J Ind Microbiol Biotechnol
DOI 10.1007/s10295-016-1862-3

METABOLIC ENGINEERING AND SYNTHETIC BIOLOGY - ORIGINAL PAPER

Tailor-made transcriptional biosensors for optimizing microbial cell factories

Brecht De Paepe[1] · Gert Peters[1] · Pieter Coussement[1] · Jo Maertens[1] · Marjan De Mey[1]

Received: 25 August 2016 / Accepted: 30 October 2016

Abstract Monitoring cellular behavior and eventually properly adapting cellular processes is key to handle the enormous complexity of today's metabolic engineering questions. Hence, transcriptional biosensors bear the potential to augment and accelerate current metabolic engineering strategies, catalyzing vital advances in industrial biotechnology. The development of such transcriptional biosensors typically starts with exploring nature's richness. Hence, in a first part, the transcriptional biosensor architecture and the various modi operandi are briefly discussed, as well as experimental and computational methods and relevant ontologies to search for natural transcription factors and their corresponding binding sites. In the second part of this review, various engineering approaches are reviewed to tune the main characteristics of these (natural) transcriptional biosensors, i.e., the response curve and ligand specificity, in view of specific industrial biotechnology applications, which is illustrated using success stories of transcriptional biosensor engineering.

Keywords Transcriptional biosensor · Transcriptional regulation · Prokaryotes · Ligand specificity engineering · Response curve engineering

Introduction

By unravelling and rewiring the metabolism of living organisms through metabolic engineering, industrial biotechnology enables the production of a variety of enzymes, pharmaceuticals, fuels, foods, and chemical building blocks. However, due to the high intrinsic complexity of metabolic networks, the development of new robust microbial cell factories is laborious and time-consuming. Therefore, new engineering tools are required that can accelerate the classic design-build-test-learn cycle and, consequently, strain and product development [103, 167]. Recent technological advances, e.g., multiplex genome engineering and degenerate gene synthesis, already instigate the design and construction of libraries of billions of cellular variants producing an ever expanding array of chemicals and enzymes [122]. Nevertheless, screening and selecting the optimal phenotypes from these libraries as well as dynamically controlling and quantifying the in vivo synthesis of these compounds remain tedious and a main bottleneck in the biological engineering cycle [41, 44, 103, 122, 163].

In this context, nature's myriad of transcriptional regulatory systems bears the potential to augment and accelerate metabolic engineering strategies, catalyzing vital advances in industrial biotechnology. Such regulation mechanisms allow intra- and extracellular monitoring of endo- or exogenous small molecules, ions, and changes in physical parameters [93]. Furthermore, these in vivo monitoring systems can be reprogrammed into biosensors which broadcast output signals of choice and, therefore, enable researchers to monitor and even fine-tune directly what transpires inside a living cell. In prokaryotes, the majority of signal transduction occurs via widely distributed one-component regulatory mechanisms consisting of allosteric DNA-binding transcription factors which display a great diversity of small-molecule specificities [35, 154]. Accordingly, metabolic engineers and synthetic biologists started reorganizing these prokaryotic transcriptional regulatory circuits into

✉ Marjan De Mey
Marjan.DeMey@UGent.be

[1] Department of Biochemical and Microbial Technology, Ghent University, Coupure Links 653, 9000 Ghent, Belgium

Published online: 11 November 2016

transcriptional biosensors. In this context, we also refer to the advances in the field of CRISPR-based activation and repression for biosensor circuit development [11, 34, 156, 159, 171]. These biosensors are typically constructed to perform one of three functions depending on what genes the biosensor circuit controls: optical readouts through, e.g., fluorescent protein expression, dynamic pathway control through, e.g., biosynthetic pathway regulation and selection strategies through, e.g., antibiotic resistance regulation [83].

This set of applications brings forth an enormous surge in high-throughput metabolic engineering capacity, closing the gap between generating genetic diversity and, subsequently, screening, selecting, quantifying and controlling these biological libraries [44]. Ideally, transcriptional biosensors can be designed to gratify any predefined response curve and any small-molecule specificity, i.e., the two fundamental characteristics defining a biosensor. However, like all biological systems, transcription factor-promoter pairs have evolved to fulfill a specific regulatory function in their native metabolic network. Despite the great potential of transcriptional biosensors, the lack of biosensor engineering principles currently limits their versatility necessary in metabolic engineering. Preceded by an overview of the basic architecture and regulatory mechanisms of transcriptional biosensors, this review focuses on the current efforts and future concepts that enable the engineering of a biosensor's ligand specificity, on the one hand, and engineering a biosensor's response curve, on the other hand. Prokaryotic single input/single output biosensor circuits are the main subject of this review, as these are the most implemented biological circuits for biosensor design due to their modularity.

Transcriptional biosensors: a beginner's guide

A transcriptional biosensor consists of the following: (1) a transcription factor (TF) that, preferentially specifically detects a ligand molecule (input) with its ligand-binding domain (LBD) which results in a conformational change in its DNA-binding domain (DBD), (2) a transcription factor binding site (TFBS) on the DNA level, on which the TF binds, modulating RNA polymerase (RNAP)-promoter association, which, in turn, initiates transcription of (3) an output signal of choice (see Fig. 1a). Despite the seemingly simplistic nature of such a biological circuit, these prokaryotic one-component systems comprise more than 50 different TF families [110, 120]. Each of these TF families has differing ligand- and DNA-binding domains and, consequently, demonstrates, more or less, differing ligand and TFBS specificities and differing mechanisms for ligand recognition and transcriptional regulation [35, 45].

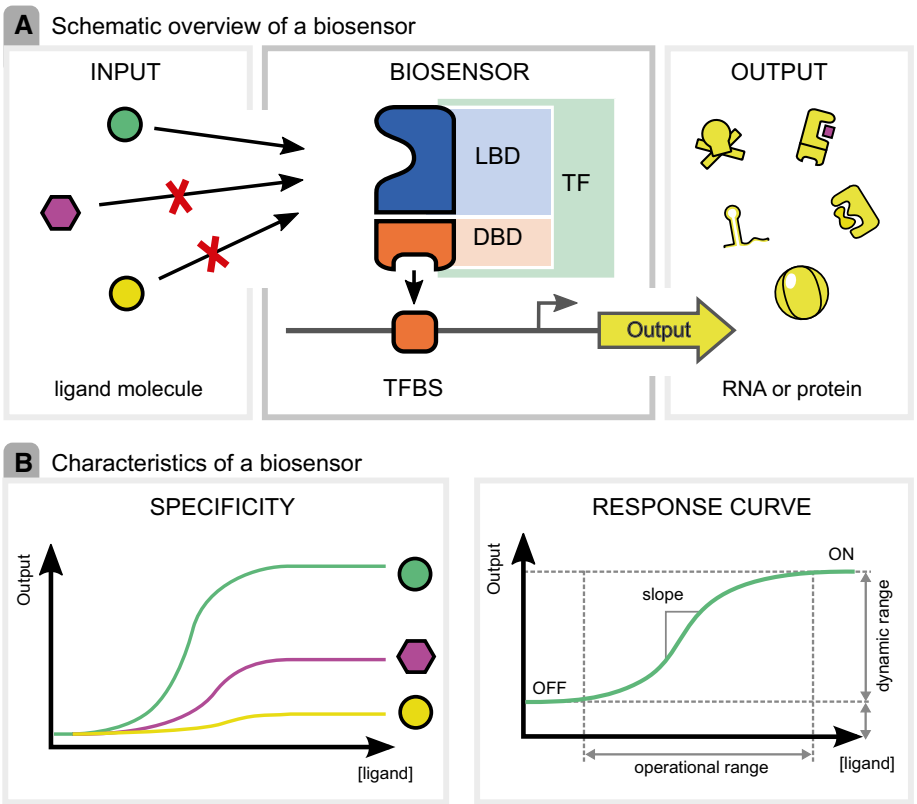

Fig. 1 **a** Schematic representation of the biosensor architecture, with specific small ligand molecules as *input* and either RNA or protein as *output*, e.g., fluorescent proteins, RNA devices, enzymes, toxic genes, orthogonal operators, or β-galactosidases. **b** Two defining characteristics of a biosensor: the biosensor specificity determined by the sensitivity profile for different ligand molecules and the *response curve* which is defined by the *slope of the curve* and the operational and dynamic range

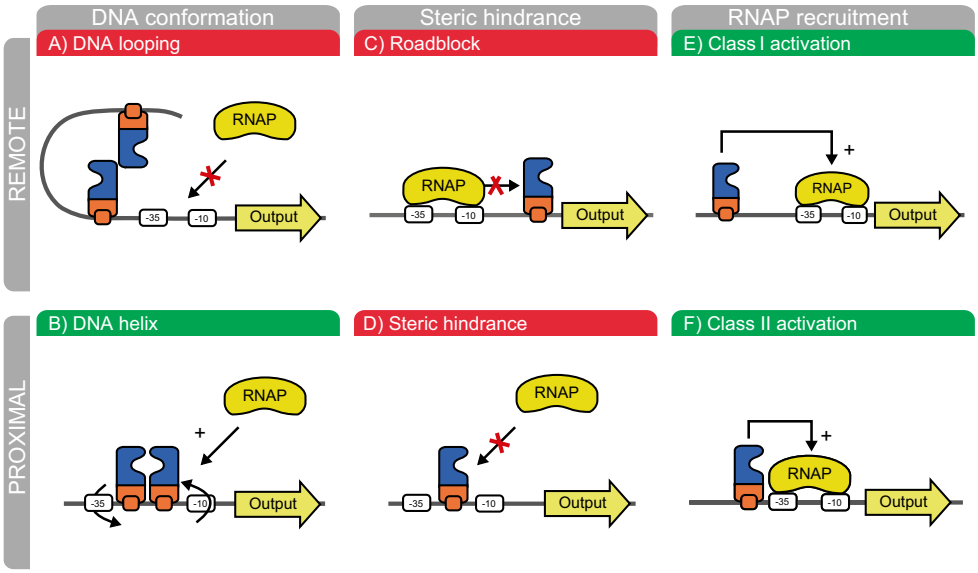

Fig. 2 Schematic representation of the six different modes of action of transcriptional regulation classified based on the three underlying regulatory mechanisms, DNA conformation, steric hindrance and RNA polymerase (RNAP) recruitment and the distance of effect, from an engineering perspective, being either proximal (between −60 and +60 relative to the transcription start site, TSS) or remote (*outside* −60 or +60 relative to the TSS)

The two fundamental characteristics defining a transcriptional biosensor, from an engineering point of view, are the specificity of a TF for a certain ligand molecule and the response curve, describing the output signal in function of the ligand concentration (see Fig. 1b). The specificity of a biosensor for a certain ligand molecule can be defined as the relative difference between the increase in output signal upon binding of this ligand in comparison with binding other potential ligand molecules from a predefined set of interest (see Fig. 1b). Logically, it should be noted that biosensors with a broad specificity range can still be used for certain applications if none of the other potential ligand molecules are present in the cell. The input/output relation is represented by the biosensor response curve, which, in itself, is defined by three main parameters: (1) the operational range, including the half-maximal threshold; (2) the dynamic range; and (3) the sensitivity or slope of the response curve (see Fig. 1b) [13, 14, 19, 125]. The operational range is the range of ligand concentrations over which the output signal exhibits significant changes. The half-maximal threshold is defined as the ligand concentration at which half of the maximal output signal is reached. The dynamic range is the range defined by the highest measurable output signal, or ON state, and the lowest measurable output signal, or OFF state, in correspondence with the operational range [125]. Leaky expression is observed when this OFF state already demonstrates transcriptional initiation when no ligand is present, or, in the case of repression, when this OFF state cannot be achieved for any level of ligand added [4]. The sensitivity for the ligand molecule, or slope of the response curve, is largely the result of the cooperativity of ligand binding by TFs and results in an either more digital- or more analog-like behavior of the biosensor [145]. Cooperative binding arises when, if a ligand is already bound by a TF, the further addition of ligand molecules alters the affinity of the TF for this ligand, either positively or negatively [4, 142]. This concept only applies for TFs that form di- or oligomeric complexes and bind the corresponding ligand molecule as such.

These biosensor response curves can be fit with the widely used Hill equation to provide a semi-empirical approach capable of providing insight in cooperativity and ligand affinity of the TF and operational and dynamic range characteristics [4, 63, 158]. Alternatively, thermodynamic models can be used to fit the generated response data based on probability dynamics of TF- and RNAP bindings to the DNA [13, 14]. These models can predict transcriptional regulation response for several classes of regulatory mechanisms and, vice versa, reveal underlying mechanisms for a given data set of unknown transcriptional control [13, 14]. In the following, the predominant types of TFBS, DBDs, and LBDs, found in nature, are considered. Subsequently, a concise overview is given of different underlying regulatory mechanisms of these ligand-dependent regulatory systems (see Fig. 2).

Transcriptional biosensor architecture

Transcription factor binding sites

TFBSs are specific DNA sequences which demonstrate an affinity for a specific TF. These binding sites can vary in length, base pair motif and position relative to the consensus core promoter and to additional TFBSs, if present. Commonly, TFBS are between 12 and 30 base pairs (bp) long and often possess an intrinsic symmetry due to homodimeric or homo-oligomeric binding of their corresponding TFs. Consequently, direct repeats and

palindromic sequences (inverted repeats) are the most common TFBS structure [62]. Typically, the half-sites of direct repeats contain a spacing sequence which spans a multiplicity of one complete DNA helix turn (10.5 nucleotides, nt) to ensure that the two half-sites are located at the same helical face and, therefore, ensure optimal TF binding [120]. Similarly, TF–RNAP interaction is generally optimal when the bound TF is located at the same side of the DNA helix as the bound RNAP. Accordingly, the nucleotide spacer sequences between TFBS and core promoter, for instance, with most class I activators (see Sect. "Transcriptional regulatory mechanisms" and Fig. 2e), also match complete helical turns (10.5 nt) [8, 62]. In addition, more complex compositions of TFBSs may occur when dealing with homo-oligomeric TFs, e.g., with transcriptional regulation through DNA looping [120]. Within a genome, several operons can be controlled by one unique TF which binds on variations of the corresponding consensus TFBS sequence upstream of the respective operon locations. This degeneracy of TFBSs enables different TF affinities and, consequently, adaptable levels of control for distinct genes [7]. An artificial distinction may be made in the location of the TFBSs, being situated either at the core promoter (proximal, between −60 and +60 relative to the transcription start site, TSS) or up- or downstream of the core promoter (remote, outside −60 or +60). From an engineering point of view, this distinction is of importance when applying mutagenesis techniques to specifically target either the core promoter, TFBSs or both (see Sects. "Engineering biosensor ligand specificity" and "Engineering the biosensor response curve"). Notably, binding sites as distant as −243 bp have been reported [166].

DNA-binding domains

The DBD of a TF specifically binds with a corresponding TFBS through non-covalent interactions between the protein side chains and the exposed base pairs within the DNA grooves [17]. It should be noted that TFs with structurally similar DBD motifs may demonstrate different mechanisms of DNA binding, as observed with different members of the TetR family [35]. In addition, within a specific TF family, the position of the DBD in the protein is usually conserved, whereas across different TF families, the DBD may vary from an N-terminal, over a central, to a C-terminal position. The position of the DBD is correlated with the mechanism of transcriptional regulation for that particular TF/TFBS pair. As such, N-terminal DBDs are mostly observed in repressors, whereas C-terminal DBDs are mostly observed in activators [35, 111]. Despite this trend, there is no apparent functional reason for the relation between DBD location and the TF being either an activator or a repressor, which suggests evolutionary motives. In addition, the large LysR family demonstrates both activating and repressing mechanisms while having a clear, conserved N-terminal DBD [111]. The majority of DBDs has one of three basic protein motifs, with the winged helix-turn-helix (HTH) motif the most commonly observed [45, 110]. First, the canonical HTH motif, consists of three helices in an open configuration with the characteristic turn between the second helix, which has a stabilizing function, and the third helix, which recognizes the TFBS in the major DNA groove [5, 17]. Noteworthy, the HTH motif is not stable as a stand-alone protein domain and, therefore, is always part of a larger DBD to enable proper functionality. Furthermore, TFBS recognition can also occur by interacting with other regions of the DBD [106]. Second, the winged HTH motif (wHTH) comprises an elaboration on the classic HTH motif with the addition of a β-strand hairpin unit (wing) at the C-terminus of the DBD [5]. Third, the less common ribbon-helix-helix (RHH) motif uses a β-sheet, instead of the classic α-helix, to interact with the TFBS [45, 135].

Ligand-binding domains

The LBD of a TF enables small-molecule dependency of transcriptional regulation through the conformational transduction of the ligand-binding input signal. This signal is transmitted either to the DBD, to alter DNA binding, or, in some cases, to an RNAP-activating domain, to alter RNAP recruitment, accordingly [45]. TFs bind a wide variety of ligand molecules ranging from amino-acids, over sugars, metal ions, fatty acids and aromatic to terpenoids, flavonoids and antibiotics, among others [35, 44, 45, 109, 125]. In contrast to the highly conserved DBDs, LBDs bear a great diversity in amino acid sequence due to the immense array of chemical structures of possible ligands. Correspondingly, prediction of domain structure, let alone prediction of potential ligands, based on existing structures, remains difficult [78]. Besides a ligand-binding functionality, many TFs possess dimerization domains and some RNAP-activating domains, whether or not as a part of the LBD. In addition to direct TF-RNAP protein–protein interactions for transcriptional activation, some activators, such as members of the XylR/NtrC family, are ATP dependent and contain the AAA+ fold for ATPase activity [118]. Upon ligand binding, this functionality provides energy for the TF to reconfigure RNAP into an open conformation which, subsequently, initiates transcription [23, 161]. The majority of TFs form homodimeric or homo-oligomeric complexes to bind their respective palindromic TFBSs with each an HTH domain per monomer [45, 62]. These TFs contain a protein dimerization interface to interact with additional TF copies to achieve cooperativity which is of great importance for the sigmoid nature of the biosensor curve [120].

Transcriptional regulatory mechanisms

One-component regulatory systems generally modulate transcription via three basic mechanisms. These mechanisms result either in activation or in repression of transcription by interfering directly or indirectly with the protein/DNA interaction of the RNA polymerase (RNAP) and the core promoter sequence (see Fig. 2). In a first mechanism, TFs create conformational changes in the DNA itself either by looping it, making the promoter inaccessible for RNAP (see Fig. 2a), or reorienting the −35 and −10 consensus boxes, by twisting the DNA helix, for improved accessibility by RNAP (see Fig. 2b). In a second mechanism, exclusive to repression, TFs structurally hinder transcription either by sterically hindering the binding of RNAP to the core promoter (see Fig. 2d) or by binding at the start codon and, consequently, blocking transcription elongation as a roadblock mechanism (see Fig. 2c). The third mechanism, exclusive to activation, involves the recruitment of RNAP either by interacting with the α-subunit (αCTD) of RNAP (class I activation, see Fig. 2e) or by promoting directly the binding of the σ-subunit of RNAP to the promoter (class II activation, see Fig. 2f) [20, 62]. Class I activators bind to TFBSs upstream of the core promoter, while Class II activators bind at the −35 consensus box [21]. Each of these three basic mechanisms can be subdivided into two distinct modes of regulation based on the position of the TF/TFBS pair relative to the core promoter/RNAP pair (see Fig. 2). As mentioned earlier, TFBSs may be considered to be located either at the core promoter (proximal, see Fig. 2b, d, f) or located up- or downstream from the core promoter (remote, see Fig. 2a, c, e), for engineering purposes.

Some TFs are exclusively activators, such as LuxR, or exclusively repressors, such as LacI and TetR, while others demonstrate both functionalities depending on the target TFBSs at specific promoters, such as LysR [109, 110, 147]. Interestingly, in *Escherichia coli* (*E. coli*), these dual regulators account for almost 40% of all TFs. In addition, some TFs, such as AraC, are bound to the TFBS in the absence of the ligand molecule, repressing transcription initiation, which is relieved upon ligand binding [147]. These systems may offer a more stringent level of control with high induction ratios [55, 147].

From an engineering point of view, these six modes of regulation are of great importance to determine, on the one hand, the boundaries and, on the other hand, the degrees of freedom of the biosensor design space. Biosensor characteristics, such as spacing between multiple TFBS copies, distance from TFBS to the core promoter, TF-TFBS and TF-RNAP affinities, among many others, are all aspects that either create engineering opportunities or set strict boundaries for biosensor functionality.

Classification and mining of biosensor parts

TFs, the key components in biosensors, are classified based on amino acid sequence conservation of their DBD and identified by homology-based predictions using Hidden Markov Model (HMM) profiles from databases, such as DBD, Superfamily, and Pfam [46, 110, 120, 164, 165]. For instance, the complete set of 304 TFs in *E. coli* was classified into 78 evolutionary families each demonstrating a variable number of members and regulatory mechanisms [112]. Notably, the LysR family is the largest known family in bacteria, followed by the AraC- and TetR families [120]. Furthermore, the abundance of each of the TF families varies in different organisms [120]. A more detailed overview of the TF families and their characteristics is reviewed elsewhere [45, 106, 110, 112, 120].

Moreover, a growing number of databases are available which contain experimental information on TFs, the genes they regulate and their TFBSs, in addition to genome-wide predictions of TFBSs, TFs, and regulons [120]. For instance, while RegulonDB [49] and EcoCyc [74] provide a specific information on transcriptional regulation in *E. coli*, DBD [164], RegTransBase [30] and Prodoric [53] provide regulatory information on prokaryotes in general [62, 120]. In addition, a part database, PAMDB, was recently created which comprises data from 135 publications containing a total of 118 characterized biological circuits and 165 genetic parts in *E. coli* [66]. The availability of reported and predicted TF/TFBS pairs to detect specific ligands is critical in the development of novel biosensors. Consequently, several in silico, in vitro, and in vivo mining strategies were developed to find the correct TF for a ligand and, accordingly, find the matching TFBS for a TF.

With the aim of identifying a TF for a specific ligand, databases are available which contain information that specifically links ligand molecules with their corresponding prokaryotic TF, i.e., RegPrecise [105], RegulonDB [49], RegTransBase [30], and BioNemo [22]. Despite the fact that these databases are great tools in this search, they are still limited to the scarce ligand–TF interactions reported in the literature. Recently, a web-based tool, SensiPath, was created, which combines these databases with additional databases containing enzymatic reactions and chemical structure libraries [39]. This tool enlarges the known set of detectable compounds by screening for enzymatic transformations that convert non-detectable compounds into detectable ones, also dubbed Sensing-Enabling Metabolic Pathways (SEMPs) [80]. The concept of SEMPs was substantiated in several studies, creating indirect biosensors for lindane, toluene, salicylaldehyde, 3-hydroxypropionate, organophosphorous compounds, cocaine, hippuric acid, parathion, nitroglycerin and 2-chloro-4-nitrophenol [28, 81, 99, 123, 139, 170]. On the other hand, to relinquish the

dependence on prior research for TF discovery, Substrate-Induced Gene Expression Screening (SIGEX) enables the screening of metagenomes for substrate-induced gene expression and can be used to identify TFs, in addition to catabolic genes [153]. Using this method with benzoate as induction substrate, 58 positive clones were obtained, from which both TFs as well as catabolic genes could be identified [153].

From the TFBS perspective, the growing field of transcriptome analysis provides the necessary high-throughput tools to discover TFBS/TF interactions, e.g., DNA microarrays, mRNA sequencing and chromatin immunoprecipitation (ChIP) (discussed in more detail elsewhere, [36, 62, 72, 82, 114, 120, 140, 160]). In general, transcriptome comparisons between wild-type and TF knockout strains are used to identify TF/TFBS pairs [62]. A different approach on whole-genome transcript arrays was used to search for a farnesyl pyrophosphate-responsive (FPP, an isoprenoid precursor) promoter [36]. Here, the toxicity of FPP in *E. coli* was exploited to screen for FPP-responsive promoters based on the transcriptome differences between two strains, either expressing an FPP-catabolic enzyme or a non-catalytic variant. The identified promoters were used for dynamic control of the amorphadiene biosynthesis pathway with a doubling of microbial production as a result [36]. In a recent study, the Alon library, comprising circa 2000 different *E. coli* promoter-*gfpmut2* fusions [172], was used to screen and select for galactose-responsive and L-phenylalanine-responsive promoters by toggled fluorescence-activated cell sorting (FACS) [91]. Based on the literature, the corresponding TF was identified for an identified L-phenylalanine-responsive and, subsequently, used for the creation of a biosensor to successfully screen a mutant library for enhanced L-phenylalanine production [91]. Finally, TFBSs can also be predicted using several computational methods which use sets of DNA sequences, reported to be regulated by a certain TF, to generate TFBS motifs (reviewed in more detail elsewhere, [62, 120]). Subsequently, these motifs can be used to scan genomes for putative TFBSs to further expand the TF-TFBS repertoire [42, 62, 94, 120, 143, 151].

Biosensor applications for optimizing microbial cell factories

Transcriptional biosensors are used mainly to perform one of three key functions in the development and optimization of microbial production strains with industrial significance, namely, high-throughput screening, adaptive laboratory evolution and dynamic pathway control [44, 83, 92, 115, 125, 131, 163, 175, 176]. These three key biosensor functionalities are exemplified in the following. For a more detailed overview of these biosensor applications, the reader is referred elsewhere [83, 92, 163].

High-throughput screening

Biosensors controlling the expression of an easily screenable output, e.g., a fluorescent protein, enable high-throughput screening of mutant libraries to select for the desired phenotypes [83]. These biosensors can be used to directly quantify the intracellular or extracellular concentration of the corresponding ligand molecule, without the need for labor-intensive analysis techniques, such as gas or high-performance liquid chromatography (GC and HPLC) [41, 83]. Recently, different biosensors were developed to enable the real-time monitoring of the microbial production of 3-hydroxypropionate (3HP), acrylate, glucarate and muconate [123, 124]. In addition, several process parameters in the microbial production of 3HP were optimized using this biosensor, resulting in a 23-fold increase in production [123]. This concept of biosensor-driven high-throughput screening was also applied to screen enzymes for their capacity to degrade lignin [90]. To this end, a biosensor was developed which detects several lignin degradation products, such as ferulic acid, p-coumaric acid, and sinapic acid [90]. In a different study, an NADPH-responsive biosensor was developed to screen an alcohol dehydrogenase mutant library using FACS [138]. Here, an improved alcohol dehydrogenase variant, with an eightfold improved K_M value for its substrate 4-methyl-2-pentanone, was selected [138]. Furthermore, biosensor-driven FACS for high-throughput screening of mutant enzyme libraries was also successfully used to identify improved mutant enzymes for lysine and histidine biosynthesis [12, 132].

Adaptive laboratory evolution

Besides, biosensors can be used to control the expression of a selection marker to enrich mutant libraries with phenotypes demonstrating the desired traits. Usually, a toggled selection workflow is implemented to eliminate evolutionary escapees. This workflow typically comprises the iteration of, first, a negative selection, second, the induction of the biosynthetic pathway, and third, positive selection for the desired phenotype [115]. Such an approach was applied to the increase microbial production of naringenin and glucaric acid by 36- and 22-fold, respectively [115]. To this end, biosensors controlling the expression of a dual selector gene were developed, optimized to reduce the evolutionary escape rate, and finally, applied in up to four rounds of directed evolution with toggled selection [115]. A different approach was pursued by Chou and Keasling who developed both a tyrosine- and isopentenyl diphosphate (IPP)-responsive biosensors to directly control the expression of a mutator gene, and as such the mutation rate [29]. This adaptive laboratory evolution workflow improved tyrosine and IPP biosynthesis up to fivefold and threefold,

respectively, and led to the identification of several genes affecting these metabolite pools [29]. Furthermore, instead of using specific selector genes under the control of the biosensor of choice, a biosensor controlling the expression of a fluorescent protein was used in combination with FACS to develop an adaptive laboratory evolution workflow applying an artificial selective pressure [93]. This method enabled the identification of several L-valine producing *Corynebacterium glutamicum* variants demonstrating higher growth rates and increased titers and a three- to fourfold reduction of by-product formation [93]. A recent study demonstrated the use of biosensors to dynamically control the cell population to increase the fraction of high-performing nongenetic variants, consequently, improving the overall performance of the population [169]. This approach rewards high-performing cells in a population with a competitive advantage in cell growth as the biosensor output, e.g., tetracycline resistance or the expression of an essential gene. For example, this method was applied to tyrosine and fatty acid production which improved both threefold [169].

Dynamic pathway control

Genetic circuits can be constructed using biosensors which control the biosynthetic pathway of interest, thus enabling dynamic control [102]. For example, dynamic control of a biosynthetic pathway enables the fine-tuning of enzyme activity in accordance with the level of ligand present to counteract the toxic accumulation of the ligand [83]. This concept was applied to dynamically control the biosynthesis of fatty acids in *E. coli* for biodiesel production through the use of a newly developed fatty acid/acyl-CoA-responsive biosensor [174]. The fatty acid ethyl ester (FAEE) biosynthetic pathway was divided into three modules of which FadR, the transcription factor of choice, inhibited two modules which form by-products. This inhibition is relieved when fatty acids, which bind FadR, accumulate, as such increasing FAEE production. This dynamic pathway regulation increased production threefold to 28% of the theoretical maximum [174]. A similar approach was used to dynamically control fatty acid production by constructing a FapR-based malonyl-CoA-responsive biosensor which exercises a negative feedback on the biosynthetic pathway [84]. Here, it was demonstrated that this regulatory effect increased the fatty acid titer and productivity by 34% and 33%, respectively [84].

Engineering biosensor ligand specificity

Biosensor-driven metabolic engineering is only feasible if a TF is available which specifically detects the target molecule in a complex metabolic network without significant interference of other compounds. Therefore, as an alternative to exploring nature for existing TFs, characterized TFs can be engineered to create custom specificity for the ligand of choice. In addition, biological engineers aspire the microbial production and detection of a vast array of natural compounds, whether or not native to the host's metabolism, and even new-to-nature compounds. Consequently, the creation of custom ligand specificities for known TFs is a critical bottleneck in biosensor development. In the following, the four main strategies to overcome this hurdle are described (see Fig. 3), ranging from the least to the most foreknowledge-dependent methods. First, random mutagenesis methods are discussed, followed by structure-based site-directed mutagenesis methods. Next, the creation of chimeric TFs is explained. Finally, the currently applied computational tools are discussed.

Random mutagenesis methods

Randomly mutating a TF, or parts of it, can generate mutant TF libraries which can be screened for TF variants responsive to a non-natural ligand of interest. However, completely shifting the specificity of a TF from its natural ligand molecule to a novel ligand of choice is not a straightforward process, even if both molecules have similar structures [47]. Parallel to other ligand-binding proteins, a TF is likely to mutate into a more promiscuous, ancestral TF form as a first step towards the divergence of a novel specificity [1, 48, 67, 87]. Screening mutant TF libraries for custom specificity consistently gives rise to these ancestral TFs which have a broader specificity range, incorporating the non-natural ligand of choice, rather than a complete switch in specificity. However, these mutant TFs can be further matured towards the true ligand specificity of choice using more comprehensive mutational strategies [1, 48, 70, 87, 150]. Random mutagenesis methods for TF engineering, which generally create TF variants with a broadening in specificity range, can be applied either on the complete TF amino acid sequence or on the LBD (see Fig. 3a). Targeting the complete TF sequence can reveal unexpected mutational hotspots, while targeting only the LBD may ensure the preservation of the natural DNA-binding efficiency potentially resulting in higher mutational success rates.

For example, the complete amino acid sequence of the quorum-sensing LuxR TF from *Vibrio fischeri* was targeted by error-prone PCR (epPCR) and DNA shuffling techniques to examine and alter the acyl-homoserine lactone (acyl-HSL) specificity and sensitivity [31]. Using this workflow, LuxR variants were generated that responded to a broader range of acyl-HSLs while retaining the sensitivity for the natural ligand molecule [31]. This study was expanded with a modified positive/negative screening

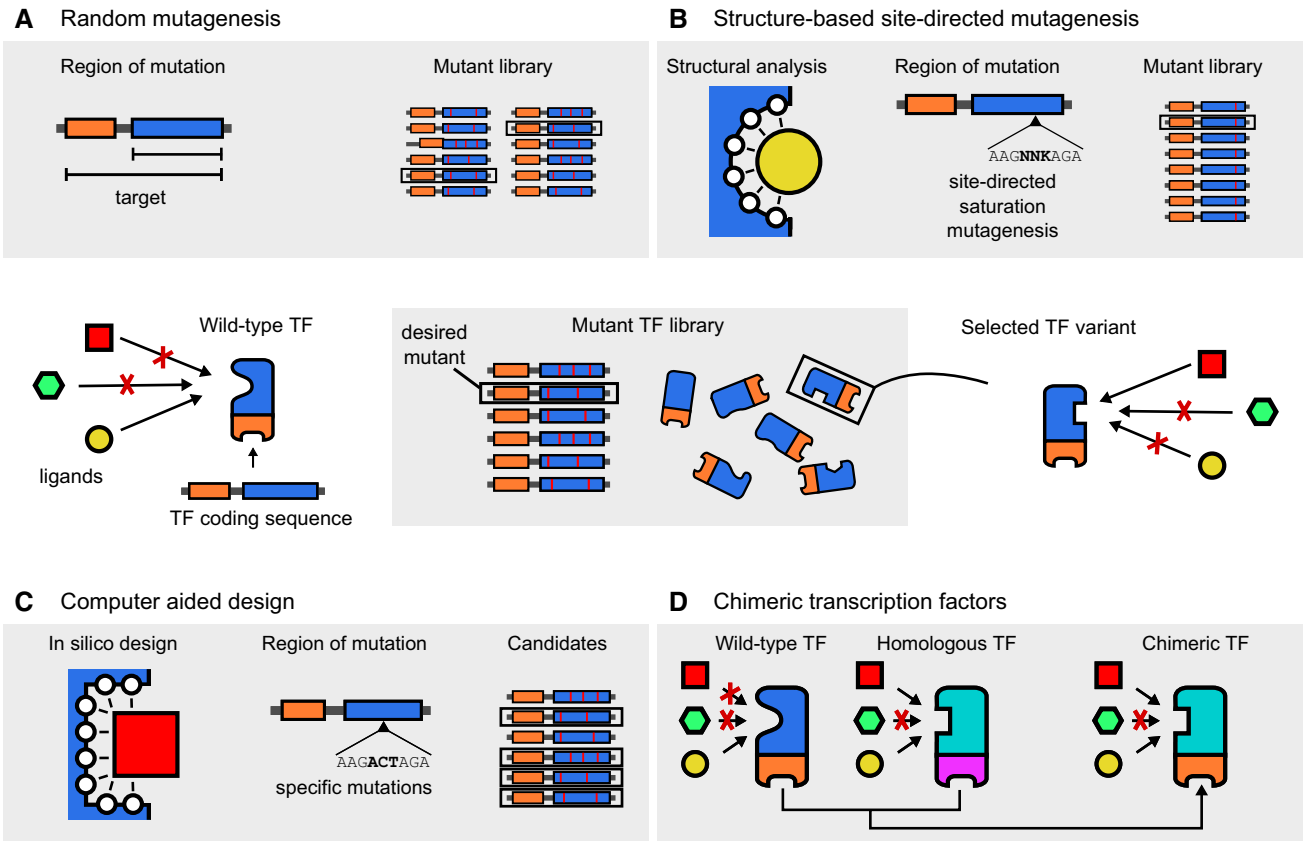

Fig. 3 Overview of the four main strategies to create transcription factors (TFs) with custom ligand specificity profiles: **a** random mutagenesis targeting the complete TF amino acid sequence or the ligand-binding domain (LBD, *blue*), **b** structure-based mutagenesis targeting specific amino-acids in, e.g., the ligand-binding pocket, **c** computer-aided design of TF amino acid sequences with novel ligand-binding pockets and **d** generation of chimeric transcription factors by combining the DNA-binding domain (DBD, *orange*) of a specific TF with the LBD, demonstrating a desired ligand specificity, of a homologous TF

method to eliminate mutant LuxR variants which still responded to the natural ligand molecules, consequently, narrowing down the specificity range to wanted ligand compounds only [32]. Here, random mutagenesis was applied on the N-terminal LBD, which enabled the creation of true custom ligand specificity by tipping the balance to an increased sensitivity for straight-chain acyl-HSLs while diminishing the sensitivity for unwanted compounds [32]. The flexibility of customizing the specificity of the LuxR TF was also validated through directed evolution of the complete TF for short-chain acyl-HSLs [56]. In a different study, the complete TF sequence of the alkane-responsive AlkSp TF from *Pseudomonas putida* was targeted for directed evolution-based random mutagenesis to modulate its broad specificity. Mutant TFs were successfully generated from which a biosensor was developed for specific short-chain alkane detection [119].

Instead of targeting the complete TF sequence, the C-terminal LBD of NahR, a LysR-type salicylate-responsive TF, was randomly mutated to generate TF variants with altered specificity for different aromatic compounds, in particular benzoate [24]. From this mutational library, several mutant TFs were isolated with novel specificity profiles, such as augmented benzoate and salicylamide specificity, in addition to augmented overall sensitivity. Interestingly, these specific changes were mostly due to single amino acid substitutions which may indicate the importance of these residues in ligand recognition of NahR [24]. A similar approach was undertaken to successfully enhance the sensitivity for and broaden the specificity range for 11 phenolic ligand compounds for a DmpR-based pollutant biosensor [168]. In a different study, an XylR mutant library was screened for enhanced DNT responsiveness to understand how TFs acquire novel specificities [48]. Here, both the linker and the LBD were targeted for mutagenesis. The resulting DNT-responsive mutants retained their sensitivity for their natural ligands and demonstrated a broadening in ligand specificity. Interestingly, all amino acid substitutions were located outside the binding pocket of the LBD, within protein regions responsible for conformational signal transduction to the DBD upon ligand binding [48]. Similarly, directed evolution was used by Lönneborg and coworkers

to evolve the salicylate-responsive LysR-type DntR regulator to a 2,4-dinitrotoluene (DNT)-specific variant [85]. Through directed evolution, both the linker and LBD of the TF were randomly mutated, followed by the recombination of previously gained mutations which resulted in several DntR variants with a greatly improved response for DNT and other aromatic compounds [85]. Finally, the specificity of TetR was successfully shifted from tetracycline to a tetracycline analog by four rounds of in vitro directed evolution through mutagenesis of the non-DBD parts of the TF [134]. Amino acid substitutions in the extended ligand-binding pocket were discovered which are determinative in the flexibility of ligand binding in TetR [134].

Structure-based site-directed mutagenesis methods

When the crystal structure of a TF and its ligand-binding pocket is available, site-saturation and site-directed mutagenesis can be used to target specific residues which interact directly with the ligand molecule (see Fig. 3b). As such, random or rational substitutions of specific amino-acids can optimize a binding pocket to bind a novel ligand with differing functional groups in comparison with natural ligands.

For instance, site-saturation mutagenesis was used to target a small set of five specific amino-acid positions in the ligand-binding pocket of the L-arabinose-responsive AraC TF from *E. coli* [147]. A positive/negative screening method, accelerated by FACS, was used to successfully select mutant TFs for D-arabinose induction. Although these mutant TFs did not respond to other sugars, adding multiple decoy ligands in the negative screening step could ensure the desired custom specificity in future engineering experiments. In parallel, an additional mutant library was created for the same purpose by epPCR on the LBD sequence of AraC. Despite the creation of D-arabinose-responsive mutants, these mutants did not demonstrate inducibility to the same extent as the mutants spawned by the site-directed approach [147]. In an ensuing study, it was demonstrated that the level of sensitivity, when engineering AraC for altered ligand specificity, can be fine-tuned by introducing mutations in the N-terminal arm of AraC [149]. Moreover, in three additional studies, the same AraC TF was redesigned, through site-directed mutagenesis and FACS-mediated dual screening, for mevalonate, triacetic acid lactone, and ectoine responsiveness, respectively [27, 146, 148]. From each novel AraC variant, a biosensor was constructed and applied in the high-throughput screening of biosynthetic pathway libraries, producing their respective ligand molecules, to successfully select for improved microbial production [27, 146, 148]. A similar approach was employed to redesign a previously developed TetR variant [134] to narrow down its newly gained broader specificity [61]. Using site-saturation mutagenesis to target specific residues in the ligand-binding pocket, the sensitivity for other ligand molecules was successfully reduced while maintaining the sensitivity for the ligand of choice, thus increasing genuine specificity [61]. Furthermore, site-saturation mutagenesis was used to target the linker region of the previously discussed TetR variant to demonstrate long-range mutational effects on DNA binding, signal transduction, ligand binding, and specificity [60, 61]. A more rational structure-based mutagenesis approach was used to alter the specificity of the quorum-sensing LuxR-type, TraR TF. Four amino-acid positions were identified by ligand docking to be critical in the formation of hydrogen bonds with the natural ligand molecule, an acyl-HSL. By substituting these residues with specific amino-acids, mutant TFs were created which demonstrated an increased specificity for acyl-HSLs with either a different 3-oxo substituent group or a shorter acyl-substituent chain [25].

The success of these purely structure-based techniques, to identify specificity-determining residues, can be improved by comparing existing orthologous and paralogous TFs [98]. This concept enabled the identification of several amino-acid positions in LacI, critical for both DNA- and ligand binding. Interestingly, these identified amino-acids constitute only a small fraction of all residues at these binding interfaces which accentuates the difficulty of determining critical residues with random or structure-based mutagenesis [98]. In a recent study, an analogous approach was applied to engineer MopR, a TF from the NtrC family, for a broader specificity range [118]. As such, the non-natural ligand, 2,5-dimetylphenol, which is a natural ligand for the TF DmpR of the same family, was included in the specificity range of MopR. This was accomplished by substituting two amino-acids in the ligand-binding pocket of MopR for the two corresponding amino-acids from the ligand-binding pocket of DmpR, consequently mimicking its binding region [118].

Chimeric transcription factors

Chimeric TFs can be created by recombining DBDs and LBDs from different homo- or even heterologous proteins to generate new combinations of reported functionalities (see Fig. 3d). Concurrently, new insights can be gained in inter-domain contacts and functional contributions of non-conserved TF regions [95, 96]. The advantage of this technique is that, from a holistic point of view, domain recombination is not expected to change the functionality of an independent domain. Therefore, chimeragenesis increases the predictability of novel TF designs by permitting the recombination of, for example, DBDs with highly characterized DNA-binding dynamics with non-native LBDs to create new ligand specificities in a given

biosensor framework, or vice versa [116]. From an atomistic view, chimeric TFs, with non-conserved amino-acids at the interface of LBD and DBD, could acquire different inter-domain interactions and, consequently, different functionalities. As such, amino acid substitutions in these non-conserved interfaces provide additional engineering opportunities for fine-tuning TFs [95, 96].

More than a decade ago, an XylR combinatorial/mutant library was created by shuffling LBDs originating from the homologous XylR, DmpR, and TbuT proteins, each responding to a specific set of ligand molecules [51]. Besides the native LBD coding sequences, additional single mutations were introduced into these LBDs to further increase the library size. Subsequently, this library was screened for TF variants that responded to four different non-native aromatic ligands. For each of these four molecules, TF variants could be isolated showing a clear response upon ligand addition. These novel ligand specificities were acquired by abolishing possible signal transduction constraints through chimeragenesis, by modulating the binding pocket through mutagenesis or by a combination of both [51]. A similar workflow was applied to create a specific 1,2,4-trichlorobenzene-responsive XylR variant for the development of a γ-hexachlorocyclohexane biosensor in combination with an SEMP (see earlier) [99].

More recently, a synthetic library of chimeric LacI/GalR transcriptional regulators was created by joining the LacI DBD and corresponding linker sequence to nine different LBDs from the same TF family, each binding different sugar or purine molecules [95, 96, 152]. This research demonstrated that chimeric LacI/GalR TFs retained the native ligand specificity of the respective LBD despite "mismatched" DBD–LBD interfaces, which is of great interest for chimeric biosensor design. Even the mode of transcriptional regulation was preserved, be it either activation or repression. Second, the non-conserved linker positions were targeted for mutagenesis to investigate their sequence–function relationship. These inter-domain positions displayed a significant functional importance and could be used to fine-tune repression levels. Finally, the tetramerization functionality of the wild-type LacI, which is absent in these chimeras, was mimicked by repeating the TFBSs to mediate dimer–dimer interactions. This TFBS repetition and the corresponding inter-operator spacing are additional parameters that enhance the flexibility of biosensor design [95]. This family of nine ligand-specific chimeric LacI/GalR TFs was used to build several multi-input transcriptional logic gates based on the concept that they have the same DBD, and thus bind to the same TFBS, but have a different ligand molecule as input [137].

A broader concept of chimeric modularity was implemented by Chou and Keasling to generate a synthetic TF for the isoprenoid precursor, isopentenyl diphosphate (IPP) [29]. This chimeric TF consists of three parts: (1) an enzyme that binds the ligand of choice; (2) a linker that transduces the ligand-binding signal of the enzyme part to (3) the DBD that responds to this signal by regulating RNA polymerase binding. For this, the native arabinose-responsive C-terminal LBD of the known AraC TF was replaced by an IPP isomerase, Idi, chosen for its ability to dimerize upon IPP binding. In contrast to the activating mechanism of the wild-type AraC, this chimeric TF represses transcription when IPP is supplemented. To further enhance the biosensor response dynamics (see Sect. "Engineering the biosensor response curve"), different IPP sensors were created by applying epPCR on the former chimeric TF. One of these synthetic TFs was used as part of an IPP-biosensor circuit that adaptively controlled the mutation rate in a lycopene producing strain through a feedback-regulated evolution of phenotype (FREP) system. In this way, the mutation rate and fluorescent protein production were increased if low levels of IPP were detected by the biosensor. As such, phenotypes were directed to and selected for IPP accumulation based on fluorescence with a twofold increase in production as a result [29].

Computer-aided design

Computational methods can aid the de novo design of ligand-binding pockets for custom specificity by sampling from a significantly larger in silico mutagenic space (~10^{76}) in comparison with directed evolution-based mutagenesis methods (~10^{14}) [41, 47, 70, 86]. Consequently, the TF test space is reduced leading to a more efficient screening workflow and a more valuable mutational data set to aid future TF engineering attempts. In addition, although random and structure-based site-directed mutagenesis can generate custom specificities for novel ligand molecules similar to the natural ligands (see earlier), these methods are still constrained by the structural characteristics of the native binding pocket [47, 70, 86]. Therefore, computational protein design may be favorable for the creation of custom specificities for ligand molecules dissimilar to the natural ligand molecule (see Fig. 3c) [150]. Once a TF variant has acquired a sensitivity for such a novel dissimilar ligand molecule, the specificity may be further matured by additional mutational and computational techniques [88].

For example, a combination of several computational tools was used to redesign the LBD of a 4-hydroxybenzoate-responsive TF, PobR, to create a TF variant with a shift in specificity towards the non-natural ligand, 3,4-dihydroxybenzoate [68]. This workflow circumvented the need for a reported crystal structure of this specific TF, by applying Rosetta-based comparative modeling at the LBD level [37]. As a method for computational library design, a compromise could be made between targeting a small set

of rational substitutions, on the one hand, and constructing an unwieldy large random mutagenesis library, on the other hand. As such, Rosetta-based ligand docking was used, from which a mutational library of every critical Rosetta-identified position was created [33]. Besides successfully altering the specificity of this TF, a rich data set was created of non-random mutations which can provide insights for future TF engineering attempts and for scrutinizing Rosetta-based models [68].

Computational protein design in combination with a cell-free transcription-translation (TX-TL) preliminary screening method was used to create mutational variants of the TetR-family TF, QacR, with novel ligand specificities [88]. To demonstrate the power of computational methods, the ligand of interest, vanillin, and the TF, QacR, were chosen independently from one another whereby this ligand is dissimilar in structure to the natural ligand molecules of QacR. Potential vanillin ligand-binding sites were predicted based on the crystal structure of QacR using Phoenix Match [77] from which multiple critical amino-acid positions were selected and 17 mutant sequences were computationally generated containing a variable number of specific amino acid substitutions [2, 77]. After in vitro and subsequent in vivo screening, mutants could be isolated which demonstrated an increased sensitivity for vanillin. Although further TF optimization would be preferential for biosensor application, these mutants are valuable mutational starting points which probably could not have been obtained with classic directed evolution techniques. Although the potential toxicity of mutant TFs is not accounted for, the cell-free screening method applied in [88] is a fast and reliable preliminary screening method which could be used for screening ligand/TF couples.

In a recent study, a combination of computational protein design, protein-wide site-saturation and random mutagenesis was used to generate novel LacI specificities for four different ligand molecules, gentiobiose, fucose, lactitol, and sucralose, with decreasing similarity to common LacI ligands [150]. To this end, Rosetta was used to successfully design LacI variants which responded to the three most dissimilar ligands and demonstrated similar inducibility to the wild-type LacI. In comparison, straightforward random mutagenesis by epPCR was applied to achieve the same goal resulting in either less responsive variants for fucose and lactitol or none at all for sucralose. However, it should be noted that comparable results as the Rosetta-based workflow could be achieved with epPCR but would require a much more labor-intensive approach. Furthermore, gentiobiose-responsive LacI variants could be identified by protein-wide single site-saturation mutagenesis and subsequent screening. This approach generated single amino acid substitutions in the binding pocket, dimerization interface or DBD which all led to a similar increase in gentiobiose-specificity, thus indicating that distant substitutions could be as potent in changing sensitivity as substitutions in the ligand-binding pocket. Finally, a combinatorial mutational shuffling approach was successfully used to decrease the sensitivity for IPTG while further increasing the newly gained sensitivity for the novel ligands to mature true specificity for these novel ligands [150].

Engineering the biosensor response curve

Transcriptional biosensors are developed for a wide variety of specific applications [83, 91, 175, 176] each imposing specific requirements in terms of operational and dynamic ranges (see Sect. "Transcriptional biosensors: a beginner's guide"). To tune these response curve parameters, the TF, the TFBS and the core promoter, and the global biosensor architecture can be altered. Logically, different approaches are to be preferred in accordance with the regulatory mechanism, either activation or repression, and the TF family. In the following section, the state-of-the-art engineering concepts are discussed for each of these biosensor parts (see Fig. 4).

Engineering at the transcription factor level

The TF plays the pivotal role in any biosensor system through the direct interaction with both the ligand and the TFBSs, in addition to potential interactions with TF copies and RNAP. Hence, on the protein level, mutant TFs can be created and selected which demonstrate modified interactions, thus altering the response curve of the biosensor (see Fig. 4a). On a more general level, the expression level of the TF relative to the number of the TFBSs can be modulated to tune these interactions in a more straightforward and predictable manner in comparison with more tedious protein and TFBS engineering techniques (see Fig. 4b) [158].

Creating custom transcription factor variants

Mutating the LBD of a TF can spawn TF variants with increased sensitivity for their corresponding ligand, therefore, altering the response curve (see Sect. "Engineering biosensor ligand specificity") [54, 136]. However, besides the ligand-binding affinity, also the transduction efficiency of the binding signal, the DNA-binding affinity and the dimerization efficiency influence the shape of the response curve. Consequently, TF mutagenesis methods range from random to site-directed methods, corresponding to mutational targets ranging from the complete amino acid sequence, over the DBD and non-conserved linker sequence, to specific amino-acid residues with reported

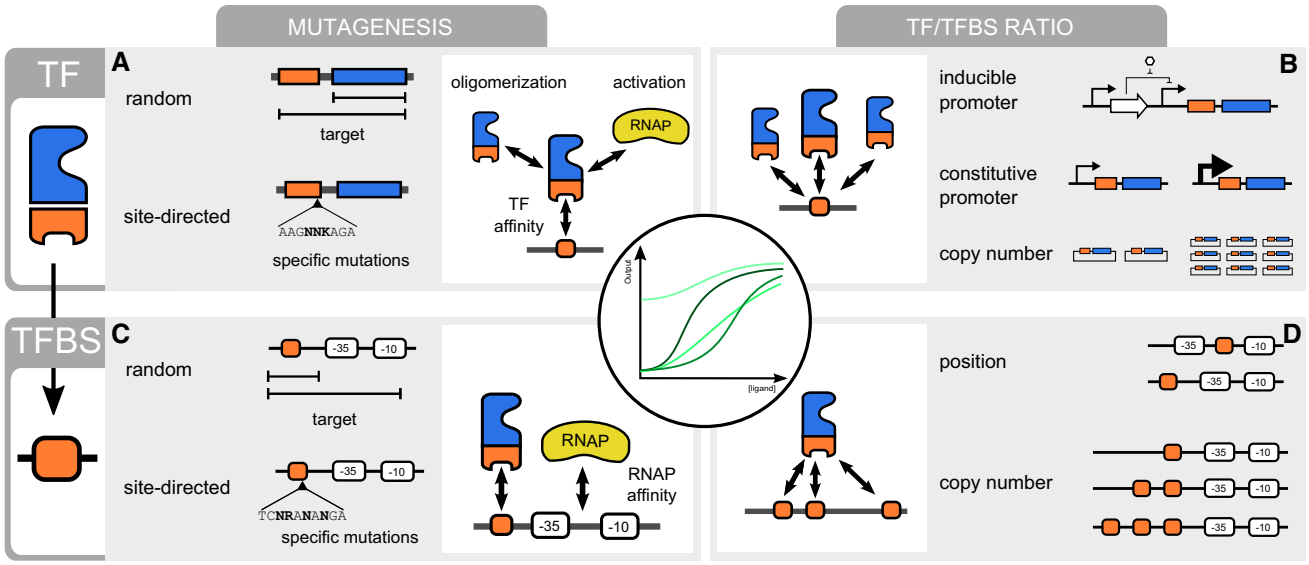

Fig. 4 Schematic overview of the discussed engineering strategies for the customization of a biosensor's response curve. The classification of these four strategies is based on, on the one hand, the applied method, being either mutagenesis or modulation of the TF/TFBS ratio and, on the other hand, being either directed to the transcription factor (TF) or the transcription factor binding sites (TFBS). In addition, for each of these four engineering strategies, a schematic representation is depicted indicating the effects of the applied strategy on the different biosensor dynamics

functionalities (see Fig. 4a). As mentioned previously (see Sect. "Engineering biosensor ligand specificity"), the specificity and sensitivity for a ligand molecule are two inseparable characteristics, and therefore, the applied mutagenesis techniques and their targets are predominantly similar. As such, in many studies aiming to alter the specificity of a TF, an altered overall sensitivity was also observed (see Sect. "Engineering biosensor ligand specificity") [24, 31, 48, 85, 168].

For example, the prevalent LacI repression system was subjected to a directed evolution study, covering the complete amino acid sequence, to increase ligand sensitivity at low IPTG concentrations [130]. The operational range was decreased and could be shifted towards lower IPTG concentrations, while the dynamic range could be increased through a set of mutations scattered across the different TF domains and its linker junction. Furthermore, a balance was observed between the mutational stabilization of, on the one hand, the inducer-bound form and, on the other hand, the DNA-bound form of the TF [130]. Similarly, directed evolution was performed on the complete DmpR gene to increase the maximum response signal of a biosensor for organophosphorus compounds [28]. Interestingly, instead of the LBD, the domain responsible for activating RNAP upon ligand binding appeared to be the mutational hotspot in this screening for improved dynamic range. Several mutants demonstrated the desired increase in operational and dynamic ranges but also a higher level of leaky expression in comparison with the wild-type TF. This may indicate a more semi-constitutive expression of the fluorescent reporter gene rather than an increase in sensitivity for the ligand molecule [28]. More intricately, the ligand sensitivity of a naphatalene/salicylate-biosensor was increased via rational engineering by implementing site-directed mutagenesis at the LBD and dimerization domain of this LysR-type TF, NahR [136]. Only two residues, previously identified to be involved in ligand binding, were targeted [24]. This approach created eight biosensor variants with differing operational and dynamic ranges, some of them displaying increased sensitivity (up to 50-fold) compared to the wild type [24, 136]. Furthermore, a comparable strategy was used to alter the sensitivity of the phenolic compound-responsive TF, DmpR [54]. First, the ligand-binding pocket for phenol and similar compounds was predicted using the ligand-docking algorithm, LIGSITE [59, 65], and a model of the N-terminal LBD domain. Second, key ligand-binding residues were identified by docking analysis and targeted for site-directed mutagenesis to successfully increase the sensitivity for these compounds [54].

Instead of targeting the ligand-binding parts of the TF, mutagenesis of the DBD or linker region can generate TF variants with various DNA-binding affinities or signal transduction efficiencies and, accordingly, different response curves of the corresponding biosensors. As such, the HTH motif of TetR was mutated to increase its DNA-binding affinity for two different TFBSs in two different studies [57, 58]. In addition to the HTH positions in direct contact with the TFBS, other mutational sites

in this domain were involved in the augmentation of the DNA-binding affinity. These residues are probably associated with adjusting the position of the HTH motif at the target DNA [57, 58]. To investigate the long-range effects of amino acid substitutions in LacI, several mutants were created based on previously reported critical residues in the linker region [144]. These specific LacI variants exhibited differing properties, such as enhanced ligand sensitivity, TFBS affinity or even both [144]. Additional critical amino-acid residues in the non-conserved linker region of LacI were predicted to be responsible for the determination of TFBS specificity based on a statistical analysis of sequence alignments from orthologous and paralogous members of the LacI/GalR family, in combination with the previously reported critical residues [96, 144]. Subsequently, these predictions were experimentally evaluated in a LacI/GalR chimeric TF containing the wild-type DBD and linker from LacI and the LBD of GalR (LLhG). The constructed chimeric LacI variants demonstrated a variety of changed functionalities ranging from enhanced dynamic range, over altered TFBS affinity, to changed TFBS specificity which emphasizes the importance of the linker region in LacI TFs and potentially other TF families [96, 173].

Finally, mutagenesis of the TF sequence can also be used to reverse the regulatory direction of a biosensor upon ligand binding. In this fashion, the TetR repressor was inverted to generate several revTetR variants which display higher DNA-binding affinity upon ligand binding instead of the contrary [133]. These mutations were clustered at the LBD/DBD interface although it was proven that even one amino-acid exchange was sufficient to reverse the TetR functionality [71, 133]. In a subsequent study, the sensitivity of this revTetR repression was further increased by site-directed mutagenesis [75].

Tuning the transcription factor expression level

The variation of the expression level of the TF, relative to the ligand concentration and the number of TFBSs, can alter the ligand/TF and TF/TFBS interplay and, consequently, the response curve of the corresponding biosensor [52]. The expression level of a TF can be tuned via, roughly, three basic approaches (see Fig. 4b). First, the TF gene can be incorporated in an additional inducible regulatory system to enable tuning of TF expression over a continuous range [18, 158]. Second, the expression level can be varied discretely by introducing either predefined or randomized constitutive promoter and/or ribosome-binding site (RBS) libraries [3, 38, 97, 129, 158, 170]. Third, a double plasmid architecture, with the TF gene and TFBS-reporter gene located on two different plasmids, permits the independent variation of TF expression level and TFBS copy number by utilizing differing origins of replication generating differing plasmid copy numbers. The same approach can be implemented by integrating either the TF gene or the TFBS-reporter gene on (multiple) chromosomal locations [6, 18, 130]. The concept of introducing competing TFBSs to titrate available TFs away from the TFBS of the biosensor will be discussed in the next section, engineering at the TFBS level.

The first two methods were used in parallel to successfully alter the response curves of three biosensors, based on the repressors TetR and ArsR and the activator LuxR [158]. The arabinose-inducible AraC regulatory system and a set of constitutive promoters (Anderson collection, Registry of Standard Biological Parts [73]) were used to modulate the TF expression level over a continuous and discrete range, respectively. Substantiated by the parameters of the fitted Hill function, the tailored repressor-based biosensors demonstrated increased ligand sensitivity and dynamic range when the respective TFs were less abundantly expressed. As such, it was hypothesized that due to the lower abundance of these repressors, typically bound to the TFBS to prevent transcription initiation, a lower amount of ligand molecules is required to relieve inhibition. The opposite phenomenon was observed for the engineered activator-based biosensor which displayed increased ligand sensitivity and dynamic range when the TF was more abundantly expressed. Following the same rationale, less ligand molecules would be required to form activating ligand/TF complexes due to higher TF abundance [158]. In a different study, the dynamic range of a malonyl-CoA-inducible FapR-based biosensor was optimized by fine-tuning the FapR expression level using the same arabinose-inducible AraC regulatory system. Interestingly, a clear discrepancy was observed between TF expression levels which were too low or too high, resulting in either no significant repression at all or unabolishable repression at any ligand concentration, respectively. The optimized biosensor, with a dynamic range delineated by this observed maximal and minimal output value, was successfully used to dynamically control the expression of a malonyl-CoA producing enzyme to mitigate potential toxicity upon overexpression [84]. Similarly, a distinct optimal transcription initiation rate, originating from a constitutive promoter library, was observed for the expression of two activators, XylS and HbpR, with regard to the increase of the dynamic range of their biosensor circuits [170]. The corresponding developed biosensors may serve as monitoring systems for environmental pollution through aromatic compound detection [170]. In another study, combinations of three plasmid backbones, with differing copy number, and two constitutive promoters, with differing transcription initiation rate, were used to determine the optimal LacI expression level to ensure the efficiency of a directed evolution strategy to create novel LacI variants [130]. The TFBS, core promoter and reporter

gene were located separately from the TF gene on the host chromosome. A strong LacI promoter in combination with a low copy number plasmid was identified as optimal for screening LacI mutants demonstrating switch-like behavior at low IPTG concentrations. During screening, however, this relatively low expression level may have caused leaky expression which may have resulted in a decrease in mutational success rate. Consequently, the expression levels of several identified mutants were increased using a high-copy-number plasmid backbone resulting in a decrease in leaky expression of the reporter gene and an increase in induction ratio [130].

Noteworthy, the necessity of an inducible TF expression system for biosensor applications may be undesirable and can easily be circumvented by introducing a constitutive promoter with the desired transcription initiation rate. Moreover, if present, the naturally occurring autoregulation mechanism of a transcriptional regulatory system may be exploited to stabilize TF expression. Such a negative feedback system can ensure that optimal TF concentration limits are not surpassed and, consequently, the distribution of TF copies among cells is tightened [10].

Engineering at the transcription factor binding site/promoter level

Binding of the TF to the TFBS directly influences the interaction of RNAP with the core promoter (−35 and −10 hexameric consensus boxes) through different regulatory mechanisms depending on the TF/TFBS pair used for biosensor construction (see Sect. "Transcriptional biosensors: a beginner's guide"). In parallel to TF engineering, on the DNA level, the specific nucleotide sequence of both the TFBS and the core promoter region can be altered to modulate the TF/TFBS and RNAP/core promoter interactions and, accordingly, the response curve (see Fig. 4c) [4, 6, 19]. Furthermore, the position relative to the core promoter and other TFBS copies and the multiplicity of the TFBS, and potential competing binding sites, are essential factors in transcriptional regulation which can be engineered to tune the biosensor response curve (see Fig. 4d) [4, 19].

Creating custom transcription factor binding site/promoter variants

Core promoter sequences can be customized to bind RNAP at different kinetic constants to customize transcription initiation rates [3, 38, 97]. Furthermore, in a biosensor system, this RNAP binding is in itself directly influenced by TF binding on the TFBS. Similarly, customizing the TFBS sequence can change the affinity of the TF towards the TFBS and, therefore, influences the transcriptional regulation process of the biosensor. From the core promoter perspective, core promoter sequences which demonstrate low RNAP-binding rates are generally more easily repressed due to steric hindrance upon TF binding to the TFBS [89]. However, these weak promoters usually demonstrate a low maximum reporter expression. On the other hand, strong core promoters may be less efficiently repressed by TF binding [89]. From the TFBS perspective, a TFBS for which the TF demonstrates a high binding affinity is expected to require a lower amount of TF to achieve a certain level of repression in comparison with a TFBS for which the TF demonstrates only a low binding affinity [13, 14, 50, 155]. In general, targets for mutagenesis, to create novel TF and RNAP-binding affinities, range from complete intergenic regions, over specific TFBS or promoter consensus regions, to specific base-pairs to customize the response curve (see Fig. 4c).

For instance, the response curve of a FapR-based malonyl-CoA biosensor was tuned by introducing random base-pair substitutions in the −35 and −10 boxes of the core promoter and in the TFBS sequence [84]. The response curves of these biosensor variants were generally shifted vertically, demonstrating a concurrent change in both the maximum output signal and the amount of leaky expression [84]. In a similar study, the response curve of a PcaU-based biosensor was optimized for more sensitive detection of 3,4-dyhydroxybenzoate, a nylon precursor [69]. The core promoter region was randomized at small set of specific nucleotides which resulted in biosensor variants demonstrating the desired increase in dynamic range with, however, also an increase in leaky expression [69]. A different illustrative study aimed at the development of a set of orthogonal anhydrotetracycline (aTc)-responsive TetR biosensor pairs [141]. First, TF sequences were obtained from different prokaryotic genomes based both on TetR homology and reported regulatory functionality without prior knowledge on their corresponding native TFBSs [117]. Therefore, synthetic TFBS libraries were designed based on the characteristics of known TFBSs of TetR homologues, followed by in vitro consensus analysis and, subsequently, in vivo selection of the TFBSs for which the highest dynamic range was observed [160]. By identifying possible cross reactions of the newly obtained TF/TFBS pair library, 16 pairs were identified which can be used orthogonally for genetic circuit building [141]. Although the scope of this study was to screen for novel orthogonal TF/TFBS pairs, this workflow can also be applied directly to tune response curves by changing the TFBS sequence and position for both reported as putative TFs in a high-throughput manner.

A combination of TF engineering and TFBS engineering was used to construct novel TF-TFBS pairs with altered TFBS-binding specificity and affinity [40]. Based on comparative genomics analysis [121], three amino-acid

positions in the DBD of the cAMP receptor protein (CRP) were mutated to create eight CRP variants. The eight predicted cognate TFBSs were created by mutating the three nucleotides in each half-site of the corresponding palindromic TFBS which were predicted to interact with the DBD. Four out of eight TF-TFBS pairs were able to activate expression of the reporter gene with one of them only showing weak activation. In addition, by randomly mutating the six central nucleotides in the corresponding TFBS, the dynamic range was increased [40].

Finally, core promoter/TFBS engineering can be applied for the adaptation of native and non-native biosensor machinery to more orthogonal and more compatible biosensor variants, respectively, with respect to the host organism of choice. As an example, the introduction of the TF gene and complete intergenic region from an *Acinetobacter* sp. in *E. coli* resulted in an impracticable signal-to-noise ratio for biosensor application [69]. This problem was partially resolved by introducing *E. coli* consensus RBS, −35 and −10 box sequences. However, the introduction of the consensus core promoter boxes likely generated a more constitutive promoter resulting in an increase in leaky expression of the biosensor [69].

Tuning position and multiplicity of transcription factor binding sites…

The position of a TFBS relative to the core promoter has a direct influence on the transcriptional regulatory mechanism of a biosensor. As discussed previously (see Sect. "Transcriptional biosensors: a beginner's guide"), also the relative position on the DNA helix of the TFBS relative the core promoter sequence influences the response of some regulatory mechanisms [8, 62]. Consequently, the modulation of the relative TFBS position directly influences the mechanistic interplay of the TF, TFBS, RNAP and core promoter and, therefore, the response curve characteristics (see Fig. 4d). Interestingly, when a TF is located away from its natural locations by placing the TFBS at different positions, other regulatory mechanisms might come into play [50, 107, 108, 126]. This dynamic interplay becomes even more complex when the TF of interest interacts with additional TF copies to regulate transcription. This di- or oligomerization implies cooperative binding and the presence of multiple TFBSs which in itself provides new engineering opportunities. As such, varying the number of TFBSs and their relative position to one another can be used to tune the biosensor response curve (see Fig. 4d).

For instance, when increasing the sequence length of the spacer between the core promoter and the TFBS, local optima, alternated with local minima, of efficient transcriptional regulation are observed at spacer length intervals corresponding to complete helical turns (10.5 nt) [8, 16, 43, 62, 100]. This phenomenon was observed for both the dimeric AraC TF and the tetrameric LacI TF which cooperatively interact with multiple TFBSs to form a DNA loop [9, 16, 43, 79, 100]. The introduction of complete helical turns between the two TFBSs responsible for DNA looping preserves repressibility, while half-integral turns cause hindered repression. In the latter situation, the bound TFs are probably located at opposite sides of the DNA helix which prevents DNA looping and, therefore, efficient repression [9, 16, 43, 79, 100]. Interestingly, DNA looping of a regulatory circuit also enhances and stabilizes repression levels with respect to changing TF expression levels [155]. Simple repression through steric hindrance is also influenced by the (helical) position of the TFBS relative to the core promoter region [50, 76]. The position of the single TFBS of a LacI-based simple repression system was varied upstream with single base-pair resolution to investigate the effects on repression [50]. Similarly, the helical period of the DNA helix recurred in the interval between the local repression optima [50].

In another study, various response curves were generated for an aTc/TetR-based repressive system by combinatorially varying the number and position of the corresponding single TFBS [101]. Seven core promoter/TFBS combinations were constructed with one, two or three TFBS copies and differing TFBS spacer length. In addition, a computational model was developed to predict response curves and output signal noise depending on the position and multiplicity of identical TFBSs with decreasing predictability for an increasing number of TFBSs. For the single copy TFBS circuits, leaky expression increased, concurrent with a decrease in the slope of response curve, as the TFBS is located further away from the core promoter region. Furthermore, the constructed triple copy TFBS circuit demonstrated a lower amount of leaky expression in comparison with any other tested circuit. The generation of noise in the output signal seemed to increase as the TFBS was located more proximally to the core promoter region which was likely related to the corresponding change in leaky expression [101]. Although the study of this prokaryotic transcriptional regulation system was performed in the eukaryote *Saccharomyces cerevisiae*, the collected information should, to a certain extent, be transferable to prokaryotic systems because of the straight-forward nature of this steric hindrance mechanism for repression.

In parallel with the previously discussed modulation of TF expression levels, the availability of TFs for transcriptional regulation can be controlled by the introduction of multiple copies of alternative TFBSs located on, e.g., several plasmid copies. These competing TFBSs, or decoys, can bind the corresponding TF to modulate the TF/TFBS ratio of the biosensor circuit and, consequently, alter the response curve. In addition, this concept reflects the fact

that even in nature, for some TFs, the number of corresponding chromosomal TFBSs could be large enough for TF titration [128].

To illustrate, a high-copy number plasmid, containing a LacI binding site, was used to titrate away available LacI TFs from a repressed gene of interest [162]. More specifically, the expression of an antibiotic resistance gene, under control of LacI repression, could be effectively induced by transforming the strain with said high-copy number plasmid, instead of supplementing IPTG. This approach enables the selection of plasmids which only contain a LacI TFBS and makes the use of resistance genes obsolete [162]. Similarly, the effect of plasmid copy number on repression was observed for the aTc/TetR repression system while maintaining a constant TF expression level [89]. Different plasmids with varying copy number were compared which contained the core promoter, TFBS and reporter gene. Here, TetR repression was less efficient for high-copy number plasmids in comparison with low copy number plasmids probably due to the changed TF/TFBS ratio as well as a decoy effect of an increase in unspecific binding sites [89]. In a more recent study, the simple LacI repression system was studied, independent of IPTG induction, with the aim of examining the interplay between the number of TFs and the number of corresponding and competing TFBSs [18]. The LacI repressor was put under control of the aTc/TetR regulatory system to adjust its intracellular expression level and regulated, in turn, the transcription initiation rate of a fluorescent protein. The copy number of the TFBSs, relative to the expression level of the TF, was altered ranging from single to high-copy numbers, with the latter generated by multiple chromosomal integrations or high-copy number plasmids. In addition, competitor TFBSs were introduced on plasmids, which did not contain a reporter gene to express. Interestingly, when the TF was distributed among an abundance of multiple TFBSs, the effect of an increase in TF copies had no significant effect on the repression of the reporter protein. When the amount of available TFs exceeded the TFBS copy number, this effect switched to a more normal, single copy-like response. The abrupt, or switch-like, nature of this phenomenon is largely dependent on the affinity the TF demonstrates for the TFBS, in addition to the distribution width of the plasmid copy number among cells. Furthermore, a thermodynamic model was constructed which enables the prediction for any such TFBS/TF configurations [18]. In a parallel study, a mechanistic model was developed, describing the effects of TF titration on reporter expression levels for simple repression systems and repression systems with DNA looping [128]. Finally, the effect of chromosomal gene location of the TFBS, core promoter and reporter gene parts of the LacI repression system was investigated while conserving the location of the LacI gene [15]. The IPTG-response curves systematically shifted vertically, without a change in the response curve slope, to lower expression levels with increasing distance from the origin. This effect was proven to be independent of the relative distance between LacI and its TFBS and was mainly the result of differences in gene copy number in relation to the chromosomal gene location. This effect can be exploited to predictably tune the response curve for chromosomal located biosensor machinery [15].

Engineering the global biosensor architecture

Besides aforementioned engineering strategies targeting the TF or TFBS (see Sects. "Engineering biosensor ligand specificity" and "Engineering the biosensor response curve"), several other strategies might be used to tune a biosensor's response curve [4, 6, 19, 115]. Instead of focusing on the underlying transcriptional regulatory system, such strategies can target the global biosensor architecture to engineer the response curve. To this end, the ligand input levels and reporter output levels can be tuned, in addition to the construction of a cascaded biosensor by introducing additional regulatory systems in series [64, 104, 115, 157, 170]. These techniques can generally be applied independent of the chosen TF/TFBS pair and, consequently, are more easily transferable to other biosensor designs and can be used to engineering multiple distinctive biosensor designs in parallel [115].

Engineering the ligand input level

The operational range can be tuned by modifying the amount of ligand molecules which are effectively detected by the biosensor. If the intracellular concentration of the ligand molecule is decreased by the expression of an exporter protein, the operational range of the biosensor can be shifted to higher ligand concentrations [115]. Likely, the opposite can be achieved by the expression of an importer protein. For example, the operational range of the aTc/TetR-based regulatory circuit was shifted tenfold higher for both the lower and upper ligand concentration limit by expressing the aTc exporter protein, TetA. In addition, this operational shift could be customized by tuning the expression of this exporter with the arabinose/AraC-based regulatory system [115].

Engineering the reporter output level

The dynamic range is determined by the mRNA output, which is, for example, translated into a fluorescent protein, in proportion to the ligand input. This output level can be modulated by changing, e.g., the mRNA stability, the translation (initiation) rate and the protein degradation rate to

change the response curve without interference of the transcriptional regulatory circuit of the biosensor. These concepts are reviewed in more detail elsewhere [4, 6, 19, 113]. For example, the translation initiation rate of the fluorescent reporter protein was varied to customize the response curve of an HbpR-based biosensor for aromatic compound detection [170]. More specifically, a library of predefined RBS sequences with reported relative strengths was introduced and the resulting response curves analyzed. This approach enabled the increase of the dynamic range of the biosensor without directly altering the cognate TF/TFBS regulatory system [170]. In a similar effort, the RBS strength was varied to modulate selector expression levels and, consequently, the evolutionary escape rate of a TtgR-based naringenin biosensor [26, 115]. This biosensor was successfully applied for toggled selection during an iterative genome-wide mutagenesis approach to increase naringenin production [115]. Besides improving the translation initiation rate of an output protein, the degradation of said protein can be modulated by adding different proteolytic tags to the output protein [4, 6, 115]. This approach was validated by appending different degradation tags to the selector output protein of a naringenin biosensor to reduce the evolutionary escape rate during a genome-wide directed evolution approach for optimizing microbial naringenin production. As the degradation rates increased with differing tags, the evolutionary escape rate was reduced with, however, also a reduction in operational range [115].

Biosensor cascades

A biosensor response curve can also be tuned by creating a linear single input/single output biosensor cascade which contains one or more additional TF/TFBS pairs in series. The development of such a cascaded biosensor enables inversion of the input/output relation, tunable and fixed-gain amplification of the output signal and modulation of response times, among others [64, 104, 127, 145, 157].

As an example, the aTc/TetR-based regulatory circuit was expanded with up to two extra TF/TFBS pairs each regulating TF expression of the subsequent link in the biosensor cascade [64]. Here, the increase in cascade length directly altered the operational and dynamic range, with a clear increase in sensitivity and decrease of the half-maximal threshold as a result. However, a downside of this approach was observed with regard to cell–cell variability and response delay time which both increased as the cascade length increased [64]. In a different study, a feedback-based amplifier was successfully developed for a single input/single output biosensor circuit [104]. More specifically, a constitutively active, ligand-independent variant of the LuxR repressor was used as input of an amplifier circuit to regulate the transcription initiation rates of both a fluorescent reporter protein and the LuxR variant itself. The output signal of the aTc/TetR-based biosensor was amplified by putting an additional copy of the LuxR variant gene under control of the TetR repressor and co-transforming the amplifier circuit on a separate plasmid into the biosensor host. When the ligand, aTc, is supplemented, TetR repression is inhibited and transcription of the LuxR gene variant is initiated. Subsequently, this LuxR variant initiates transcription of the fluorescent protein gene and the additional LuxR variant gene expression which acts as the amplifier by initiating transcription even more. This amplifier circuit increased the maximum output signal, the operational range, and sensitivity of the biosensor response curve [104]. A different approach was used by Wang and coworkers to develop both a fixed-gain and a tunable-gain amplifier circuit which can be integrated in a biosensor circuit [157]. Here, a set of two complexing ultrasensitive, ligand-independent activator proteins were introduced as the direct output of an arsenic-responsive ArsR-based biosensor to regulate the expression of a fluorescent reporter gene in the next chain of the cascade. This fixed-gain amplifier circuit increased the dynamic range as well as the ligand sensitivity of the biosensor response curve. A tunable variant of this circuit was constructed by introducing an arabinose/AraC-based circuit which controlled the transcription initiation rate of a third protein which inhibits the formation of the activating complex of the other two proteins in the amplifier circuit. Noteworthy, no apparent increase in response delay time or noise was observed when implementing these amplifier circuits [157]. Finally, a set of 15 "sensitivity tuners" were developed based on phage-derived TF-TFBS pairs which can be put in series into a biosensor circuit of choice. These circuits were able to increase the maximal output signal and slope of the arabinose/AraC-based biosensor [4, 177].

Conclusions and perspectives

Metabolism heavily relies on transcriptional regulatory systems to optimally control cellular processes with a view to optimal performance under a wide range of environmental conditions. These regulatory systems are ideal starting points for the development of biosensors for industrial biotechnology applications to dynamically control de novo assembled pathways and to screen combinatorial microbial cell factory libraries, among others. In this respect, in view of the ever increasing set of biotechnologically produced molecules and despite the already comprehensive transcription factor libraries, the further discovery of novel ligands, transcription factors and binding sites through part mining from nature will spark future applications when exploiting the next-generation sequencing tools and extensive computational methods.

However, the operational and dynamic ranges and ligand specificity of such natural biosensors typically do not fully align with the strict requirements of specific applications, e.g., the operational range of natural biosensor is generally too low for screening libraries of excessive producers. Hence, extensive tuning of the biosensor's main characteristics is of utmost importance. To this end, various techniques ranging from full-blown mutagenesis to computer-aided design and ranging from fluorescence-activated cell sorting to kinetic characterization have successfully been applied whereby novel ligand specificities and/or the dynamic and operational range have been drastically altered.

Nevertheless, despite numerous success stories in reengineering the transcriptional biosensors' main characteristics, true incremental knowledge acquisition is limited due the incredibly large number of parameters defining both the specificity and the dose response curve. Consequently, many engineering efforts have thus far relied on mostly ad hoc engineering strategies with limited predictability and scarce opportunities for true standardization. To streamline the development process of novel biosensors, the advancement of a holistic computational framework that considers the various interactions of different nature will be of utmost importance. Undoubtedly, this will accelerate the biosensor development process, which is of great interest for metabolic engineering and will enable industrial biotechnology to generate more efficient and faster microbial cell factories for an expanding collection of natural and non-natural molecules.

Acknowledgements The first two authors hold a Ph.D. Grant from the Institute for the Promotion of Innovation through Science and Technology in Flanders (IWT-Vlaanderen). This research was also supported by the project "CondEx" (FWO-G.0321.13 N) and "Nano-3Bio" from the EU 7th Framework Programme (FP7/2007-2013) under Grant Agreement No. 613931.

References

1. Aharoni A, Gaidukov L, Khersonsky O, Gould SM, Roodveldt C, Tawfik DS (2004) The "evolvability" of promiscuous protein functions. Nat Genet 37:73. doi:10.1038/ng1482
2. Allen BD, Mayo SL (2009) An efficient algorithm for multistate protein design based on faster. J Comput Chem 31:904–916. doi:10.1002/jcc.21375
3. Alper H, Fischer C, Nevoigt E, Stephanopoulos G (2005) Tuning genetic control through promoter engineering. Proc Natl Acad Sci 102:12678–12683. doi:10.1073/pnas.0504604102
4. Ang J, Harris E, Hussey BJ, Kil R, McMillen DR (2013) Tuning response curves for synthetic biology. ACS Synth Biol 2:547–567. doi:10.1021/sb4000564
5. Aravind L, Anantharaman V, Balaji S, Babu MM, Iyer LM (2005) The many faces of the helix-turn-helix domain: transcription regulation and beyond. FEMS Microbiol Rev 29:231–262. doi:10.1016/j.femsre.2004.12.008
6. Arpino JAJ, Hancock EJ, Anderson J, Barahona M, Stan G-BV, Papachristodoulou A, Polizzi K (2013) Tuning the dials of synthetic biology. Microbiology 159:1236–1253. doi:10.1099/mic.0.067975-0
7. Balleza E, López-Bojorquez LN, Martínez-Antonio A, Resendis-Antonio O, Lozada-Chávez I, Balderas-Martínez YI, Encarnación S, Collado-Vides J (2009) Regulation by transcription factors in bacteria: beyond description. FEMS Microbiol Rev 33:133–151. doi:10.1111/j.1574-6976.2008.00145.x
8. Barnard A, Wolfe A, Busby S (2004) Regulation at complex bacterial promoters: how bacteria use different promoter organizations to produce different regulatory outcomes. Curr Opin Microbiol 7:102–108. doi:10.1016/j.mib.2004.02.011
9. Becker NA, Peters JP, Lionberger TA, Maher LJ (2013) Mechanism of promoter repression by Lac repressor-DNA loops. Nucleic Acids Res 41:156–166. doi:10.1093/nar/gks1011
10. Becskei A, Serrano L (2000) Engineering stability in gene networks by autoregulation. Nature 405:590–593. doi:10.1038/35014651
11. Bikard D, Jiang W, Samai P, Hochschild A, Zhang F, Marraffini LA (2013) Programmable repression and activation of bacterial gene expression using an engineered CRISPR-Cas system. Nucleic Acids Res 41:7429–7437. doi:10.1093/nar/gkt520
12. Binder S, Schendzielorz G, Stäbler N, Krumbach K, Hoffmann K, Bott M, Eggeling L (2012) A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. Genome Biol 13:R40. doi:10.1186/gb-2012-13-5-r40
13. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Kuhlman T, Phillips R (2005) Transcriptional regulation by the numbers: applications. Curr Opin Genet Dev 15:125–135. doi:10.1016/j.gde.2005.02.006
14. Bintu L, Buchler NE, Garcia HG, Gerland U, Hwa T, Kondev J, Phillips R (2005) Transcriptional regulation by the numbers: models. Curr Opin Genet Dev 15:116–124. doi:10.1016/j.gde.2005.02.007
15. Block DHS, Hussein R, Liang LW, Lim HN (2012) Regulatory consequences of gene translocation in bacteria. Nucleic Acids Res 40:8979–8992. doi:10.1093/nar/gks694
16. Bond LM, Peters JP, Becker NA, Kahn JD, Maher LJ (2010) Gene repression by minimal lac loops in vivo. Nucleic Acids Res 38:8072–8082. doi:10.1093/nar/gkq755
17. Brennan RG, Matthews BW (1989) The helix-turn-helix DNA binding motif. J Biol Chem 264:1903–1906
18. Brewster RC, Weinert FM, Garcia HG, Song D, Rydenfelt M, Phillips R (2014) The transcription factor titration effect dictates level of gene expression. Cell 156:1312–1323. doi:10.1016/j.cell.2014.02.022
19. Brophy JAN, Voigt CA (2014) Principles of genetic circuit design. Nat Methods 11:508–520. doi:10.1038/nmeth.2926
20. Browning DF, Busby SJW (2004) The regulation of bacterial transcription initiation. Nat Rev Microbiol 2:57–65. doi:10.1038/nrmicro787
21. Browning DF, Busby SJW (2016) Local and global regulation of transcription initiation in bacteria. Nat Rev Microbiol 14:638–650. doi:10.1038/nrmicro.2016.103
22. Carbajosa G, Trigo A, Valencia A, Cases I (2009) Bionemo: molecular information on biodegradation metabolism. Nucleic Acids Res 37:D598–D602. doi:10.1093/nar/gkn864
23. De Carlo S, Chen B, Hoover TR, Kondrashkina E, Nogales E, Nixon BT (2006) The structural basis for regulated assembly and function of the transcriptional activator NtrC. Genes Dev 20:1485–1495. doi:10.1101/gad.1418306
24. Cebolla A, Sousa C, De Lorenzo V (1997) Effector specificity mutants of the transcriptional activator NahR of naphthalene degrading *Pseudomonas* define protein sites involved in

binding of aromatic inducers. J Biol Chem 272:3986–3992. doi:10.1074/jbc.272.7.3986
25. Chai Y, Winans SC (2003) Site-directed mutagenesis of a LuxR-type quorum-sensing transcription factor: alteration of autoinducer specificity. Mol Microbiol 51:765–776. doi:10.1046/j.1365-2958.2003.03857.x
26. Chen H, Bjerknes M, Kumar R, Jay E (1994) Determination of the optimal aligned spacing between the Shine-Dalgarno sequence and the translation initiation codon of *Escherichia coli* mRNAs. Nucleic Acids Res 22:4953–4957
27. Chen W, Zhang S, Jiang P, Yao J, He Y, Chen L, Gui X, Dong Z, Tang S-Y (2015) Design of an ectoine-responsive AraC mutant and its application in metabolic engineering of ectoine biosynthesis. Metab Eng 30:149–155. doi:10.1016/j.ymben.2015.05.004
28. Chong HH, Ching CB (2016) Development of colorimetric-based whole-cell biosensor for organophosphorus compounds by engineering transcription regulator DmpR. ACS Synth Biol. doi:10.1021/acssynbio.6b00061
29. Chou HH, Keasling JD (2013) Programming adaptive control to evolve increased metabolite production. Nat Commun 4:2595. doi:10.1038/ncomms3595
30. Cipriano MJ, Novichkov PN, Kazakov AE, Rodionov DA, Arkin AP, Gelfand MS, Dubchak I (2013) RegTransBase–a database of regulatory sequences and interactions based on literature: a resource for investigating transcriptional regulation in prokaryotes. BMC Genom 14:213
31. Collins CH, Arnold FH, Leadbetter JR (2005) Directed evolution of *Vibrio fischeri* LuxR for increased sensitivity to a broad spectrum of acyl-homoserine lactones. Mol Microbiol 55:712–723. doi:10.1111/j.1365-2958.2004.04437.x
32. Collins CH, Leadbetter JR, Arnold FH (2006) Dual selection enhances the signaling specificity of a variant of the quorum-sensing transcriptional activator LuxR. Nat Biotechnol 24:708–712. doi:10.1111/j.1467-9655.2006.00359_31.x
33. Combs SA, DeLuca SL, DeLuca SH, Lemmon GH, Nannemann DP, Nguyen ED, Willis JR, Sheehan JH, Meiler J (2013) Small-molecule ligand docking into comparative models with Rosetta. Nat Protoc 8:1277–1298. doi:10.1038/nprot.2013.074
34. Cress BF, Jones JA, Kim DC, Leitz QD, Englaender JA, Collins SM, Linhardt RJ, Koffas MAG (2016) Rapid generation of CRISPR/dCas9-regulated, orthogonally repressible hybrid T7-lac promoters for modular, tuneable control of metabolic pathway fluxes in *Escherichia coli*. Nucleic Acids Res 44:4472–4485. doi:10.1093/nar/gkw231
35. Cuthbertson L, Nodwell JR (2013) The TetR family of regulators. Microbiol Mol Biol Rev 77:440–475. doi:10.1128/MMBR.00018-13
36. Dahl RH, Zhang F, Alonso-Gutierrez J, Baidoo E, Batth TS, Redding-Johanson AM, Petzold CJ, Mukhopadhyay A, Lee TS, Adams PD, Keasling JD (2013) Engineering dynamic pathway regulation using stress-response promoters. Nat Biotechnol 31:1039–1046. doi:10.1038/nbt.2689
37. Das R, Baker D (2008) Macromolecular modeling with rosetta. Annu Rev Biochem 77:363–382. doi:10.1146/annurev.biochem.77.062906.171838
38. Davis JH, Rubin AJ, Sauer RT (2011) Design, construction and characterization of a set of insulated bacterial promoters. Nucleic Acids Res 39:1131–1141. doi:10.1093/nar/gkq810
39. Delépine B, Libis V, Carbonell P, Faulon JL (2016) SensiPath: computer-aided design of sensing-enabling metabolic pathways. Nucleic Acids Res 44:W226–231. doi:10.1093/nar/gkw305
40. Desai TA, Rodionov DA, Gelfand MS, Alm EJ, Rao CV (2009) Engineering transcription factors with novel DNA-binding specificity using comparative genomics. Nucleic Acids Res 37:2493–2503. doi:10.1093/nar/gkp079
41. Dietrich JA, McKee AE, Keasling JD (2010) High-throughput metabolic engineering: advances in small-molecule screening and selection. Annu Rev Biochem 79:563–590. doi:10.1146/annurev-biochem-062608-095938
42. Djordjevic M, Sengupta AM, Shraiman BI (2003) A biophysical approach to transcription factor binding site discovery. Genome Res 13:2381–2390. doi:10.1101/gr.1271603
43. Dunn TM, Hahn S, Ogden S, Schleif RF (1984) An operator at −280 base pairs that is required for repression of *araBAD* operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. Proc Natl Acad Sci USA 81:5017–5020
44. Eggeling L, Bott M, Marienhagen J (2015) Novel screening methods—biosensors. Curr Opin Biotechnol 35:30–36. doi:10.1016/j.copbio.2014.12.021
45. Fernandez-López R, Ruiz R, de la Cruz F, Moncalián G (2015) Transcription factor-based biosensors enlightened by the analyte. Front Microbiol 6:648. doi:10.3389/fmicb.2015.00648
46. Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M (2014) Pfam: the protein families database. Nucleic Acids Res 42:D222–D230. doi:10.1093/nar/gkt1223
47. Galvão TC, de Lorenzo V (2006) Transcriptional regulators à la carte: engineering new effector specificities in bacterial regulatory proteins. Curr Opin Biotechnol 17:34–42. doi:10.1016/j.copbio.2005.12.002
48. Galvão TC, Mencía M, De Lorenzo V (2007) Emergence of novel functions in transcriptional regulators by regression to stem protein types. Mol Microbiol 65:907–919. doi:10.1111/j.1365-2958.2007.05832.x
49. Gama-Castro S, Salgado H, Peralta-Gil M, Santos-Zavaleta A, Muniz-Rascado L, Solano-Lira H, Jimenez-Jacinto V, Weiss V, Garcia-Sotelo JS, Lopez-Fuentes A, Porron-Sotelo L, Alquicira-Hernandez S, Medina-Rivera A, Martinez-Flores I, Alquicira-Hernandez K, Martinez-Adame R, Bonavides-Martinez C, Miranda-Rios J, Huerta AM, Mendoza-Vargas A, Collado-Torres L, Taboada B, Vega-Alvarado L, Olvera M, Olvera L, Grande R, Morett E, Collado-Vides J (2011) RegulonDB version 7.0: transcriptional regulation of *Escherichia coli* K-12 integrated within genetic sensory response units (gensor units). Nucleic Acids Res 39:D98–D105. doi:10.1093/nar/gkq1110
50. Garcia HG, Sanchez A, Boedicker JQ, Osborne M, Gelles J, Kondev J, Phillips R (2012) Operator sequence alters gene expression independently of transcription factor occupancy in bacteria. Cell Rep 2:150–161. doi:10.1016/j.celrep.2012.06.004
51. Garmendia J, Devos D, Valencia A, De Lorenzo V (2001) À la carte transcriptional regulators: unlocking responses of the prokaryotic enhancer-binding protein XylR to non-natural effectors. Mol Microbiol 42:47–59. doi:10.1046/j.1365-2958.2001.02633.x
52. Glascock CB, Weickert MJ (1998) Using chromosomal lacIQ1 to control expression of genes on high-copy-number plasmids in *Escherichia coli*. Gene 223:221–231
53. Grote A, Klein J, Retter I, Haddad I, Behling S, Bunk B, Biegler I, Yarmolinetz S, Jahn D, Münch R (2009) PRODORIC (release 2009): a database and tool platform for the analysis of gene regulation in prokaryotes. Nucleic Acids Res 37:D61–D65
54. Gupta S, Saxena M, Saini N, Mahmooduzzafar Kumar R, Kumar A (2012) An effective strategy for a whole-cell biosensor based on putative effector interaction site of the regulatory DmpR protein. PLoS One 7:e43527. doi:10.1371/journal.pone.0043527
55. Guzman LM, Belin D, Carson MJ, Beckwith J (1995) Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter. J Bacteriol 177:4121–4130

56. Hawkins AC, Arnold FH, Stuermer R, Hauer B, Leadbetter JR (2007) Directed evolution of *Vibrio fischeri* LuxR for improved response to butanoyl-homoserine lactone. Appl Environ Microbiol 73:5775–5781. doi:10.1128/AEM.00060-07
57. Helbl V, Hillen W (1998) Stepwise selection of TetR variants recognizing tet operator 4C with high affinity and specificity. J Mol Biol 276:313–318. doi:10.1006/jmbi.1997.1540
58. Helbl V, Hillen W, Tiebel B (1998) Stepwise selection of TetR variants recognizing tet operator 6C with high affinity and specificity. J Mol Biol 276:319–324. doi:10.1006/jmbi.1997.1539
59. Hendlich M, Rippmann F, Barnickel G (1997) Ligsite: automatic and efficient detection of potential small molecule-binding sites in proteins. J Mol Graph Model 15:359–363. doi:10.1016/S1093-3263(98)00002-3
60. Henssler E-M, Bertram R, Wisshak S, Hillen W (2005) Tet repressor mutants with altered effector binding and allostery. FEBS J 272:4487–4496. doi:10.1111/j.1742-4658.2005.04868.x
61. Henssler E-M, Scholz O, Lochner S, Gmeiner P, Hillen W (2004) Structure-based design of Tet repressor to optimize a new inducer specificity. Biochemistry 43:9512–9518. doi:10.1021/bi049682j
62. van Hijum SAFT, Medema MH, Kuipers OP (2009) Mechanisms and evolution of control logic in prokaryotic transcriptional regulation. Microbiol Mol Biol Rev 73:481–509. doi:10.1128/MMBR.00037-08
63. Hill AV (1913) The combinations of haemoglobin with oxygen and with carbon monoxide. Biochem J 7:471
64. Hooshangi S, Thiberge S, Weiss R (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. Proc Natl Acad Sci 102:3581–3586. doi:10.1073/pnas.0408507102
65. Huang B, Schroeder M (2006) Ligsitecsc: predicting ligand binding sites using the connolly surface and degree of conservation. BMC Struct Biol 6:19. doi:10.1186/1472-6807-6-19
66. Huynh L, Tagkopoulos I (2016) A parts database with consensus parameter estimation for synthetic circuit design. ACS Synth Biol. doi:10.1021/acssynbio.5b00205
67. Jensen RA (1976) Enzyme recruitment in evolution of new function. Annu Rev Microbiol 30:409–425. doi:10.1146/annurev.mi.30.100176.002205
68. Jha RK, Chakraborti S, Kern TL, Fox DT, Strauss CEM (2015) Rosetta comparative modeling for library design: engineering alternative inducer specificity in a transcription factor. Proteins Struct Funct Bioinforma 83:1327–1340. doi:10.1002/prot.24828
69. Jha RK, Kern TL, Fox DT, Strauss CEM (2014) Engineering an *Acinetobacter* regulon for biosensing and high-throughput enzyme screening in *E. coli* via flow cytometry. Nucleic Acids Res 42:8150–8160. doi:10.1093/nar/gku444
70. Jha RK, Kern TL, Kim Y, Tesar C, Jedrzejczak R, Joachimiak A, Strauss CEM (2016) A microbial sensor for organophosphate hydrolysis exploiting an engineered specificity switch in a transcription factor. Nucleic Acids Res 44:8490–8500. doi:10.1093/nar/gkw687
71. Kamionka A, Bogdanska-Urbaniak J, Scholz O, Hillen W (2004) Two mutations in the tetracycline repressor change the inducer anhydrotetracycline to a corepressor. Nucleic Acids Res 32:842–847. doi:10.1093/nar/gkh200
72. Kawakami E, Nakaoka S, Ohta T, Kitano H (2016) Weighted enrichment method for prediction of transcription regulators from transcriptome and global chromatin immunoprecipitation data. Nucleic Acids Res 44:5010–5021. doi:10.1093/nar/gkw355
73. Kelly JR, Rubin AJ, Davis JH, Ajo-Franklin CM, Cumbers J, Czar MJ, de Mora K, Glieberman AL, Monie DD, Endy D (2009) Measuring the activity of BioBrick promoters using an in vivo reference standard. J Biol Eng 3:4. doi:10.1186/1754-1611-3-4
74. Keseler IM, Mackie A, Peralta-Gil M, Santos-Zavaleta A, Gama-Castro S, Bonavides-Martinez C, Fulcher C, Huerta AM, Kothari A, Krummenacker M, Latendresse M, Muniz-Rascado L, Ong Q, Paley S, Schroder I, Shearer AG, Subhraveti P, Travers M, Weerasinghe D, Weiss V, Collado-Vides J, Gunsalus RP, Paulsen I, Karp PD (2013) EcoCyc: fusing model organism databases with systems biology. Nucleic Acids Res 41:D605–D612. doi:10.1093/nar/gks1027
75. Klotzsche M, Ehrt S, Schnappinger D (2009) Improved tetracycline repressors for gene silencing in mycobacteria. Nucleic Acids Res 37:1778–1788. doi:10.1093/nar/gkp015
76. Lanzer M, Bujard H (1988) Promoters largely determine the efficiency of repressor action. Proc Natl Acad Sci USA 85:8973–8977
77. Lassila JK, Privett HK, Allen BD, Mayo SL (2006) Combinatorial methods for small-molecule placement in computational enzyme design. Proc Natl Acad Sci USA 103:16710–16715. doi:10.1073/pnas.0607691103
78. Le TBK, Stevenson CEM, Fiedler HP, Maxwell A, Lawson DM, Buttner MJ (2011) Structures of the TetR-like simocyclinone efflux pump repressor, SimR, and the mechanism of ligand-mediated derepression. J Mol Biol 408:40–56. doi:10.1016/j.jmb.2011.02.035
79. Lee DH, Schleif RF (1989) In vivo DNA loops in *araCBAD*: size limits and helical repeat. Proc Natl Acad Sci USA 86:476–480
80. Libis V, Delépine B, Faulon JL (2016) Sensing new chemicals with bacterial transcription factors. Curr Opin Microbiol 33:105–112. doi:10.1016/j.mib.2016.07.006
81. Libis V, Delépine B, Faulon J-L (2016) Expanding biosensing abilities through computer-aided design of metabolic pathways. ACS Synth Biol. doi:10.1021/acssynbio.5b00225
82. Lihu A, Holban Ş (2015) A review of ensemble methods for de novo motif discovery in ChIP-Seq data. Brief Bioinform 16:964–973. doi:10.1093/bib/bbv022
83. Liu D, Evans T, Zhang F (2015) Applications and advances of metabolite biosensors for metabolic engineering. Metab Eng 31:35–43. doi:10.1016/j.ymben.2015.06.008
84. Liu D, Xiao Y, Evans BS, Zhang F (2015) Negative feedback regulation of fatty acid production based on a malonyl-CoA sensor–actuator. ACS Synth Biol 4:132–140. doi:10.1021/sb400158w
85. Lönneborg R, Varga E, Brzezinski P (2012) Directed evolution of the transcriptional regulator DntR: isolation of mutants with improved DNT-response. PLoS One 7:e29994. doi:10.1371/journal.pone.0029994
86. Looger LL, Dwyer MA, Smith JJ, Hellinga HW (2003) Computational design of receptor and sensor proteins with novel functions. Nature 423:185–190. doi:10.1038/nature01556
87. de Lorenzo V, Pérez-Martín J (1996) Regulatory noise in prokaryotic promoters: how bacteria learn to respond to novel environmental signals. Mol Microbiol 19:1177–1184
88. de los Santos ELC, Meyerowitz JT, Mayo SL, Murray RM (2016) Engineering transcriptional regulator effector specificity using computational design and in vitro rapid prototyping: developing a vanillin sensor. ACS Synth Biol 5:287–295. doi:10.1021/acssynbio.5b00090
89. Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in *Escherichia coli* via the LacR/O, the TetR/O and AraC/I1-I2 regulatory elements. Nucleic Acids Res 25:1203–1210. doi:10.1093/nar/25.6.1203
90. Machado LFM, Dixon N (2016) Development and substrate specificity screening of an in vivo biosensor for the detection of biomass derived aromatic chemical building blocks. Chem Commun 52:11402–11405. doi:10.1039/C6CC04559F

91. Mahr R, von Boeselager RF, Wiechert J, Frunzke J (2016) Screening of an *Escherichia coli* promoter library for a phenylalanine biosensor. Appl Microbiol Biotechnol 100:6739–6753. doi:10.1007/s00253-016-7575-8
92. Mahr R, Frunzke J (2016) Transcription factor-based biosensors in biotechnology: current state and future prospects. Appl Microbiol Biotechnol 100:79–90. doi:10.1007/s00253-015-7090-3
93. Mahr R, Gätgens C, Gätgens J, Polen T, Kalinowski J, Frunzke J (2015) Biosensor-driven adaptive laboratory evolution of L-valine production in *Corynebacterium glutamicum*. Metab Eng 32:184–194. doi:10.1016/j.ymben.2015.09.017
94. Mathelier A, Wasserman WW (2013) The next generation of transcription factor binding site prediction. PLoS Comput Biol 9:e1003214. doi:10.1371/journal.pcbi.1003214
95. Meinhardt S, Manley MW, Becker NA, Hessman JA, Maher LJ, Swint-Kruse L (2012) Novel insights from hybrid LacI/GalR proteins: family-wide functional attributes and biologically significant variation in transcription repression. Nucleic Acids Res 40:11139–11154. doi:10.1093/nar/gks806
96. Meinhardt S, Swint-Kruse L (2008) Experimental identification of specificity determinants in the domain linker of a LacI/GalR protein: bioinformatics-based predictions generate true positives and false negatives. Proteins Struct Funct Bioinforma 73:941–957. doi:10.1002/prot.22121
97. De Mey M, Maertens J, Lequeux GJ, Soetaert WK, Vandamme EJ (2007) Construction and model-based analysis of a promoter library for *E. coli*: an indispensable tool for metabolic engineering. BMC Biotechnol 7:34. doi:10.1186/1472-6750-7-34
98. Mirny LA, Gelfand MS (2002) Using orthologous and paralogous proteins to identify specificity-determining residues in bacterial transcription factors. J Mol Biol 321:7–20. doi:10.1016/S0022-2836(02)00587-9
99. Mohn WW, Garmendia J, Galvao TC, de Lorenzo V (2006) Surveying biotransformations with à la carte genetic traps: translating dehydrochlorination of lindane (gamma-hexachlorocyclohexane) into lacZ-based phenotypes. Environ Microbiol 8:546–555. doi:10.1111/j.1462-2920.2006.00983.x
100. Müller J, Oehler S, Müller-Hill B (1996) Repression of lac promoter as a function of distance, phase and quality of an auxiliary lac operator. J Mol Biol 257:21–29. doi:10.1006/jmbi.1996.0143
101. Murphy KF, Balazsi G, Collins JJ (2007) Combinatorial promoter design for engineering noisy gene expression. Proc Natl Acad Sci 104:12726–12731. doi:10.1073/pnas.0608451104
102. Nielsen AAK, Der BS, Shin J, Vaidyanathan P, Paralanov V, Strychalski EA, Ross D, Densmore D, Voigt CA (2016) Genetic circuit design automation. Science 352:aac7341. doi:10.1126/science.aac7341
103. Nielsen J, Keasling JD (2016) Engineering cellular metabolism. Cell 164:1185–1197. doi:10.1016/j.cell.2016.02.004
104. Nistala GJ, Wu K, Rao CV, Bhalerao KD (2010) A modular positive feedback-based gene amplifier. J Biol Eng 4:4. doi:10.1186/1754-1611-4-4
105. Novichkov PS, Kazakov AE, Ravcheev DA, Leyn SA, Kovaleva GY, Sutormin RA, Kazanov MD, Riehl W, Arkin AP, Dubchak I, Rodionov DA (2013) RegPrecise 3.0–a resource for genome-scale exploration of transcriptional regulation in bacteria. BMC Genom 14:745. doi:10.1186/1471-2164-14-745
106. Pabo CO, Sauer RT (1992) Transcription factors: structural families and principles of DNA recognition. Annu Rev Biochem 61:1053–1095. doi:10.1146/annurev.biochem.61.1.1053
107. Pavco PA, Steege DA (1990) Elongation by *Escherichia coli* RNA polymerase is blocked in vitro by a site-specific DNA binding protein. J Biol Chem 265:9960–9969
108. Pavco PA, Steege DA (1991) Characterization of elongating T7 and SP6 RNA polymerases and their response to a roadblock generated by a site-specific DNA binding protein. Nucleic Acids Res 19:4639–4646
109. Perez-Rueda E, Collado-Vides J (2000) The repertoire of DNA-binding transcriptional regulators in *Escherichia coli* K-12. Nucleic Acids Res 28:1838–1847. doi:10.1093/nar/28.8.1838
110. Pérez-Rueda E, Collado-Vides J, Segovia L (2004) Phylogenetic distribution of DNA-binding transcription factors in bacteria and archaea. Comput Biol Chem 28:341–350. doi:10.1016/j.compbiolchem.2004.09.004
111. Pérez-Rueda E, Gralla JD, Collado-Vides J (1998) Genomic position analyses and the transcription machinery. J Mol Biol 275:165–170. doi:10.1006/jmbi.1997.1465
112. Pérez-Rueda E, Tenorio-Salgado S, Huerta-Saquero A, Balderas-Martínez YI, Moreno-Hagelsieb G (2015) The functional landscape bound to the transcription factors of *Escherichia coli* K-12. Comput Biol Chem 58:93–103. doi:10.1016/j.compbiolchem.2015.06.002
113. Peters G, Coussement P, Maertens J, Lammertyn J, De Mey M (2015) Putting RNA to work: translating RNA fundamentals into biotechnological engineering practice. Biotechnol Adv 33:1829–1844. doi:10.1016/j.biotechadv.2015.10.011
114. Petzold CJ, Chan LJG, Nhan M, Adams PD (2015) Analytics for metabolic engineering. Front Bioeng Biotechnol 3:135. doi:10.3389/fbioe.2015.00135
115. Raman S, Rogers JK, Taylor ND, Church GM (2014) Evolution-guided optimization of biosynthetic pathways. Proc Natl Acad Sci 111:17803–17808. doi:10.1073/pnas.1409523111
116. Raman S, Taylor N, Genuth N, Fields S, Church GM (2014) Engineering allostery. Trends Genet 30:521–528. doi:10.1016/j.tig.2014.09.004
117. Ramos JL, Martínez-Bueno M, Molina-Henares AJ, Terán W, Watanabe K, Zhang X, Gallegos MT, Brennan R, Tobes R (2005) The TetR family of transcriptional repressors. Microbiol Mol Biol Rev 69:326–356. doi:10.1128/MMBR.69.2.326-356.2005
118. Ray S, Gunzburg MJ, Wilce M, Panjikar S, Anand R (2016) Structural basis of selective aromatic pollutant sensing by the effector binding domain of MopR, an NtrC family transcriptional regulator. ACS Chem Biol 11:2357–2365. doi:10.1021/acschembio.6b00020
119. Reed B, Blazeck J, Alper H (2012) Evolution of an alkane-inducible biosensor for increased responsiveness to short-chain alkanes. J Biotechnol 158:75–79. doi:10.1016/j.jbiotec.2012.01.028
120. Rodionov DA (2007) Comparative genomic reconstruction of transcriptional regulatory networks in bacteria. Chem Rev 107:3467–3497. doi:10.1021/cr068309+
121. Rodionov DA, Dubchak IL, Arkin AP, Alm EJ, Gelfand MS (2005) Dissimilatory metabolism of nitrogen oxides in bacteria: comparative reconstruction of transcriptional networks. PLoS Comput Biol 1:e55. doi:10.1371/journal.pcbi.0010055
122. Rogers JK, Church GM (2016) Multiplexed engineering in biology. Trends Biotechnol 34:198–206. doi:10.1016/j.tibtech.2015.12.004
123. Rogers JK, Church GM (2016) Genetically encoded sensors enable real-time observation of metabolite production. Proc Natl Acad Sci USA 113:2388–2393. doi:10.1073/pnas.1600375113
124. Rogers JK, Guzman CD, Taylor ND, Raman S, Anderson K, Church GM (2015) Synthetic biosensors for precise gene control and real-time monitoring of metabolites. Nucleic Acids Res 43:7648–7660

125. Rogers JK, Taylor ND, Church GM (2016) Biosensor-based engineering of biosynthetic pathways. Curr Opin Biotechnol 42:84–91. doi:10.1016/j.copbio.2016.03.005
126. Rojo F (1999) Repression of transcription initiation in bacteria. J Bacteriol 181:2987–2991
127. Rosenfeld N, Alon U (2003) Response delays and the structure of transcription networks. J Mol Biol 329:645–654. doi:10.1016/S0022-2836(03)00506-0
128. Rydenfelt M, Cox RS, Garcia H, Phillips R (2014) Statistical mechanical model of coupled transcription from multiple promoters due to transcription factor titration. Phys Rev E 89:12702. doi:10.1103/PhysRevE.89.012702
129. Salis HM, Mirsky EA, Voigt CA (2009) Automated design of synthetic ribosome binding sites to control protein expression. Nat Biotechnol 27:946–950. doi:10.1038/nbt.1568
130. Satya Lakshmi O, Rao NM (2008) Evolving lac repressor for enhanced inducibility. Protein Eng Des Sel 22:53–58. doi:10.1093/protein/gzn069
131. Schallmey M, Frunzke J, Eggeling L, Marienhagen J (2014) Looking for the pick of the bunch: high-throughput screening of producing microorganisms with biosensors. Curr Opin Biotechnol 26C:148–154. doi:10.1016/j.copbio.2014.01.005
132. Schendzielorz G, Dippong M, Grünberger A, Kohlheyer D, Yoshida A, Binder S, Nishiyama C, Nishiyama M, Bott M, Eggeling L (2014) Taking control over control: use of product sensing in single cells to remove flux control at key enzymes in biosynthesis pathways. ACS Synth Biol 3:21–29. doi:10.1021/sb400059y
133. Scholz O, Henßler E-M, Bail J, Schubert P, Bogdanska-Urbaniak J, Sopp S, Reich M, Wisshak S, Köstner M, Bertram R, Hillen W (2004) Activity reversal of Tet repressor caused by single amino acid exchanges. Mol Microbiol 53:777–789. doi:10.1111/j.1365-2958.2004.04159.x
134. Scholz O, Köstner M, Reich M, Gastiger S, Hillen W (2003) Teaching TetR to recognize a new inducer. J Mol Biol 329:217–227. doi:10.1016/S0022-2836(03)00427-3
135. Schreiter ER, Drennan CL (2007) Ribbon–helix–helix transcription factors: variations on a theme. Nat Rev Microbiol 5:710–720. doi:10.1038/nrmicro1717
136. Shin HJ (2010) Development of highly-sensitive microbial biosensors by mutation of the nahR regulatory gene. J Biotechnol 150:246–250. doi:10.1016/j.jbiotec.2010.09.936
137. Shis DL, Hussain F, Meinhardt S, Swint-Kruse L, Bennett MR (2014) Modular, multi-input transcriptional logic gating with orthogonal LacI/GalR family chimeras. ACS Synth Biol 3:645–651. doi:10.1021/sb500262f
138. Siedler S, Schendzielorz G, Binder S, Eggeling L, Bringer S, Bott M (2014) SoxR as a single-cell biosensor for NADPH-consuming enzymes in *Escherichia coli*. ACS Synth Biol 3:41–47. doi:10.1021/sb400110j
139. Silva-Rocha R, de Lorenzo V (2014) Engineering multicellular logic in bacteria with metabolic wires. ACS Synth Biol 3:204–209. doi:10.1021/sb400064y
140. Smith AD, Sumazin P, Das D, Zhang MQ (2005) Mining ChIP-chip data for transcription factor and cofactor binding sites. Bioinformatics 21:i403–i412. doi:10.1093/bioinformatics/bti1043
141. Stanton BC, Nielsen AAK, Tamsir A, Clancy K, Peterson T, Voigt CA (2014) Genomic mining of prokaryotic repressors for orthogonal logic gates. Nat Chem Biol 10:99–105. doi:10.1038/nchembio.1411
142. Stefan MI, Le Novère N (2013) Cooperative binding. PLoS Comput Biol 9:e1003106. doi:10.1371/journal.pcbi.1003106
143. Stormo GD, Tan K (2002) Mining genome databases to identify and understand new gene regulatory systems. Curr Opin Microbiol 5:149–153. doi:10.1016/S1369-5274(02)00309-0
144. Swint-Kruse L, Zhan H, Fairbanks BM, Maheshwari A, Matthews KS (2003) Perturbation from a distance: mutations that alter LacI function through long-range effects. Biochemistry 42:14004–14016. doi:10.1021/bi035116x
145. Tabor JJ, Groban ES, Voigt CA (2009) Systems biology and biotechnology of *Escherichia coli*. Syst Biol Biotechnol Escherichia coli. doi:10.1007/978-1-4020-9394-4
146. Tang S-Y, Cirino PC (2011) Design and application of a mevalonate-responsive regulatory protein. Angew Chemie Int Ed 50:1084–1086. doi:10.1002/anie.201006083
147. Tang S-Y, Fazelinia H, Cirino PC (2008) AraC regulatory protein mutants with altered effector specificity. J Am Chem Soc 130:5267–5271. doi:10.1021/ja7109053
148. Tang S-Y, Qian S, Akinterinwa O, Frei CS, Gredell JA, Cirino PC (2013) Screening for enhanced triacetic acid lactone production by recombinant *Escherichia coli* expressing a designed triacetic acid lactone reporter. J Am Chem Soc 135:10099–10103. doi:10.1021/ja402654z
149. Tang SY, Cirino PC (2010) Elucidating residue roles in engineered variants of AraC regulatory protein. Protein Sci 19:291–298. doi:10.1002/pro.310
150. Taylor ND, Garruss AS, Moretti R, Chan S, Arbing MA, Cascio D, Rogers JK, Isaacs FJ, Kosuri S, Baker D, Fields S, Church GM, Raman S (2015) Engineering an allosteric transcription factor to respond to new ligands. Nat Methods 13:177–183. doi:10.1038/nmeth.3696
151. Tompa M, Li N, Bailey TL, Church GM, De Moor B, Eskin E, Favorov AV, Frith MC, Fu Y, Kent WJ, Makeev VJ, Mironov AA, Noble WS, Pavesi G, Pesole G, Régnier M, Simonis N, Sinha S, Thijs G, van Helden J, Vandenbogaert M, Weng Z, Workman C, Ye C, Zhu Z (2005) Assessing computational tools for the discovery of transcription factor binding sites. Nat Biotechnol 23:137–144
152. Tungtur S, Egan SM, Swint-Kruse L (2007) Functional consequences of exchanging domains between LacI and PurR are mediated by the intervening linker sequence. Proteins Struct Funct Bioinforma 68:375–388. doi:10.1002/prot.21412
153. Uchiyama T, Abe T, Ikemura T, Watanabe K (2005) Substrate-induced gene-expression screening of environmental metagenome libraries for isolation of catabolic genes. Nat Biotechnol 23:88–93. doi:10.1038/nbt1048
154. Ulrich LE, Koonin EV, Zhulin IB (2005) One-component systems dominate signal transduction in prokaryotes. Trends Microbiol 13:52–56. doi:10.1016/j.tim.2004.12.006
155. Vilar JMG, Leibler S (2003) DNA looping and physical constraints on transcription regulation. J Mol Biol 331:981–989. doi:10.1016/S0022-2836(03)00764-2
156. Vora S, Tuttle M, Cheng J, Church G (2016) Next stop for the CRISPR revolution: rNA-guided epigenetic regulators. FEBS J 283:3181–3193. doi:10.1111/febs.13768
157. Wang B, Barahona M, Buck M (2014) Engineering modular and tunable genetic amplifiers for scaling transcriptional signals in cascaded gene networks. Nucleic Acids Res 42:9484–9492. doi:10.1093/nar/gku593
158. Wang B, Barahona M, Buck M (2015) Amplification of small molecule-inducible gene expression via tuning of intracellular receptor densities. Nucleic Acids Res 43:1955–1964. doi:10.1093/nar/gku1388
159. Wang Z, Cirino PC (2016) New and improved tools and methods for enhanced biosynthesis of natural products in microorganisms. Curr Opin Biotechnol 42:159–168. doi:10.1016/j.copbio.2016.05.003
160. Warren CL, Kratochvil NCS, Hauschild KE, Foister S, Brezinski ML, Dervan PB, Phillips GN, Ansari AZ (2006) Defining the sequence-recognition profile of DNA-binding

molecules. Proc Natl Acad Sci USA 103:867–872. doi:10.1073/pnas.0509843102
161. Weiss DS, Batut J, Klose KE, Keener J, Kustu S (1991) The phosphorylated form of the enhancer-binding protein NTRC has an ATPase activity that is essential for activation of transcription. Cell 67:155–167. doi:10.1016/0092-8674(91)90579-N
162. Williams SG, Cranenburgh RM, Weiss AM, Wrighton CJ, Sherratt DJ, Hanak JA (1998) Repressor titration: a novel system for selection and stable maintenance of recombinant plasmids. Nucleic Acids Res 26:2120–2124
163. Williams TC, Pretorius IS, Paulsen IT (2016) Synthetic evolution of metabolic productivity using biosensors. Trends Biotechnol 34:371–381. doi:10.1016/j.tibtech.2016.02.002
164. Wilson D, Charoensawan V, Kummerfeld SK, Teichmann SA (2008) DBD–taxonomically broad transcription factor predictions: new content and functionality. Nucleic Acids Res 36:D88–D92. doi:10.1093/nar/gkm964
165. Wilson D, Pethica R, Zhou Y, Talbot C, Vogel C, Madera M, Chothia C, Gough J (2009) SUPERFAMILY–sophisticated comparative genomics, data mining, visualization and phylogeny. Nucleic Acids Res 37:D380–D386. doi:10.1093/nar/gkn762
166. Wilson RL, Urbanowski ML, Stauffer GV (1995) DNA binding sites of the LysR-type regulator GcvA in the gcv and gcvA control regions of *Escherichia coli*. J Bacteriol 177:4940–4946
167. Winkler JD, Erickson K, Choudhury A, Halweg-Edwards AL, Gill RT (2015) Complex systems in metabolic engineering. Curr Opin Biotechnol 36:107–114. doi:10.1016/j.copbio.2015.08.002
168. Wise AA, Kuske CR (2000) Generation of novel bacterial regulatory proteins that detect priority pollutant phenols. Appl Environ Microbiol 66:163–169. doi:10.1128/AEM.66.1.163-169.2000
169. Xiao Y, Bowen CH, Liu D, Zhang F (2016) Exploiting nongenetic cell-to-cell variation for enhanced biosynthesis. Nat Chem Biol 12:339–344. doi:10.1038/nchembio.2046
170. Xue H, Shi H, Yu Z, He S, Liu S, Hou Y, Pan X, Wang H, Zheng P, Cui C, Viets H, Liang J, Zhang Y, Chen S, Zhang HM, Ouyang Q (2014) Design, construction, and characterization of a set of biosensors for aromatic compounds. ACS Synth Biol 3:1011–1014. doi:10.1021/sb500023f
171. Zalatan JG, Lee ME, Almeida R, Gilbert LA, Whitehead EH, La Russa M, Tsai JC, Weissman JS, Dueber JE, Qi LS, Lim WA (2015) Engineering complex synthetic transcriptional programs with CRISPR RNA scaffolds. Cell 160:339–350. doi:10.1016/j.cell.2014.11.052
172. Zaslaver A, Bren A, Ronen M, Itzkovitz S, Kikoin I, Shavit S, Liebermeister W, Surette MG, Alon U (2006) A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. Nat Methods 3:623–628. doi:10.1038/nmeth895
173. Zhan H, Taraban M, Trewhella J, Swint-Kruse L (2008) Subdividing repressor function: dNA binding affinity, selectivity, and allostery can be altered by amino acid substitution of non-conserved residues in a LacI/GalR homologue. Biochemistry 47:8058–8069. doi:10.1021/bi800443k
174. Zhang F, Carothers JM, Keasling JD (2012) Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. Nat Biotechnol 30:354–359. doi:10.1038/nbt.2149
175. Zhang F, Keasling J (2011) Biosensors and their applications in microbial metabolic engineering. Trends Microbiol 19:323–329. doi:10.1016/j.tim.2011.05.003
176. Zhang J, Jensen MK, Keasling JD (2015) Development of biosensors and their application in metabolic engineering. Curr Opin Chem Biol 28:1–8. doi:10.1016/j.cbpa.2015.05.013
177. University of Cambridge (2009) International Genetically Engineered Machine (iGEM). http://2009.igem.org/Team:Cambridge. Accessed 20 Aug 2016