Dimensional Single-Cell CRISPR Screens. *Cell* **2018**, *175*, 1141–1155.e16.
103. Liu, Y.; Yu, C.; Daley, T. P.; et al. CRISPR Activation Screens Systematically Identify Factors That Drive Neuronal Fate and Reprogramming. *Cell Stem Cell* **2018**, *23*, 758–771.e8.
104. Shifrut, E.; Carnevale, J.; Tobin, V.; et al. Genome-Wide CRISPR Screens in Primary Human T Cells Reveal Key Regulators of Immune Function. *Cell* **2018**, *175*, 1958–1971. e15.
105. Kim, H. S.; Lee, K.; Kim, S. J.; et al. Arrayed CRISPR Screen with Image-Based Assay Reliably Uncovers Host Genes Required for Coxsackievirus Infection. *Genome Res*. **2018**, *28*, 859–868.
106. Henser-Brownhill, T.; Monserrat, J.; Scaffidi, P. Generation of an Arrayed CRISPR-Cas9 Library Targeting Epigenetic Regulators: From High-Content Screens to In Vivo Assays. *Epigenetics* **2017**, *12*, 1065–1075.
107. Starkuviene, V.; Kallenberger, S. M.; Beil, N.; et al. High-Density Cell Arrays for Genome-Scale Phenotypic Screening. *SLAS Discov*. **2019**, *24*, 274–283.
108. Metzakopian, E.; Strong, A.; Iyer, V.; et al. Enhancing the Genome Editing Toolbox: Genome Wide CRISPR Arrayed Libraries. *Sci. Rep*. **2017**, *7*, **2244**.
109. Li, Z.; Ivanov, A. A.; Su, R.; et al. The OncoPPi Network of Cancer-Focused Protein-Protein Interactions to Inform Biological Insights and Therapeutic Strategies. *Nat. Commun*. **2017**, *8*, 14356.
110. Fang, H.; ULTRA-DD Consortium, De Wolf, H.; et al. A Genetics-Led Approach Defines the Drug Target Landscape of 30 Immune-Related Traits. *Nat. Genet*. **2019**, *51*, 1082–1091.
111. Jin, H. J.; Jung, S.; DebRoy, A. R.; et al. Identification and Validation of Regulatory SNPs That Modulate Transcription Factor Chromatin Binding and Gene Expression in Prostate Cancer. *Oncotarget*. **2016**, *7*, 54616–54626.
112. Gardner, T. S.; Cantor, C. R.; Collins, J. J. Construction of a Genetic Toggle Switch in *Escherichia coli*. *Nature* **2000**, *403*, 339–342.
113. Basu, S.; Gerchman, Y.; Collins, C.; et al. A Synthetic Multicellular System for Programmed Pattern Formation. *Nature* **2005**, *434*, 1130–1134.
114. Park, S. H.; Zarrinpar, A.; Lim, W. A. Rewiring MAP Kinase Pathways Using Alternative Scaffold Assembly Mechanisms. *Science* **2003**, *299*, 1061–1064.
115. Kramer, B. P.; Viretta, A. U.; Daoud-El-Baba, M.; et al. An Engineered Epigenetic Transgene Switch in Mammalian Cells. *Nat. Biotechnol*. **2004**, *22*, 867–870.
116. Deyell, M.; Ameta, S.; Nghe, P. Large Scale Control and Programming of Gene Expression Using CRISPR. *Semin. Cell Dev. Biol*. **2019**, *96*, 124–132.
117. Heidari, R.; Shaw, D. M.; Elger BS CRISPR and the Rebirth of Synthetic Biology. *Sci. Eng. Ethics* **2017**, *23*, 351–363.
118. European Commission. Opinion on Synthetic Biology I Definition. http://ec.europa.eu/health/scientific_committees/emerging/docs/scenihr_o_044.pdf (accessed Feb 21, 2021).
119. Lundin, A.; Porritt, M. J.; Jaiswal, H.; et al. Development of an ObLiGaRe Doxycycline Inducible Cas9 System for Pre-Clinical Cancer Drug Discovery. *Nat. Commun*. **2020**, *11*, 4903.
120. Brophy, J. A. N.; Voigt, C. A. Principles of Genetic Circuit Design. *Nat. Methods* **2014**, *11*, 508–520.
121. Weber, W.; Fussenegger, M. Synthetic Gene Networks in Mammalian Cells. *Curr. Opin. Biotechnol*. **2010**, *21*, 690–696.
122. Liu, W.; Yuan, J. S.; Stewart, C. N., Jr. Advanced Genetic Tools for Plant Biotechnology. *Nat. Rev. Genet*. **2013**, *14*, 781–793.
123. Uings, I. J.; Farrow, S. N. Cell Receptors and Cell Signalling. *Mol. Pathol*. **2000**, *53*, 295–299.
124. Hauser, A. S.; Attwood, M. M.; Rask-Andersen, M.; et al. Trends in GPCR Drug Discovery: New Agents, Targets and Indications. *Nat. Rev. Drug Discov*. **2017**, *16*, 829–842.
125. Eglen, R. M.; Reisine, T. Photoproteins: Important New Tools in Drug Discovery. *Assay Drug Dev. Technol*. **2008**, *6*, 659–672.
126. Barnea, G.; Strapps, W.; Herrada, G.; et al. The Genetic Design of Signaling Cascades to Record Receptor Activation. *Proc. Natl. Acad. Sci. U.S.A*. **2008**, *105*, 64–69.
127. Brown, A. J.; Dyos, S. L.; Whiteway, M. S.; et al. Functional Coupling of Mammalian Receptors to the Yeast Mating Pathway Using Novel Yeast/Mammalian G Protein α-Subunit Chimeras. *Yeast* **2000**, *16*, 11–22.
128. Pausch, M. H. G-Protein-Coupled Receptors in *Saccharomyces cerevisiae*: High-Throughput Screening Assays for Drug Discovery. *Trends Biotechnol*. **1997**, *15*, 487–494.

129. Minic, J.; Persuy, M.-A.; Godel, E.; et al. Functional Expression of Olfactory Receptors in Yeast and Development of a Bioassay for Odorant Screening. *FEBS J*. **2005**, *272*, 524–537.
130. Sachpatzidis, A.; Benton, B. K.; Manfredi, J. P.; et al. Identification of Allosteric Peptide Agonists of CXCR4. *J. Biol. Chem*. **2003**, *278*, 896–907.
131. Shaw, W. M.; Yamauchi, H.; Mead, J.; et al. Engineering a Model Cell for Rational Tuning of GPCR Signaling. *Cell* **2019**, *177*, 782–796.e27.
132. Ryan, D. P.; Matthews, J. M. Protein–Protein Interactions in Human Disease. *Curr. Opin. Struct. Biol*. **2005**, *15*, 441–446.
133. Michnick, S. W.; Ear, P. H.; Manderson, E. N.; et al. Universal Strategies in Research and Drug Discovery Based on Protein-Fragment Complementation Assays. *Nat. Rev. Drug Discov*. **2007**, *6*, 569–582.
134. Remy, I.; Michnick, S. W. Visualization of Biochemical Networks in Living Cells. *Proc. Natl. Acad. Sci. U.S.A*. **2001**, *98*, 7678–7683.
135. MacDonald, M. L.; Lamerdin, J.; Owens, S.; et al. Identifying Off-Target Effects and Hidden Phenotypes of Drugs in Human Cells. *Nat. Chem. Biol*. **2006**, *2*, 329–337.
136. Patel, A.; Murray, J.; McElwee-Whitmer, ; et al. A Combination of Ultrahigh Throughput PathHunter and Cytokine Secretion Assays to Identify Glucocorticoid Receptor Agonists. *Anal. Biochem*. **2009**, *385*, 286–292.
137. Hwang, Y.-C.; Chu, J. J.-H.; Yang, P. L.; et al. Rapid Identification of Inhibitors That Interfere with Poliovirus Replication Using a Cell-Based Assay. *Antiviral Res*. **2008**, *77*, 232–236.
138. Yu, X.; Sainz, B.; Uprichard, S. L. Development of a Cell-Based Hepatitis C Virus Infection Fluorescent Resonance Energy Transfer Assay for High-Throughput Antiviral Compound Screening. *Antimicrob. Agents Chemother*. **2009**, *53*, 4311–4319.
139. Horswill, A. R.; Savinov, S. N.; Benkovic, S. J. A Systematic Method for Identifying Small-Molecule Modulators of Protein-Protein Interactions. *Proc. Natl. Acad. Sci. U.S.A*. **2004**, *101*, 15591–15596.
140. Tavassoli, A.; Benkovic, S. J. Genetically Selected Cyclic-Peptide Inhibitors of AICAR Transformylase Homodimerization. *Angew. Chem. Int. Ed*. **2005**, *44*, 2760–2763.
141. Tavassoli, A.; Lu, Q.; Gam, J.; et al. Inhibition of HIV Budding by a Genetically Selected Cyclic Peptide Targeting the Gag–TSG101 Interaction. *ACS Chem. Biol*. **2008**, *3*, 757–764.
142. Miranda, E.; Nordgren, I. K.; Male, A. L.; et al. A Cyclic Peptide Inhibitor of HIF-1 Heterodimerization That Inhibits Hypoxia Signaling in Cancer Cells. *J. Am. Chem. Soc*. **2013**, *135*, 10418–10425.
143. Birts, C. N.; Nijjar, S. K.; Mardle, C. A.; et al. A Cyclic Peptide Inhibitor of C-Terminal Binding Protein Dimerization Links Metabolism with Mitotic Fidelity in Breast Cancer Cells. *Chem. Sci*. **2013**, *4*, 3046.
144. Grkovic, S.; Hardie, K. M.; Brown, M. H.; et al. Interactions of the QacR Multidrug-Binding Protein with Structurally Diverse Ligands: Implications for the Evolution of the Binding Pocket. *Biochemistry* **2003**, *42*, 15226–15236.
145. Ramos, J. L.; Martínez-Bueno, M.; Molina-Henares, A. J.; et al. The TetR Family of Transcriptional Repressors. *Microbiol. Mol. Biol. Rev*. **2005**, *69*, 326–356.
146. Fussenegger, M.; Morris, R. P.; Fux, C.; et al. Streptogramin-Based Gene Regulation Systems for Mammalian Cells. *Nat. Biotechnol*. **2000**, *18*, 1203–1208.
147. Aubel, D.; Morris, R.; Lennon, B.; et al. Design of a Novel Mammalian Screening System for the Detection of Bioavailable, Non-Cytotoxic Streptogramin Antibiotics. *J. Antibiot*. **2001**, *54*, 44–55.
148. Gossen, M.; Freundlieb, S.; Bender, G.; et al. Transcriptional Activation by Tetracyclines in Mammalian Cells. *Science* **1995**, *268*, 1766–1769.
149. Scott, L. H.; Mathews, J. C.; Flematti, G. R.; et al. An Artificial Yeast Genetic Circuit Enables Deep Mutational Scanning of an Antimicrobial Resistance Protein. *ACS Synth. Biol*. **2018**, *7*, 1907–1917.
150. Weber, W.; Schoenmakers, R.; Keller, B.; et al. A Synthetic Mammalian Gene Circuit Reveals Antituberculosis Compounds. *Proc. Natl. Acad. Sci. U.S.A*. **2008**, *105*, 9994–9998.
151. Gonzalez-Nicolini, V.; Fux, C.; Fussenegger, M. A Novel Mammalian Cell-Based Approach for the Discovery of Anticancer Drugs with Reduced Cytotoxicity on Non-Dividing Cells. *Invest. New Drugs* **2004**, *22*, 253–262.
152. Cuthbertson, L.; Nodwell, J. R. The TetR family of regulators. *Microbiol. Mol. Biol. Rev*. **2013**, *77*, 440–475.
153. Tang, S.-Y.; Qian, S.; Akinterinwa, O.; et al. Screening for Enhanced Triacetic Acid Lactone Production by Recombinant *Escherichia coli* Expressing a Designed Triacetic Acid Lactone Reporter. *J. Am. Chem. Soc*. **2013**, *135*, 10099–10103.
154. Nelson, J. W.; Plummer, M. S.; Blount, K. F.; et al. Small Molecule Fluoride Toxicity Agonists. *Chem. Biol*. **2015**, *22*, 527–534.
155. Blind, M.; Blank, M. Aptamer Selection Technology and Recent Advances. *Mol. Ther. Nucleic Acids* **2015**, *4*, e223.
156. Buskirk, A. R.; Ong, Y.-C.; Gartner, Z. J.; et al. Directed Evolution of Ligand Dependence: Small-Molecule-Activated Protein Splicing. *Proc. Natl. Acad. Sci. U.S.A*. **2004**, *101*, 10505–10510.
157. Yen, L.; Svendsen, J.; Lee, J.-S.; et al. Exogenous Control of Mammalian Gene Expression through Modulation of RNA Self-Cleavage. *Nature* **2004**, *431*, 471–476.
158. Patridge, E.; Gareiss, P.; Kinch, M. S.; et al. An Analysis of FDA-Approved Drugs: Natural Products and Their Derivatives. *Drug Discov. Today* **2016**, *21*, 204–207.
159. Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014 *J. Nat. Prod*. **2016**, *79*, 629–661.
160. Teague, S. J.; Davis, A. M.; Leeson, P. D.; et al. The Design of Leadlike Combinatorial Libraries. *Angew. Chem. Int. Ed*. **1999**, *38*, 3743–3748.
161. Davis, A. M.; Plowright, A. T.; Valeur, E. Directing Evolution: The Next Revolution in Drug Discovery? *Nat. Rev. Drug Discov*. **2017**, *16*, 681–698.
162. de la Torre, B. G.; Albericio, F. The Pharmaceutical Industry in 2018. An Analysis of FDA Drug Approvals from the Perspective of Molecules. *Molecules* **2019**, *24*, 809/1–809/12.

163. Harvey, A. L.; Edrada-Ebel, R.; Quinn, R. J. The Re-Emergence of Natural Products for Drug Discovery in the Genomics Era. *Nat. Rev. Drug Discov*. **2015**, *14*, 111–129.
164. Charlop-Powers, Z.; Pregitzer, C. C.; Lemetre, C.; et al. Urban Park Soil Microbiomes Are a Rich Reservoir of Natural Product Biosynthetic Diversity. *Proc. Natl. Acad. Sci. U.S.A*. **2016**, *113*, 14811–14816.
165. Cimermancic, P.; Medema, M. H.; Claesen, J.; et al. Insights into Secondary Metabolism from a Global Analysis of Prokaryotic Biosynthetic Gene Clusters. *Cell* **2014**, *158*, 412–421.
166. Klein, J.; Heal, J. R.; Hamilton, W. D. O.; et al. Yeast Synthetic Biology Platform Generates Novel Chemical Structures as Scaffolds for Drug Discovery. *ACS Synth. Biol*. **2014**, *3*, 314–323.
167. Naesby, M.; Nielsen, S. V. S.; Nielsen, C. A. F.; et al. Yeast Artificial Chromosomes Employed for Random Assembly of Biosynthetic Pathways and Production of Diverse Compounds in *Saccharomyces cerevisiae*. *Microb. Cell Fact*. **2009**, *8*, 45.
168. Ajikumar, P. K.; Xiao, Wen-Hai, T.; et al. Isoprenoid Pathway Optimization for Taxol Precursor Overproduction in *Escherichia coli*. *Science* **2010**, *330*, 70–74.
169. Paddon, C. J.; Westfall, P. J.; Pitera, D. J.; et al. High-Level Semi-Synthetic Production of the Potent Antimalarial Artemisinin. *Nature* **2013**, *496*, 528–532.
170. Thodey, K.; Galanie, S.; Smolke, C. D. A Microbial Biomanufacturing Platform for Natural and Semisynthetic Opioids. *Nat. Chem. Biol*. **2014**, *10*, 837–844.
171. Jakočiūnas, T.; Klitgaard, A. K.; Kontou, E. E.; et al. Programmable Polyketide Biosynthesis Platform for Production of Aromatic Compounds in Yeast. *Synth. Syst. Biotechnol*. **2020**, *5*, 11–18.
172. Awan, A. R.; Blount, B. A.; Bell, D. J.; et al. Biosynthesis of the Antibiotic Nonribosomal Peptide Penicillin in Baker's Yeast. *Nat. Commun*. **2017**, *8*, 15202.
173. Trenchard, I. J.; Smolke, C. D. Engineering Strategies for the Fermentative Production of Plant Alkaloids in Yeast. *Metab. Eng*. **2015**, *30*, 96–104.
174. Smanski, M. J.; Zhou, H.; Claesen, J.; et al. Synthetic Biology to Access and Expand Nature's Chemical Diversity. *Nat. Rev. Microbiol*. **2016**, *14*, 135.
175. Nieselt, K.; Battke, F.; Herbig, A.; et al. The Dynamic Architecture of the Metabolic Switch in *Streptomyces coelicolor*. *BMC Genom*. **2010**, *11*, 10.
176. Kushwaha, M.; Salis, H. M. A Portable Expression Resource for Engineering Cross-Species Genetic Circuits and Pathways. *Nat. Commun*. **2015**, *6*, 7832.
177. Keller, N. P. Translating Biosynthetic Gene Clusters into Fungal Armor and Weaponry. *Nature Chem. Biol*. **2015**, *11*, 671–677.
178. Tang, X.; Li, J.; Millan-Aguinaga, N.; et al. Identification of Thiotetronic Acid Antibiotic Biosynthetic Pathways by Target-Directed Genome Mining. *ACS Chem. Biol*. **2015**, *10*, 2841–2849.
179. Hsu-Hua, Y.; Manmeet, A.; Yi-Ming, C.; et al. Resistance Gene-Guided Genome Mining: Serial Promoter Exchanges in *Aspergillus nidulans* Reveal the Biosynthetic Pathway for Fellutamide B, a Proteasome Inhibitor. *ACS Chem Biol*. **2016**, *11*, 2275–2284.
180. Nyilasi, I.; Kocsube, S.; Krizsan, K.; et al. Susceptibility of Clinically Important Dermatophytes against Statins and Different Statin-Antifungal Combinations. *Med. Mycol*. **2014**, *52*, 140–148.
181. Abe, Y.; Suzuki, T.; Mizuno, T.; et al. Effect of Increased Dosage of the ML-236B (Compactin) Biosynthetic Gene Cluster on ML-236B Production in *Penicillium citrinum*. *Mol. Genet. Genom*. **2002**, *268*, 130–137.
182. Lin, H.-C.; Chooi, Y.-H.; Dhingra, S.; et al. The Fumagillin Biosynthetic Gene Cluster in *Aspergillus fumigatus* Encodes a Cryptic Terpene Cyclase Involved in the Formation of β-Trans-Bergamotene. *J. Am. Chem. Soc*. **2013**, *135*, 4616–4619.
183. Yan, Y.; Liu, Q.; Zang, X.; et al. Resistance-Gene-Directed Discovery of a Natural-Product Herbicide with a New Mode of Action. *Nature* **2018**, *559*, 415–418.
184. Yue, Q.; Li, Y.; Li, C.; et al. Genomics-Driven Discovery of a Novel Self-Resistance Mechanism in the Echinocandin-Producing Fungus *Pezicula radicicola*. *Environ. Microbiol*. **2018**, *20*, 3154–3167.
185. Hansen, B. G.; Genee, H. J.; Kaas, C. S.; et al. A New Class of IMP Dehydrogenase with a Role in Self-Resistance of Mycophenolic Acid Producing Fungi. *BMC Microbiol*. **2011**, *11*, 202.
186. By Bushley, K. E.; Raja, R.; Jaiswal,; et al. The Genome of *Tolypocladium inflatum*: Evolution, Organization, and Expression of the Cyclosporin Biosynthetic Gene Cluster. *PLoS Genet*. **2013**, *9*, e1003496.
187. Davis, A. M.; Plowright, A. T.; Valeur, E. Directing Evolution: The Next Revolution in Drug Discovery? *Nat. Rev. Drug Discov*. **2017**, *16*, 681–698.
188. Austin, R. J.; Ja, W. W.; Roberts, R. W. Evolution of Class-Specific Peptides Targeting a Hot Spot of the Gαs Subunit. *J. Mol. Biol*. **2008**, *377*, 1406–1418.
189. Young, T. S.; Young, D. D.; Ahmad, I.; et al. Evolution of Cyclic Peptide Protease Inhibitors. *Proc. Natl. Acad. Sci. U.S.A*. **2011**, *108*, 11052–11056.
190. Fischbach, M. A.; Lai, J. R.; Roche, E. D.; et al. Directed Evolution Can Rapidly Improve the Activity of Chimeric Assembly-Line Enzymes. *Proc. Natl. Acad. Sci. U.S.A*. **2007**, *104*, 11951–11956.
191. Evans, B. S.; Chen, Y.; Metcalf, W. W.; et al. Directed Evolution of the Nonribosomal Peptide Synthetase AdmK Generates New Andrimid Derivatives In Vivo. *Chem. Biol*. **2011**, *18*, 601–607.
192. Bozhüyük, K. A. J.; Fleischhacker, F.; Linck, A.; et al. De Novo Design and Engineering of Non-Ribosomal Peptide Synthetases. *Nat. Chem*. **2018**, *10*, 275–281.
193. Ramm, S.; Krawczyk, B.; Mühlenweg, A.; et al. A Self-Sacrificing *N*-Methyltransferase Is the Precursor of the Fungal Natural Product Omphalotin. *Angew. Chem. Int. Ed*. **2017**, *56*, 9994–9997.
194. Hetrick, K. J.; Walker, M. C.; van der Donk, W. A. Development and Application of Yeast and Phage Display of Diverse Lanthipeptides. *ACS Cent. Sci*. **2018**, *4*, 458–467.

195. Burkhart, B. J.; Kakkar, N.; Hudson, G. A.; et al. Chimeric Leader Peptides for the Generation of Non-Natural Hybrid RiPP Products. *ACS Cent. Sci*. **2017**, *3*, 629–638.
196. Crook, N.; Abatemarco, J.; Sun, J.; et al. In Vivo Continuous Evolution of Genes and Pathways in Yeast. *Nat. Commun*. **2016**, *7*, 13051.
197. Nishida, K.; Arazoe, T.; Yachie, N.; et al. Targeted Nucleotide Editing Using Hybrid Prokaryotic and Vertebrate Adaptive Immune Systems. *Science* **2016**, *353*, 1248.
198. Ravikumar, A.; Arzumanyan, G. A.; Obadi, M. K. A.; et al. Scalable, Continuous Evolution of Genes at Mutation Rates above Genomic Error Thresholds. *Cell* **2018**, *175*, 1946–1957.e13.
199. Halperin, S. O.; Tou, C. J.; Wong, E. B.; et al. CRISPR-Guided DNA Polymerases Enable Diversification of All Nucleotides in a Tunable Window. *Nature* **2018**, *560*, 248–252.
200. Dang, T.; Süssmuth, R. D. Bioactive Peptide Natural Products as Lead Structures for Medicinal Use. *Acc. Chem. Res*. **2017**, *50*, 1566–1576.
201. Campos, K. R.; Coleman, P. J.; Alvarez, J. C.; et al. The Importance of Synthetic Chemistry in the Pharmaceutical Industry. *Science* **2019**, *363*, eaat0805.
202. Moir, M.; Danon, J. J.; Reekie, T. A.; et al. An Overview of Late-Stage Functionalization in Today's Drug Discovery. *Expert Opin. Drug Discov*. **2019**, *14*, 1137–1149.
203. Zhao, J.; Zhang, D.; Zhang, W.; et al. Highly Selective MERTK Inhibitors Achieved by a Single Methyl Group. *J. Med. Chem*. **2018**, *61*, 10242–10254.
204. Schönherr, H.; Cernak, T. Profound Methyl Effects in Drug Discovery and a Call for New C–H Methylation Reactions. *Angew. Chem. Int. Ed*. **2013**, *52*, 12256–12267.
205. Wilcken, R.; Zimmermann, M. O.; Lange, A.; et al. Principles and Applications of Halogen Bonding in Medicinal Chemistry and Chemical Biology. *J. Med. Chem*. **2013**, *56*, 1363–1388.
206. Blakemore, D. C.; Castro, L.; Churcher, I.; et al. Organic Synthesis Provides Opportunities to Transform Drug Discovery. *Nat. Chem*. **2018**, *10*, 383–394.
207. Fasan, R.; Jennifer Kan, S. B.; Zhao, H. A Continuing Career in Biocatalysis: Frances H. Arnold. *ACS Catal*. **2019**, *9*, 9775–9788.
208. Arnold, F. H. Directed Evolution: Bringing New Chemistry to Life. *Angew. Chem. Int. Ed*. **2018**, *57*, 4143–4148.
209. Reetz, M. T.; Bocola, M.; Carballeira, J. D.; et al. Expanding the Range of Substrate Acceptance of Enzymes: Combinatorial Active-Site Saturation Test. *Angew. Chem. Int. Ed*. **2005**, *44*, 4192–4196
210. Markel, U.; Essani, K. D.; Besirlioglu, V.; et al. Advances in Ultrahigh-Throughput Screening for Directed Enzyme Evolution. *Chem. Soc. Rev*. **2020**, *49*, 233–262.
211. Li, G.; Dong, Y.; Reetz, M. T.; et al. Can Machine Learning Revolutionize Directed Evolution of Selective Enzymes? *Adv. Synth. Catal*. **2019**, *361*, 2377–2386.
212. Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod*. **2020**, *83* 770–803.
213. Newman, D. J.; Cragg, G. M. Natural Products as Sources of New Drugs from 1981 to 2014. *J. Nat. Prod*. **2016**, *79*, 629–661.
214. Zhang, K.; Shafer, B. M.; Demars, M. D.; et al. Controlled Oxidation of Remote sp3 C–H Bonds in Artemisinin via P450 Catalysts with Fine-Tuned Regio- and Stereoselectivity. *J. Am. Chem. Soc*. **2012**, *134*, 18695–18704.
215. Kolev, J. N.; O'Dwyer, K. M.; Jordan, C. T.; et al. Discovery of Potent Parthenolide-Based Antileukemic Agents Enabled by Late-Stage P450-Mediated C–H Functionalization. *ACS Chem. Biol*. **2013**, *9*, 164–173.
216. Le-Huu, P.; Heidt, T.; Claasen, B.; et al. Chemo-, Regio-, and Stereoselective Oxidation of the Monocyclic Diterpenoid β-Cembrenediol by P450 BM3. *ACS Catal*. **2015**, *5*, 1772–1780.
217. Petrović, D.; Bokel, A.; Allan, M.; et al. Simulation-Guided Design of Cytochrome P450 for Chemo- and Regioselective Macrocyclic Oxidation. *J. Chem. Inf. Model*. **2018**, *58*, 848–858.
218. Le-Huu, P.; Rekow, D.; Krüger, C.; et al. Chemoenzymatic Route to Oxyfunctionalized Cembranoids Facilitated by Substrate and Protein Engineering. *Chem. Eur. J*. **2018**, *24*, 12010–12021.
219. Lowell, A. N.; DeMars, M. D.; Slocum, S. T.; et al. Chemoenzymatic Total Synthesis and Structural Diversification of Tylactone-Based Macrolide Antibiotics through Late-Stage Polyketide Assembly, Tailoring, and C−H Functionalization. *J. Am. Chem. Soc*. **2017**, *139*, 7913–7920.
220. Li, Y.; Qin, B.; Li, X.; et al. Selective Oxidations of Cyperenoic Acid by Slightly Reshaping the Binding Pocket of Cytochrome P450 BM3. *ChemCatChem* **2018**, *10*, 559–565.
221. Loskot, S. A.; Romney, D. K.; Arnold, F. H.; et al. Enantioselective Total Synthesis of Nigelladine A via Late-Stage C–H Oxidation Enabled by an Engineered P450 Enzyme. *J. Am. Chem. Soc*. **2017**, *139*, 10196–10199.
222. Zhang, K.; Shafer, B. M.; Demars, M. D.; et al. Controlled Oxidation of Remote sp3 C–H Bonds in Artemisinin via P450 Catalysts with Fine-Tuned Regio- and Stereoselectivity. *J. Am. Chem. Soc*. **2012**, *134*, 18695–18704.
223. Kinne, M.; Poraj-Kobielska, M.; Aranda, E.; et al. Regioselective Preparation of 5-Hydroxypropranolol and 4′-Hydroxydiclofenac with a Fungal Peroxygenase. *Bioorg. Med. Chem. Lett*. **2009**, *19*, 3085–3087.
224. Zhang, J.; Huang, X.; Zhang, R. K.; et al. Enantiodivergent α-Amino C–H Fluoroalkylation Catalyzed by Engineered Cytochrome P450s. *J. Am. Chem. Soc*. **2019**, *141*, 9798–9802.
225. Zhang, R. K.; Chen, K.; Huang, X.; et al. Enzymatic Assembly of Carbon–Carbon Bonds via Iron-Catalysed sp3 C–H Functionalization. *Nature* **2019**, *565*, 67–72.
226. Yang, Y.; Cho, I.; Qi, X.; et al. An Enzymatic Platform for the Asymmetric Amination of Primary, Secondary and Tertiary C(*sp*3)–H bonds. *Nat. Chem*. **2019**, *11*, 987–993.
227. Neugebauer, M. E.; Sumida, K. H.; Pelton, J. G.; et al. A Family of Radical Halogenases for the Engineering of Amino-Acid-Based Products. *Nat. Chem. Biol*. **2019**, *15*, 1009–1016.
228. Schadt, S.; Bister, B.; Chowdhury, S. K.; et al. A Decade in the MIST: Learnings from Investigations of Drug Metabolites in Drug Development under the Metabolites in Safety Testing Regulatory Guidance. *Drug Metab. Dispos*. **2018**, *46*, 865–878.

229. Bertz, R. J.; Granneman, G. R. Use of In Vitro and In Vivo Data to Estimate the Likelihood of Metabolic Pharmacokinetic Interactions. *Clin. Pharmocokinet*. **1997**, *32*, 210–258.
230. Evans, W. E.; Relling, M. V. Pharmacogenomics: Translating Functional Genomics into Rational Therapeutics. *Science* **1999**, *286*, 487–491.
231. Di Nardo, G.; Gilardi, G. Optimization of the Bacterial Cytochrome P450 BM3 System for the Production of Human Drug Metabolites. *Int. J. Mol. Sci*. **2012**, *13*, 15901–15924.
232. Butler, C. F.; Peet, C.; Mason, A. E.; et al. Key Mutations Alter the Cytochrome P450 BM3 Conformational Landscape and Remove Inherent Substrate Bias. *J. Biol. Chem*. **2013**, *288*, 25387–25399.
233. Ren, X.; Yorke, J. A.; Taylor, E.; et al. Drug Oxidation by Cytochrome P450BM3: Metabolite Synthesis and Discovering New P450 Reaction Types. *Chem. Eur. J*. **2015**, *21*, 15039–15047.
234. Otey, C. R.; Bandara, G.; Lalonde, J.; et al. Preparation of Human Metabolites of Propranolol Using Laboratory-Evolved Bacterial Cytochromes P450. *Biotechnol. Bioeng*. **2006**, *93*, 494–499.
235. Gomez de Santos, P.; Cañellas, M.; Tieves, F.; et al. Selective Synthesis of the Human Drug Metabolite 5′-Hydroxypropranolol by an Evolved Self-Sufficient Peroxygenase. *ACS Catal*. **2018**, *8*, 4789–4799.
236. de Santos, P. G.; Cervantes, F. V.; Tieves, F.; et al. Benchmarking of Laboratory Evolved Unspecific Peroxygenases for the Synthesis of Human Drug Metabolites. *Tetrahedron* **2019**, *75*, 1827–1831.
237. Wang, Y.; Lan, D.; Durrani, R.; et al. Peroxygenases En Route to Becoming Dream Catalysts. What Are the Opportunities and Challenges? *Curr. Opin. Chem. Biol*. **2017**, *37*, 1–9.
238. Chen, K.; Arnold, F. H. Engineering New Catalytic Activities in Enzymes. *Nat. Catal*. **2020**, *3*, 203–213.
239. Wang, Z. J.; Renata, H.; Peck, N. E.; et al. Improved Cyclopropanation Activity of Histidine-Ligated Cytochrome P450 Enables the Enantioselective Formal Synthesis of Levomilnacipran. *Angew. Chem. Int. Ed*. **2014**, *53*, 6810–6813.
240. Xu, Z.; Yang, Z.; Liu, Y.; et al. Halogen Bond: Its Role Beyond Drug–Target Binding Affinity for Drug Discovery and Development. *J. Chem. Inf. Model*. **2014**, *54*, 69–78.
241. Latham, J.; Brandenburger, E.; Shepherd, S. A.; et al. Development of Halogenase Enzymes for Use in Synthesis. *Chem. Rev*. **2018**, *118*, 232–269.
242. Poor, C. B.; Andorfer, M. C.; Lewis, J. C. Improving the Stability and Catalyst Lifetime of the Halogenase RebH by Directed Evolution. *ChemBioChem* **2014**, *15*, 1286–1289.
243. Payne, J. T.; Poor, C. B.; Lewis, J. C. Directed Evolution of RebH for Site-Selective Halogenation of Large Biologically Active Molecules. *Angew. Chem. Int. Ed*. **2015**, *54*, 4226–4230.
244. Menon, B. R. K.; Brandenburger, E.; Sharif, H. H.; et al. RadH: A Versatile Halogenase for Integration into Synthetic Pathways. *Angew. Chem. Int. Ed*. **2017**, *56*, 11841–11845.
245. Hillwig, M. L.; Zhu, Q.; Ittiamornkul, K.; et al. Discovery of a Promiscuous Non-Heme Iron Halogenase in Ambiguine Alkaloid Biogenesis: Implication for an Evolvable Enzyme Family for Late-Stage Halogenation of Aliphatic Carbons in Small Molecules. *Angew. Chem. Int. Ed*. **2016**, *55*, 5780–5784.
246. Hayashi, T.; Ligibel, M.; Sager, E.; et al. Evolved Aliphatic Halogenases Enable Regiocomplementary C–H Functionalization of an Added-Value Chemical. *Angew. Chem. Int. Ed*. **2019**.
247. Duewel, S.; Schmermund, L.; Faber, T.; et al. Directed Evolution of an FeII-Dependent Halogenase for Asymmetric C (sp^3)-H Chlorination. *ACS Catal*. **2020**, *10*, 1272–1277.
248. O'Hagan, D.; Deng, H. Enzymatic Fluorination and Biotechnological Developments of the Fluorinase. *Chem. Rev*. **2015**, *115*, 634–649.
249. Thompson, S.; Zhang, Q.; Onega, M.; et al. A Localized Tolerance in the Substrate Specificity of the Fluorinase Enzyme Enables Last-Step 18F Fluorination of a RGD Peptide under Ambient Aqueous Conditions. *Angew. Chem. Int. Ed*. **2014**, *53*, 8913–8918.
250. Sun, H.; Yeo, W. L.; Lim, Y. H.; et al. Directed Evolution of a Fluorinase for Improved Fluorination Efficiency with a Non-Native Substrate. *Angew. Chem. Int. Ed*. **2016**, *55*, 14277–14280.
251. Yeo, W. L.; Chew, X.; Smith, D. J.; et al. Probing the Molecular Determinants of Fluorinase Specificity. *Chem. Commun*. **2017**, *53*, 2559–2562.
252. Atzrodt, J.; Derdau, V.; Kerr, W. J.; et al. Deuterium- and Tritium-Labelled Compounds: Applications in the Life Sciences. *Angew. Chem. Int. Ed*. **2018**, *57*, 1758–1784.
253. Isin, E. M.; Elmore, C. S.; Nilsson, G. N.; et al. Use of Radiolabeled Compounds in Drug Metabolism and Pharmacokinetic Studies. *Chem. Res. Toxicol*. **2012**, *25*, 532–542.
254. Liao, C.; Seebeck, F. P. *S*-Adenosylhomocysteine as a Methyl Transfer Catalyst in Biocatalytic Methylation Reactions. *Nat. Catal*. **2019**, *2*, 696–701.
255. Frey, R.; Hayashi, T.; Buller, R. M. Directed Evolution of Carbon–Hydrogen Bond Activating Enzymes. *Curr. Opin. Biotech*. **2019**, *60*, 29–38.
256. Devine, P. N.; Howard, R. M.; Kumar, R.; et al. Extending the Application of Biocatalysis to Meet the Challenges of Drug Development. *Nat. Rev. Chem*. **2018**, *2*, 409–421.
257. Kalos, M.; Levine, B. L.; Porter, D. L.; et al. T Cells with Chimeric Antigen Receptors Have Potent Antitumor Effects and Can Establish Memory in Patients with Advanced Leukemia. *Sci. Transl. Med*. **2011**, *3*, 95ra73.
258. Caliendo, F.; Dukhinova, M.; Siciliano, V. Engineered Cell-Based Therapeutics: Synthetic Biology Meets Immunology. *Front. Bioeng. Biotechnol*. **2019**, *7*, 43.
259. Shao, J.; Xue, S.; Yu, G.; et al. Smartphone-Controlled Optogenetically Engineered Cells Enable Semiautomatic Glucose Homeostasis in Diabetic Mice. *Sci. Transl. Med*. **2017**, *9*, eaal2298/1–eaal2298/13.
260. Krawczyk, K.; Xue, S.; Buchmann, P.; et al. Electrogenetic Cellular Insulin Release for Real-Time Glycemic Control in Type 1 Diabetic Mice. *Science* **2020**, *368*, 993–1001.

261. Chowdhury, S.; Castro, S.; Coker, C.; et al. Programmable Bacteria Induce Durable Tumor Regression and Systemic Antitumor Immunity. *Nat. Med*. **2019**, *25*, 1057–1063.
262. Leschner, S.; Weiss, S. Salmonella—Allies in the Fight Against Cancer. *J. Mol. Med*. **2010**, *88*, 763–73.
263. Lehouritis, P.; Hogan, G.; Tangney, M. Designer Bacteria as Intratumoural Enzyme Biofactories *Adv. Drug Deliv. Rev*. **2017**, *118*, 8–23.
264. Braat, H.; Rottiers, P.; Hommes, D. W.; et al. A Phase I Trial with Transgenic Bacteria Expressing Interleukin-10 in Crohn's Disease. *Clin. Gastroenterol. Hepatol*. **2006**, *4*, 754–759.
265. Takiishi, T.; Cook, D. P.; Korf, H.; et al. Reversal of Diabetes in NOD Mice by Clinical-Grade Proinsulin and IL-10-Secreting *Lactococcus lactisin* Combination with Low-Dose Anti-CD3 Depends on the Induction of Foxp3-Positive T cells. *Diabetes* **2017**, *66*, 448–459.
266. Zheng, J. H.; Nguyen, V. H.; Jiang, S.-N.; et al. Two-Step Enhanced Cancer Immunotherapy with Engineered *Salmonella typhimurium* Secreting Heterologous Flagellin. *Sci. Transl Res*. **2017**, *9*, eaak9537.
267. Pedrolli, D. B.; Ribeiro, N. V.; Squizato, P. N.; et al. Engineering Microbial Living Therapeutics: the Synthetic Biology Toolbox. *Trends Biotechnol*. **2019**, *37*, 100–115.
268. Naydich, A. D.; Nangle, S. N.; Bues, J. J.; et al. Synthetic Gene Circuits Enable Systems-Level Biosensor Trigger Discovery at the Host-Microbe Interface. *MSystems* **2019**, *4*, e00125-19.
269. Kotopka, B. J.; Smolke, C. D. Model-Driven Generation of Artificial Yeast Promoters. *Nat. Commun*. **2020**, *11*, 2113.
270. Radivojević, T.; Costello, Z.; Workman, K.; et al. A Machine Learning Automated Recommendation Tool for Synthetic Biology. *Nat. Commun*. **2020**, *11*, 4879.
271. Klein, J; Heal, J. R.; Hamilton, W. D. O.; et al. Yeast Synthetic Biology Platform Generates Novel Chemical Structures as Scaffolds for Drug Discovery. *ACS Synth. Biol*. **2014**, *3*, 314–323.
272. Chen, H.; Engkvist, O.; Olivecrona, M.; et al. The Rise of Deep Learning in Drug Discovery. *Drug Discov. Today* **2018**, *23*, 1241–1250.
273. Kotsias, P.-C.; Arus-Pous, J.; Chen, H.; et al. Direct Steering of De Novo Molecular Generation with Descriptor Conditional Recurrent Neural Networks. *Nat. Mach. Intell*. **2020**, *2*, 254–265.
274. Mason, D. M.; Friedensohn, S.; Weber, C. R.; et al. Deep Learning Enables Therapeutic Antibody Optimization in Mammalian Cells by Deciphering High-Dimensional Protein Sequence Space. *BioRxiv* **2019**. DOI: 10.1101/617860.