VRIJE
UNIVERSITEIT
BRUSSEL

Graduation thesis submitted in partial fulfilment of the requirements for the degree Master of Science in Bioengineering Sciences: Cell and Gene Biotechnology — Molecular Biotechnology

Modelling and response curve engineering of β-alanine-responsive biosensors

Vincent Crabbe

2021-2022

Promotor: Prof. Dr. ir. Eveline Peeters

Co-promotor: Prof. Dr. Sophie de Buyl

Faculty of Sciences and Bioengineering Sciences

Department of Bioengineering Sciences

Research Group of Microbiology

VRIJE
UNIVERSITEIT
BRUSSEL

Graduation thesis submitted in partial fulfilment of the requirements for the degree Master of Science in Bioengineering Sciences: Cell and Gene Biotechnology — Molecular Biotechnology

Modelling and response curve engineering of β-alanine-responsive biosensors

Vincent Crabbe

2021-2022

Promotor: Prof. Dr. ir. Eveline Peeters

Co-promotor: Prof. Dr. Sophie de Buyl

Faculty of Sciences and Bioengineering Sciences

Department of Bioengineering Sciences

Research Group of Microbiology

.....	.....
Vincent Crabbe (Author)	Date
.....	.....
Prof. Dr. ir. Eveline Peeters (Promotor)	Date
S. de Buyl	
.....	.....
Prof. Dr. Sophie de Buyl (Co-promotor)	Date

Preface

First of all, I would like to thank my promotor Prof. Dr. ir. Eveline Peeters for the opportunity of performing my Master's thesis at her lab on this interesting and interdisciplinary project. I deeply admire your contagious enthusiasm for science and your positive management style. Your intelligent feedback has been of great value throughout this year. I would also like to thank Prof. Dr. Sophie de Buyl for her continuous trust and for our frequent discussions on the subject of mathematical modelling. Next, I would like to thank my supervisor Dr. ir. Indra Bervoets. It was a pleasure to work alongside such an experienced researcher. You have always believed in me and you were always ready to stand in for my experiments. Thank you.

Further, I would like to thank all the members of the MICR research group. I really enjoyed the positive and spontaneous atmosphere in the group. Special thanks go to Drs. ir. Amber Bernauw. Thank you for the beautiful response curves you provided me with. I really enjoyed being your neighbour in the lab.

I would also like to thank my fellow student Alain Focquet. Thank you for being such a great friend and for letting me use your professional infrastructure to run my Python script. I was flattered by your interest in my thesis topic.

Finally, I would like to thank my loving family for supporting me throughout my studies. Mama, your care shepherded me through periods of stress and deadlines. Papa, thank you for sparking my interest in science and informatics. Simon, thank you for helping me put my thesis into perspective. I love you all.

Contents

1 Introduction **1**

2 Literature review **3**

2.1 Whole-cell biosensors 3

2.1.1 Mechanistic classification of biosensors 3

2.1.1.1 Metabolite-responsive transcription factor-based biosensors 3

2.1.1.2 Two-component regulatory system-based biosensors . 6

2.1.1.3 RNA-based biosensors 9

2.1.2 Applications of biosensors 12

2.1.2.1 Monitoring 12

2.1.2.2 High-throughput screening 14

2.1.2.3 Adaptive laboratory evolution 17

2.1.2.4 Metabolic flux regulation 19

2.2 Modelling biosensors 22

2.2.1 The Hill equation and ultrasensitivity 22

2.2.2 Characteristics of a biosensor 25

2.2.3 Dynamic modelling of gene expression 26

2.2.4 The phenomenological model of Mannan for biosensor response 29

2.3 β-Alanine-responsive transcription factors 33

3 Aims and objectives **38**

4 Materials and methods **40**

4.1 Mathematical modelling 40
4.1.1 Dose-response parameter estimation 40
4.1.2 Development of a parameter estimation method for the Mannan model for biosensor response 41
4.1.2.1 The parameter estimation method 41
4.1.2.2 Evaluation of the parameter estimation method 45
4.2 Molecular biology work 46
4.2.1 Materials 46
4.2.1.1 Buffers, media and solutions 46
4.2.1.2 Kits 49
4.2.1.3 Bacterial strains 49
4.2.1.4 Plasmid vectors 50
4.2.1.5 Oligonucleotide primers 52
4.2.1.6 Restriction enzymes 55
4.2.2 Methods 55
4.2.2.1 Cloning 55
4.2.2.2 *In vivo* characterisation of biosensors 58

5 Results and Discussion 60
5.1 Mathematical modelling 60
5.1.1 Systematic analysis of the Mannan model for biosensor response for the repressed-repressor architecture 60
5.1.2 Development of a parameter estimation method for the Mannan model for biosensor response 69
5.2 *In vivo* characterisation of biosensors 82
5.2.1 Plasmid systems used for the assembly of biosensors 82
5.2.2 OapR-based biosensors 84
5.2.2.1 Preliminary experiments to determine a suitable β-alanine concentration gradient and naringenin concentration 84
5.2.2.2 Tuning the transcription-factor binding site 88
5.2.3 AHOS_RS02205-based biosensors 98

Conclusion **104**

Summary **107**

Samenvatting **109**

I Supplementary material to Section 5.1 Mathematical modelling **138**

I.1 Additional simulations of the Mannan model 138

I.2 Additional parameter estimation method evaluation results 140

II Supplementary material to Section 5.2 *In vivo* characterisation of biosensors **151**

Chapter 1

Introduction

Microorganisms have developed a myriad of sensory and regulatory mechanisms to adapt their metabolic state to cellular and environmental cues. Synthetic biologists have tapped into these systems to construct whole-cell biosensors: microbial cells able to specifically and dose-dependently convert a chemical signal to a change in gene expression. Biosensors based on metabolite-responsive transcription factor (MRTFs) have been applied in diverse metabolic engineering applications, including real-time monitoring, high-throughput screening, adaptive laboratory evolution and dynamic pathway regulation. For the latter, biosensors are used to autonomously adapt metabolic flux by regulating key pathway enzymes in response to changes in intracellular metabolite concentration, and increased production titers, rates and yields have resulted from their implementation.

3-Hydroxypropionic acid is considered as one of the top value-added chemicals for microbial production from renewable feedstocks, and biosensors have been applied to improve its production through different pathways. However, the heterologous pathway starting from β-alanine has not yet been addressed in this context, which might be due to the lack of characterised biosensors responsive to β-alanine until recently. These biosensors could also prove beneficial in other applications, as β-alanine is a popular food supplement among athletes for improving their performance during high-intensity exercise.

In order to optimally benefit practical applications, biosensors need to be tunable. However, the effect of different interventions to the biosensor dose-response curve is complex and not entirely understood. Different mathematical models have been developed to assist in biosensor design, exposing interesting parameters to facilitate response curve engineering. Notably, Mannan *et al.* (2017) proposed that their simple phenomenological model is widely applicable across different MRTF-based biosensor architectures. Although frequently cited in literature, this model has not yet been further evaluated by independent researchers.

In this work, whole-cell biosensors are reviewed and their application in metabolic engineering is illustrated with a case study on 3-hydroxypropionic acid. Next, the use of mathematical models for the study of biosensors is discussed and the different transcription factors responsive to β-alanine are cited. Furthermore, simulations of the phenomenological model of Mannan *et al.* (2017) are studied to provide further insight in the effect of the tunable parameters of MRTF-based biosensors of the repressed-repressor architecture, and a parameter estimation method is implemented to fit biosensor characterisation data to the model. Finally, an array of β-alanine-responsive biosensors are characterised *in vivo*, and their response curve was engineered by tuning the MRTF expression level and mutating predicted transcription-factor binding sites.

Chapter 2

Literature review

2.1 Whole-cell biosensors

Microorganisms have evolved mechanisms to sense and adapt to cellular and environmental signals. Synthetic biologists have exploited this body of sensory systems to construct *whole-cell biosensors*: microbial cells that convert a chemical (e.g. concentration) or physical (e.g. temperature) signal into a change in gene expression. Depending on the application, the target genes code for proteins ranging from fluorescent reporters to enzymes. Whole-cell biosensors have been extensively studied to probe intracellular metabolite concentrations for applications in metabolic engineering (Qin *et al.*, 2022). Therefore, this review focuses on metabolite-responsive biosensors in microorganisms. This section sets out the major mechanistic classes and applications of these biosensors.

2.1.1 Mechanistic classification of biosensors

2.1.1.1 Metabolite-responsive transcription factor-based biosensors

Metabolite-responsive transcription factor (MRTF)-based biosensors apply MRTFs to regulate the transcription of a particular gene of interest. In nature, transcription factors bind to specific DNA [deoxyribonucleic acid] sequences on the genome (*transcription-factor binding sites* or TFBSs) to repress or activate transcription of target genes (Figure 2.1A). This leads to the distinction between *repressor* and *activator* transcription

factors, respectively. In addition to this *regulation module*, MRTFs encompass a *sensing module*. Typically, binding of the cognate metabolite (the *effector*) to the effector-binding domain of the MRTF relays conformational changes to the *DNA-binding domain* (DBD) (Fic *et al.*, 2009). Effector binding can either activate or reverse the activity of the MRTF, giving rise to four patterns (Figure 2.1B) (Mannan *et al.*, 2017). However, more complex mechanisms exist, where an MRTF both acts as repressor and activator for different operons (Pastan & Adhya, 1976; Platko *et al.*, 1990; Henry & Cronan Jr, 1991; Pittard & Davidson, 1991) or for the same operon in response to effector binding (Englesberg *et al.*, 1969; Ni'Bhriain *et al.*, 1983; Charlier *et al.*, 1993; Bhagwat *et al.*, 1997).

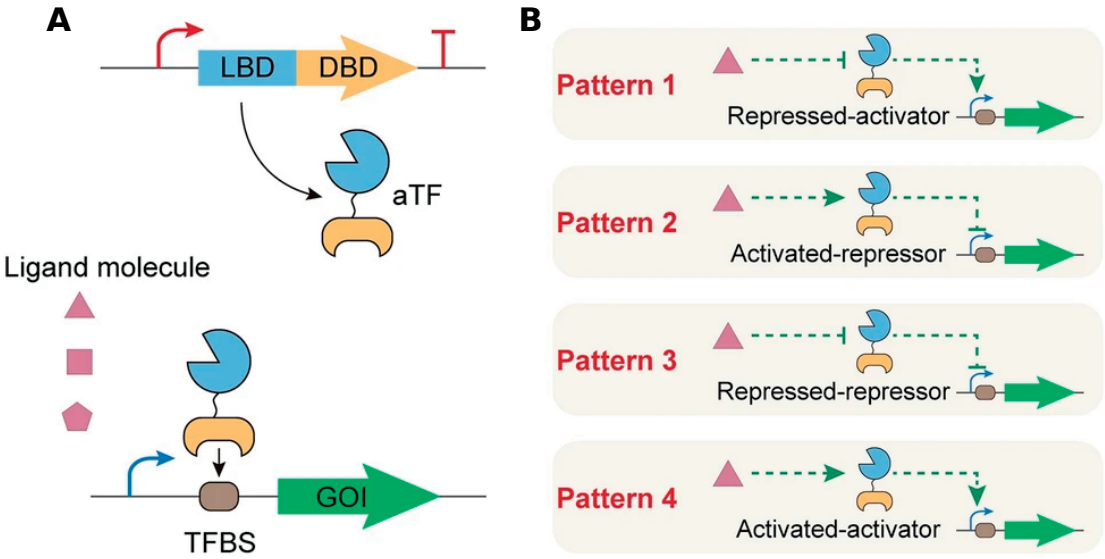

Figure 2.1: Metabolite-responsive transcription factor (MRTF)-based biosensors. **(A)** MRTFs generally consist of two domains: an effector-binding domain (EBD) responsible for binding of the effector molecule and a DNA-binding domain (DBD) responsible for interaction with the transcription factor binding site (TFBS) of a gene of interest (GOI). **(B)** Possible interactions between effector, MRTF and GOI, giving rise to four patterns of biosensors. Adapted from Wu, Du, Chen & Liu (2020).

Patterns 3 and 4 (Figure 2.1B) are overrepresented in biosensor literature as they constitute inducible systems (Wu, Du, Chen & Liu, 2020). However, some important compounds still lack a biosensor (Qin *et al.*, 2022). Hanko *et al.* (2020) formulated a method to mine metabolite-inducible transcription factors from any genome sequence to fill this gap. The researchers relied on the common feature that MRTF genes are often arranged

in a divergent orientation to the operon they regulate in order to link a putative MRTF gene to its hypothetical effector. Concretely it was assumed that if that operon encoded at least two enzymes that operate in the same metabolic pathway, the corresponding MRTF is responsive to the primary substrate metabolised by the enzymes encoded in the operon. This strategy was applied to *Cupriavidus necator* H16, a member of the *Betaproteobacteria* known for its versatile metabolism and diverse genetic regulation (Vandamme & Coenye, 2004; Pohlmann *et al.*, 2006). This way, effectors could be proposed for sixteen putative MRTFs, including predicted members of the *feast/famine regulatory protein* (FFRP) and GntR transcription factor families (see also Section 2.3). For each identified inducible system, the MRTF gene and intergenic region were amplified from the genome of *C. necator* through *polymerase chain reaction* (PCR). Biosensors were constructed by cloning this amplicon into a reporter vector so that a *monomeric red fluorescent protein* (mRFP) was placed under transcriptional control of the metabolite-inducible promoter. For fifteen of the MRTFs, a successful biosensor could be constructed, showing induced mRFP fluorescence after adding the predicted effector to the growth medium (Hanko *et al.*, 2020).

Biosensors are often sourced from an organism distinct from the host where it is applied. Heterologous biosensors expand the repertoire of perceptible metabolites and should, in principle, function orthogonal to the host's metabolism. However, differing codon usage or regulatory elements might challenge this approach (Carpenter *et al.*, 2018). Concerning previous paragraph's example, identified metabolite-inducible biosensors were tested for their functionality in the industrially important *Gammaproteobacteria Escherichia coli* and *Pseudomonas putida*. Respectively, 7 and 11 out of 15 biosensors showed a significant increase in fluorescence after ligand addition without any modification. The β-alanine (BA)-responsive MRTF OapR was insensitive to its ligand in both *E. coli* and *P. putida*. By simply replacing the upstream sequence of *oapR* with a medium-strength core promoter and strong *ribosome binding site* (RBS), a 40- and 3-fold increase in fluorescence was obtained upon ligand addition, respectively (Hanko *et al.*, 2020).

Biosensors can even be used across domains of life. Numerous reports describe the usage of biosensors sourced from *Bacteria* in the industrial production host *Saccharomyces cerevisiae* due to the lack of thoroughly characterised MRTFs in yeast and the

complexity of its transcriptional regulation network (Teo *et al.*, 2013; Umeyama *et al.*, 2013; Li *et al.*, 2015; Teo & Chang, 2015; Castaño-Cerezo *et al.*, 2020). However, this is not without challenges, as the basic transcription factor machinery is very dissimilar across domains of life, especially between *Bacteria* on the one hand and *Archaea* and *Eukaryota* on the other hand (Werner & Grohmann, 2011). Therefore, one often relies on transcriptional repression since this is easier to implement artificially. Indeed, one can easily imagine a bacterial MRTF to block transcription elongation in a eukaryotic system (according to the 'roadblock' mechanism). However, it would be more difficult to engineer the same MRTF to interface with the eukaryotic transcriptional machinery and recruit RNA polymerase (RNAP) to the promoter to activate transcription (Carpenter *et al.*, 2018). Still, the complexity of biology can lead to unintended results. In a recent publication, it was attempted to construct a biosensor in *E. coli* based on the fatty acid degradation regulator from archaeal species *Sulfolobus acidocaldarius* ($FadR_{Sa}$). In *S. acidocaldarius*, transcriptional repression by the $FadR_{Sa}$ is relieved by the specific binding of fatty acyl-coenzyme A (CoA) molecules to $FadR_{Sa}$ (Wang *et al.*, 2019). However, when attempted to transfer this system to *E. coli*, adding the effector did not induce the expression of the reporter gene *mkate2* (Sybers *et al.*, 2022). Hence, a functional biosensor could not be constructed in *E. coli* since this requires responsiveness to the effector.

MRTF-based biosensors are the most widely applied type of whole-cell biosensors (Qin *et al.*, 2022). This is due to several advantages. First, they are abundant in prokaryotes (Ulrich *et al.*, 2005) and are responsive to a vast repertoire of metabolites (Fernandez-López *et al.*, 2015). Besides, their intrinsic modularity and relative ease of engineering decrease the development time for biosensors. A drawback, however, is the relatively long response time because of the inherent time scales of transcription and translation (Liu *et al.*, 2015; De Paepe *et al.*, 2017; Carpenter *et al.*, 2018).

2.1.1.2 Two-component regulatory system-based biosensors

In *two-component regulatory system* (TCRS)-based biosensors, the sensing and regulation modules consist of two distinct proteins: a *sensor histidine kinase* and a *response regulator* (Figure 2.2). The sensor kinase is often integrated into the cytoplasmic membrane, its N-terminal *input domain* projecting into the extracellular or periplasm and the

C-terminal *transmitter domains* located in the cytoplasm (Briegel *et al.*, 2009). In response to the presence or absence of a signal sensed by the input domain, the sensor kinase autophosphorylates at a conserved histidine residue of the transmitter domain. This phosphate can subsequently be transferred to a conserved aspartate residue of the *receiver domain* of the response regulator, leading in most cases to allosteric activation of an *output domain* (Ninfa & Magasanik, 1986; Nixon *et al.*, 1986; Hess *et al.*, 1988; Kofoid & Parkinson, 1988; Stock *et al.*, 2000). Most commonly, the output domain is a DBD; as such, the response regulator constitutes a transcription factor, either a repressor or activator (Galperin, 2006). Again, TCRSs are often far more complex than presented above, as they are most often involved in larger networks (Jung *et al.*, 2012).

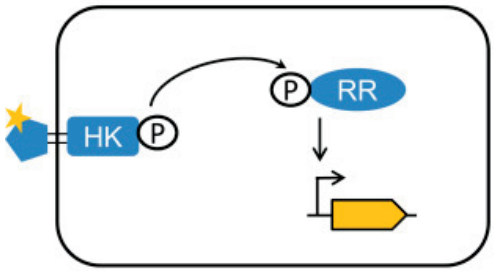

Figure 2.2: Example of a two-component regulatory system-based biosensor. After binding of a ligand (star) to the sensor histidine kinase (HK), it phosphorylates itself. The posphate group can subsequently be transferred to a response regulator (RR), activating transcription of a gene of interest in response. Adapted from Liu *et al.* (2015).

A well-studied and conserved example of a TCRS is embodied by the phosphate (Pho) regulon (Torres-Bacete *et al.*, 2021). The *E. coli* Pho regulon encompasses at least 47 genes that are induced under phosphate starvation (Lamarche *et al.*, 2008). Under limiting extracellular concentrations of inorganic phosphate (P_i), the preferred source of phosphorus, sensor kinase PhoR autophosphorylates (Wanner, 1993). PhoR~P, in turn, phosphorylates response regulator PhoB, converting it into its active state (Makino *et al.*, 1989). PhoB~P both functions as an activator and repressor of Pho regulon members, all containing a conserved *Pho box* in their promoter sequence (Makino *et al.*, 1986). For instance, the expression of *phoA* is induced under phosphate stress. It encodes bacterial alkaline phosphatase that functions in the periplasm to free P_i by hydrolysing organophosphate compounds (Wanner, 1996). Upon shifting to P_i-rich conditions, PhoR functions as a phospho-PhoB phosphatase to terminate the signal cascade (Carmany

et al., 2003). It should be noted that the currently accepted model in literature states that PhoR does not directly bind P_i; in fact, it lacks a typical periplasmic input domain altogether (Scholten & Tommassen, 1993; Uluşeker *et al.*, 2019). PhoR would sense the abundance of P_i via the PstSCAB complex (the phosphate-specific transport system) and PhoU (the chaperone-like PhoR/PhoB inhibitory protein) (Gardner *et al.*, 2014). The existence of extensive cross-regulation even enhances the complexity (Wanner *et al.*, 1988; Wanner & Wilmes-Riesenberg, 1992; Fisher *et al.*, 1995; Kim *et al.*, 1996; Verhamme *et al.*, 2002; Nishino *et al.*, 2005; Zhou *et al.*, 2005).

However complex, the Pho regulon has been used to construct TCRS-based biosensors. Torres-Bacete *et al.* (2021) constructed a synthetic P_i-starvation promoter library by replacing the -35 promoter region of the strong constitutive promoter BG42 with sequence-randomised Pho boxes (Zobel *et al.*, 2015). Degenerate primers based on a consensus sequence of Pho boxes from six different species were used to amplify the BG42 promoter to generate a combinatorial promoter library that should be portable across different species. These promoters drive the expression of the *msfgfp* (*monomeric super folder green fluorescent protein*) reporter gene (Landgraf, 2012). Selected promoters were characterised in the industrially relevant strains *E. coli* W and *P. putida* KT2440, revealing a wide range of induction strengths (often called dynamic range, see also Section 2.2.2) and host-dependent performance.

The biotechnological significance of these biosensors was demonstrated by decoupling growth and production in *P. putida*. It has previously been established that by separating a fermentation process into distinct growth and production phases, optimal productivity of the target product can be attained (Farmer & Liao, 2000; Gadkar *et al.*, 2005). Therefore, chemical inducers are often applied to manually activate the expression of the pathway enzymes when sufficient cell density has been obtained. However, adding an external inducer to the fermentation broth can complicate downstream processing or result in toxic effects (Palomares *et al.*, 2004). The synthetic P_i-depletion promoters obviate the need for such chemical inducers. During the initial growth phase, *P. putida* gradually consumed the P_i from the medium while GFP fluorescence remained at a basal level. When extracellular P_i became limiting, *P. putida* suddenly stopped growing and displayed high-level fluorescence during the expression phase. Therefore, the biomass concentration at which the shift from growth to expression occurs can be tuned by ad-

justing the initial phosphate concentration in the growth medium.

In general, TCRSs possess several interesting properties that make them suitable for application in biosensors. They offer a straightforward platform to sense membrane-impermeable compounds and are abundantly represented across *Bacteria* (Wuichet *et al.*, 2010). As the transmitter and receiver domains are generally well conserved, sensor kinases and response regulators can be interchanged to some extent (Nixon *et al.*, 1986; Kofoid & Parkinson, 1988). However, this could also lead to unwanted interactions of the TCRS-based biosensor with endogenous proteins (Wanner, 1992). Besides, the engineering and heterologous use of TCRSs is complex due to the involvement of membrane proteins and possibly unanticipated accessory proteins (Carpenter *et al.*, 2018). Finally, as with MRTFs, transcriptional regulation leads to slow response times (Liu *et al.*, 2015).

2.1.1.3 RNA-based biosensors

Last century, it has become increasingly evident that RNA [ribonucleic acid] does not merely function as a passive carrier of information, but can also adopt an active role in catalysis and gene control. *Ribozymes* can catalyse chemical reactions in the absence of proteins, and *small RNAs* (sRNAs), *T-boxes*, *RNA thermometers* and *attenuation* allow sophisticated gene regulation in *Bacteria* (Yanofsky, 1981; Kruger *et al.*, 1982; Mizuno *et al.*, 1984; Altuvia *et al.*, 1989; Grundy & Henkin, 1993). However, this review constricts the definition of *RNA-based biosensors* to include only naturally evolved and synthetic *riboswitches*. Riboswitches most commonly reside in the 5' untranslated region (UTR) of the messenger RNA (mRNA) transcript they regulate. They consist of an *aptamer domain* and an *expression platform*. The former directly binds to its cognate effector. Typically, riboswitches adopt multiple tertiary structures, one of which is stabilised on effector binding (Liberman & Wedekind, 2012). The conformational change experienced upon effector binding is transduced to the expression platform where it affects the expression of the downstream gene (Lotz & Suess, 2018).

RNA-based biosensors thus combine a sensing and regulation module in a single transcription unit. These modules can communicate through multiple mechanisms. In the first possible pattern, effector binding induces the formation of a *terminator*, leading to premature termination of transcription (Figure 2.3A) (Winkler *et al.*, 2002). Conversely, effector binding can also stabilise an alternative structure (the *antiterminator*),

leading to positive regulation (Wachsmuth *et al.*, 2013). This mechanism seems to be favoured by *Bacteria* as it allows to interrupt transcription when it is not needed, and hence save time and resources (Breaker, 2011). On the other hand, some transcripts display ligand-responsive self-cleavage activity. Effector binding can either activate or inactivate the ribozyme activity of such transcripts, leading to degradation or stabilisation of the mRNA molecule (Winkler *et al.*, 2004). Finally, the simplest RNA-based biosensors operate on the level of translation initiation. The RBS can be sequestered or liberated upon effector binding, affecting its accessibility for the ribosome and hence the expression of the target gene (Figure 2.3B) (Nahvi *et al.*, 2002; Findeiß *et al.*, 2017).

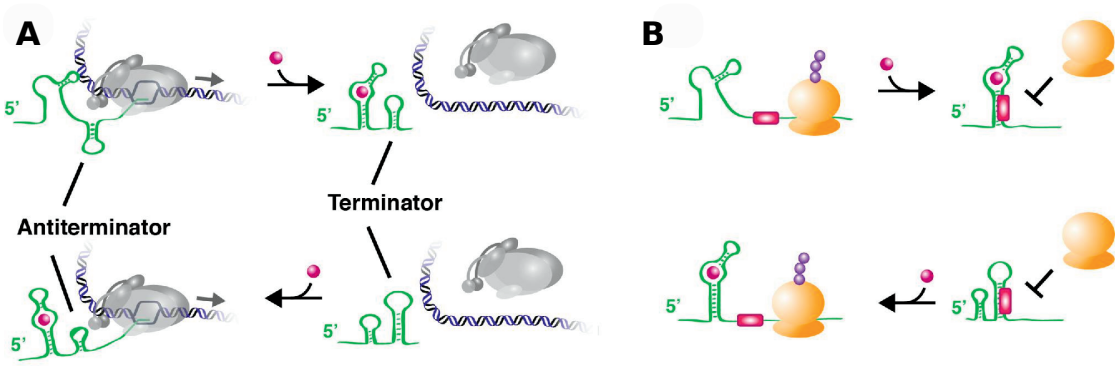

Figure 2.3: Regulatory mechanisms employed by RNA-based biosensors. **(A)** Binding of the ligand (magenta sphere) to the aptamer stabilises a terminator structure, halting transcription (above); or conversely, ligand binding stabilises an alternative structure (antiterminator), allowing transcription elongation to proceed (below). **(B)** Binding of the ligand traps the ribosome binding site (RBS, magenta box) in base-pairing, inaccessible to the ribosome (above); or conversely, ligand binding stabilises an alternative structure, freeing the RBS for the ribosome to bind to and initiate translation (below). For both panels, RNA is depected in green, RNA polymerase in grey and ribosomes in yellow. Adapted from Bervoets & Charlier (2019).

The modular nature of riboswitches has prompted many synthetic biologists to engineer artificial riboswitches *de novo*. For instance, Win & Smolke (2007) constructed synthetic RNA-based biosensors in *S. cerevisiae* by including the tobacco ringspot virus hammerhead ribozyme in the 3' UTR of the yeast-enhanced GFP gene (*yegfp*). They chose to include the ribozyme in this atypical location as excessive secondary structure in the 5' UTR represses efficient translation in eukaryotes (Pelletier & Sonenberg, 1985). Besides, cleavage in the 3' UTR results in rapid degradation of the transcript, a property

used as regulatory mechanism in *Eukaryota* and some *Bacteria* (Stoeckle & Hanafusa, 1989; Nielsen & Christiansen, 1992; Brown *et al.*, 1993; Binder *et al.*, 1994; Maeda & Wachi, 2012; López-Garrido *et al.*, 2014; Liu *et al.*, 2016; Zhu *et al.*, 2016). The ribozyme was surrounded by spacer sequences to insulate it from the flanking sequences on the transcrpt to increase the portability and modularity of the riboswitch. A synthetic aptamer responsive to theophylline was grafted in one of the loop sequences of the ribozyme in order to render its self-cleavage activity responsive to this small molecule. This way, inducible as well as repressible biosensors with a range of dynamic ranges were obtained *in vivo*, respectively. Furthermore, it was shown that for some but not all riboswitches, the aptamer could be replaced with the tetracycline miniaptamer while retaining its function. The functionality of the biosensors was proven by enabling small molecule-mediated regulation of cell growth (by including the riboswitch in the 3' UTR of the histidine biosynthesis gene *his5*) and intracellular monitoring of metabolite production (see also Section 2.1.2). For the latter application, the original theophylline-inducible biosensor was applied that shows reduced responsiveness to xanthine produced by yeast. These proof-of-concept applications show the potential of this gene-regulatory platform for diverse applications.

RNA-based biosensors have not yet reached the required maturity for mass application in metabolic engineering (Etzel & Mörl, 2017). Notwithstanding the limited knowledge of natural aptamers, synthetic aptamers can be designed to bind virtually any desired ligand through the powerful protocol designated *systematic evolution of ligands by exponential enrichment* (SELEX) (Tuerk & Gold, 1990; Lotz & Suess, 2018). However, the main bottleneck remains the signal propagation from the sensing module to the regulation module, which is challenging to engineer and not yet entirely understood (Fowler *et al.*, 2008; Ceres *et al.*, 2013; Wachsmuth *et al.*, 2013, 2015; Trausch & Batey, 2015; Hallberg *et al.*, 2017). Furthermore, RNA-based biosensors often have shortcomings concerning dynamic and operational ranges (the operational range is defined as the range of concentrations for which a dose-dependent change in output can be measured) (Hossain *et al.*, 2020). Still, RNA-based biosensors show great promise because of their faster response times and less demanding nature for cellular resources (Win & Smolke, 2007; Liu *et al.*, 2015).

2.1.2 Applications of biosensors

The fact that biosensors respond to their cognate effectors in a specific and dose-dependent manner, as discussed in Section 2.2.2, can be exploited in several applications that are divisible into four groups: *monitoring*, *high-throughput screening* (HTS), *adaptive laboratory evolution* (ALE) and *metabolic flux regulation*. In this section, these applications are explained and illustrated with a comprehensive discussion of biosensor implementations developed to increase the production of 3-hydroxypropionic acid (3-HP), a promising platform molecule. 3-HP can be produced *in vivo* through four important pathways: the CoA-dependent glycerol oxidation pathway, the CoA-independent glycerol oxidation pathway, the malonyl-CoA pathway and the BA pathway (Jiang *et al.*, 2009). Depending on the pathway at hand, investigated hosts include *E. coli*, *Klebsiella pneumoniae*, *Pseudomonas denitrificans* (nom. rej.), *Pseudomonas asiatica* and *S. cerevisiae* (Somasundar *et al.*, 2011; Rathnasingh *et al.*, 2012; Zhou *et al.*, 2013; Borodina *et al.*, 2015; Song *et al.*, 2016; Thi Nguyen *et al.*, 2021).

2.1.2.1 Monitoring

In their simplest form, whole-cell biosensors can be used to translate the presence of a specific compound to a change in gene expression, most often of a fluorescent reporter gene. Potential applications range from quality assurance in food production (Zhao *et al.*, 2021) to environmental monitoring (Ravikumar *et al.*, 2011; Fernandez-López *et al.*, 2015; Ravikumar *et al.*, 2017; Landry *et al.*, 2018; Chaurasia *et al.*, 2020; Liu *et al.*, 2020) and healthcare (Guo *et al.*, 2018; Tanna *et al.*, 2021). Monitoring of metabolite concentrations also has applications in metabolic engineering. In this context, biosensors offer the attractive feature of real-time quantification, as opposed to analytical techniques that require cumbersome sample preparation (U.S. EPA, 1996; Lee *et al.*, 2013; Rogers & Church, 2016*b*). In such a way, biosensors responsive to the desired end product can be used to follow product formation in real time, providing information to explore process parameters or mutant strains yielding increased product titres and productivities.

3-HP biosensors have been used to monitor product formation. The first 3-HP biosensors were not directly responsive to 3-HP itself because no 3-HP-responsive MRTFs were known yet. Instead, two metabolic pathways were constructed that converted 3-

HP to a detectable ligand intracellularly. The acrylate-sensitive biosensor based on the MRTF AcuR from *Cereibacter sphaeroides* emerged as the superior because of its lack of catabolite repression and sharply higher dynamic range (Sullivan *et al.*, 2011; Rogers *et al.*, 2015). Therefore, this biosensor was chosen to screen 24 production conditions for improved 3-HP titre through the malonyl-CoA pathway in *E. coli* by measuring fluorescence as a proxy for product formation. The applied concentration of cerulenin (a fatty acid biosynthesis inhibitor (Nomura *et al.*, 1972)) was varied in combination with the IPTG (isopropyl β-D-1-thiogalactopyranoside, used to induce the expression of *mcr*) concentration. The identified optimal conditions led to a 3-HP titre of 4.2 g/L, a 23-fold increase compared to earlier efforts (Rogers & Church, 2016*a*).

Later, multiple MRTFs that are directly responsive to 3-HP have been discovered and developed into biosensors (Zhou *et al.*, 2015; Hanko *et al.*, 2017, 2020; Magnus, 2021), among which the 3-HP-inducible biosensor based on the MRTF MmsR from *P. denitrificans* (Nguyen, Kim & Park, 2019; Nguyen, Ainala, Zhou & Park, 2019). Binding of 3-HP to MmsR increases its affinity for the P_{mmsA} promoter, resulting in increased expression of the reporter gene *gfp*. This biosensor was subsequently used to study 3-HP tolerance in different species. Compared to its isomer L-lactic acid, 3-HP is highly toxic at high intracellular concentrations in *E. coli*, leading to decreased growth rates and productivities. By growing *E. coli* W in the presence of increasing 3-HP concentrations in an ALE experiment, multiple mutations were accumulated in its genome, leading to much increased 3-HP tolerance. Among these mutations, disruption of the minor global transcriptional regulator gene *yieP* was entirely responsible for this increased tolerance (Nguyen-Vo *et al.*, 2019). A later study found that this effect is mainly caused by the increase in expression of the *yohJK* operon upon *yieP* mutation, that acts as a repressor of this operon. In order to obtain insight into the physiological role of YohJK, the MmsR biosensor was applied. When *yohJK* was overexpressed in *E. coli* harbouring the biosensor, a drastic decrease in GFP fluorescence was observed, corresponding to a decrease in intracellular 3-HP concentration. This observation suggests that *yohJK* encodes 3-HP exporter(s), further supported by the prediction of YohJ and YohK to be inner membrane proteins (Daley *et al.*, 2005). Notably, deletion of *yieP* yielded just as fluorescence levels (Nguyen-Vo *et al.*, 2020). Similar results were obtained in *K. pneumoniae*, where deletion of *yieP* enhanced 3-HP tolerance and production through the CoA-independent

glycerol oxidation pathway under 3-HP-stressed conditions (Nguyen-Vo *et al.*, 2022).

In addition to end product-responsive biosensors, biosensors sensitive to a limiting intermediate product can be used in metabolic engineering. In *S. cerevisiae*, the intracellular concentration of malonyl-CoA is tightly regulated, severely limiting the production of malonyl-CoA-derived products such as 3-HP (David *et al.*, 2016). Acetyl-CoA carboxylase (Acc1p), the enzyme that converts acetyl-CoA to malonyl-CoA in the cytoplasm, is regulated at the transcriptional and post-translational level (Haßlacher *et al.*, 1993). At low glucose availability, Acc1p can be phosphorylated at 15 residues causing down-regulation of its enzymatic activity (Woods *et al.*, 1994; Li *et al.*, 2014). A malonyl-CoA-responsive biosensor was developed based on the repressor FapR from *Bacillus subtilis* to improve the availability of malonyl-CoA in the cytoplasm (Schujman *et al.*, 2003). *fapR* was expressed in *S. cerevisiae* after codon-optimisation and appendage of the SV40 nuclear localisation signal to its 3' end. By introducing its TFBS *fapO* 7 bp downstream of the TATA box of a synthetic hybrid promoter consisting of the *GAL1* core promoter sequence and *TEF1* upstream activating sequence driving expression of *yegfp*, a malonyl-CoA-inducible system with low leaky expression was developed. This biosensor was used to screen thirteen phosphorylation site mutants for increased intracellular malonyl-CoA levels. When three beneficial mutations were combined (S659A, S686A and S1157A), the 3-HP titre through the malonyl-CoA pathway was increased 1.5-fold relative to the non-mutated strain. This study suggests that increased Acc1p activity benefits 3-HP production in *S. cerevisiae* (Chen *et al.*, 2018).

2.1.2.2 High-throughput screening

Previous section introduced the use of biosensors to monitor intracellular metabolite concentration in real time by measuring fluorescence as a proxy for titre. By comparing the fluorescent output of different mutants, strains with increased production of 3-HP could be identified. A downside of this approach is that these measurements occur in parallel: different mutants are spatially separated in, for instance, 96-well plates (Rogers & Church, 2016*a*). This constrains the throughput of the "test" step of the metabolic engineering cycle. Therefore, these studies could only compare a small number of mutants. However, the complexity of biology requires that many designs be assessed before an optimal producer strain can be obtained (Rogers & Church, 2016*b*).

DNA sequencing and synthesis advances have dramatically progressed the throughputs of both the ”design” and ”build” steps. However, the most significant progress would be obtained if a multiplexed solution for the ”test” step were implemented. This means that the best-performing cells can be identified and isolated from a mixture of cells containing different designs. Biosensors in combination with *fluorescence-activated cell sorting* (FACS) can be used to succeed in this objective. The biosensor couples the presence of an inconspicuous metabolite to the expression of a fluorescent reporter gene, and cells that fluoresce more brightly correlate with increased production of that target metabolite (Binder *et al.*, 2012; Rogers & Church, 2016*b*). FACS then allows sorting a mutant library in bins according to their fluorescent properties in a high-throughput fashion, resolving the screening bottleneck. Libraries consisting of up to 10^9 clones have been analysed in a single experiment, a millionfold improvement compared to conventional analytical methods (Zweigenbaum *et al.*, 1999; Santoro *et al.*, 2002).

In literature, two attempts to increase microbial 3-HP production through HTS have focused on the malonyl-CoA pathway in *S. cerevisiae*. Li *et al.* (2015) constructed a malonyl-CoA biosensor to screen a genome-wide overexpression library for genes that increased intracellular malonyl-CoA concentration, an important bottleneck in 3-HP production. The malonyl-CoA-inducible biosensor was constructed by transforming two plasmids in a 3-HP-producing yeast strain, one plasmid containing the *fapR* gene encoding a malonyl-CoA-responsive transcription factor and another containing the *tdTomato* fluorescent reporter gene under transcriptional control of a synthetic FapR-regulated promoter. Finally, an *S. cerevisiae* strain CEN.PK2-1C cDNA library was cloned under the control of the strong constitutive *TEF1* promoter on a third plasmid. This plasmid was transformed in the whole-cell biosensor, followed by overnight incubation and FACS screening.

After three rounds of FACS, each time sorting the cells showing high-level fluorescence, plasmids from 72 colonies were isolated for sequencing to identify the overexpressed genes. Of this sampling, only two genes could be identified to correspond with an increase in fluorescence: *PMP1* encoding an important subunit of the proton-ATPase PMA1, and *TPI1* coding for the enzyme triose-phosphate isomerase. When these genes were individually overexpressed in *S. cerevisiae*, 3-HP titres through the malonyl-CoA pathway were improved by 100% and 120%, respectively, after 72 h. It was hypoth-

esised that overexpression of *PMP1* led to an improved proton gradient, accelerating the uptake of biotin, a crucial cofactor of acetyl-CoA carboxylase. *TPI1*, on the other hand, would improve ATP production by redirecting the carbon flux from dihydroxyacetone phosphate to glycerol, to pyruvate. As one molecule of ATP is consumed per molecule of malonyl-CoA produced, an increased concentration of ATP should enhance malonyl-CoA accumulation. Interestingly, co-overexpression of *PMP1* and *TPI1* only led to a 47% increase in 3-HP titre. The authors attributed this antagonistic effect to the basification of the cytosol due to increased PMP1p activity, resulting in decreased Tpi1p activity. In addition, the cellular burden associated with the simultaneous overexpression of two genes might adversely affect 3-HP production. In conclusion, two novel gene targets for enhanced 3-HP production were identified from a genome-wide overexpression library by HTS (Li *et al.*, 2015).

In contrast to the ”brute force” screening approach followed in the previous example, Ferreira *et al.* (2019) applied a more rational approach guided by *flux balance analysis* (FBA). This way, 168 genes were *in silico* predicted as candidates for up- or downregulation to enhance the flux toward malonyl-CoA. To alter the expression of these target genes *in vivo*, dCas9 [’dead’ CRISPR (*clustered regularly interspaced short palindromic repeats*)-associated protein 9]-based regulation was applied, where dCas9 is fused to the tripartite activator VP64-p65-Rta (VPR). This synthetic transcription factor can be conveniently targeted to any genomic locus using a complementary *guide RNA* (gRNA), to which dCas9 binds. Interestingly, dCas-VPR can both act as activator and repressor, depending on the position with respect to the *transcription start site*. As it remains hard to predict the regulatory outcome for a given gRNA *a priori*, up to 21 gRNAs were designed per selected gene to sample a gradient of regulatory activities, giving rise to a library of 3194 gRNAs. This library was co-transformed with a FapR-based malonyl-CoA biosensor (David *et al.*, 2016) in *S. cerevisiae* CEN.PK-11C in order to enrich for gRNAs that increased the availability of malonyl-CoA. At each stage of the screening, the gRNAs were sequenced to reveal the dynamics of the sorting process. Only 49 gRNAs were significantly enriched, corresponding to 46 genes. Notably, both gene targets discovered in the previous study were on the shortlist based on FBA but did not yield significantly enriched gRNAs. The selected gRNAs were retransformed in *S. cerevisiae* and tested for their ability to increase the production of 3-HP from glucose. More than

a third of the gRNAs did not result in increased 3-HP titres. The ones that did, targeted genes with functions ranging from cofactor supply to production of malonyl-CoA precursors.

Without explicitly referring to Li *et al.* (2015), the researchers compared the effect on the 3-HP yield of their approach based on dCas9-VPR to the conventional approach for expression tuning based on the strong *TEF1* promoter. For *ALD2*, for which FBA suggested that upregulation would increase the flux to malonyl-CoA, the TEF1-based approach did not result in increased 3-HP titre, whereas the dCas9-based approach did. On the other hand, the method of Ferreira *et al.* (2019) relying on FBA primarily exposes well-annotated gene targets, neglecting unknown or not thoroughly characterised targets. However, the gRNA that led to the most significant increase in 3-HP production serendipitously regulated a second target gene: *HTA1*, encoding an essential histone protein. Notably, a second gRNA did not regulate its predicted gene at all but increased 3-HP production through upregulation of *SPO23*, a gene of unknown function. These results suggest that, although valuable to make a rough selection of gene targets, FBA can still fail to predict highly relevant genes. In conclusion, HTS leveraging whole-cell biosensors in FACS screenings has been used to improve bioproduction successfully, yet concrete applications remain challenging due to the high prevalence of false positives. Several publications have addressed this issue, promising improved HTS campaigns leading to ever-improving product titres (Raman *et al.*, 2014; Woolston *et al.*, 2018; Flachbart *et al.*, 2019).

2.1.2.3 Adaptive laboratory evolution

The previous sections introduced discrete successes in the rational engineering of metabolism towards increased product (3-HP) formation. However, rational engineering remains difficult due to the complexity underlying physiology and the need for robust methods of genetic modification. These challenges can be overcome by ALE. In brief, ALE comes down to culturing cells for an extended period of time in a desired environment. According to the principles of natural selection, microorganisms isolated at the end of the cultivation period will have accumulated mutations leading to improved fitness under the experimental conditions. Therefore, ALE does not require *a priori* knowledge to guide strain development and optimisation. A body of studies has harnessed ALE

to improve growth rate (Yu *et al.*, 2013; LaCroix *et al.*, 2015; Pfeifer *et al.*, 2017) and increase stress tolerance (Riehle *et al.*, 2003; Alcántara-Díaz *et al.*, 2004; Çakar *et al.*, 2005; Jansen *et al.*, 2005; Stoebel *et al.*, 2009; Zorraquino-Salvo *et al.*, 2014; Nguyen-Vo *et al.*, 2019) or substrate utilisation (Sonderegger & Sauer, 2003; Lee & Palsson, 2010; Hong *et al.*, 2011; Utrilla *et al.*, 2012; Latif *et al.*, 2015; Cordova *et al.*, 2016; Rajaraman *et al.*, 2016; Sandberg *et al.*, 2017). However, optimising product titres through ALE is far less evident, as the cellular burden associated with high-level production of a certain metabolite is most likely to decrease fitness (Sandberg *et al.*, 2019). This can be resolved by coupling the production of the metabolite of interest to overall energy and biomass generation. However, genetic modifications to achieve such coupling again require a thorough knowledge of metabolism and can be highly complex (Burgard *et al.*, 2003; Pharkya *et al.*, 2004; Fong *et al.*, 2005; Jantama *et al.*, 2008; Otero *et al.*, 2013; von Kamp & Klamt, 2017).

The use of whole-cell biosensors poses an elegant solution to this problem. An end product-sensitive biosensor can be used to drive the expression of a selector gene, e.g. encoding antibiotic resistance. By exposing microorganisms carrying the biosensor construct to increasing concentrations of the corresponding antibiotic, mutants with increased productivity are selected. This way, appropriate application of ALE can potentially substitute the ”design” and ”build” steps of the metabolic engineering cycle entirely.

Seok *et al.* (2018) developed a biosensor based on the 3-HP-responsive transcription factor MmsR that induced the expression of the tetracycline resistance gene in response to 3-HP accumulation. This biosensor was used to increase 3-HP titres of the CoA-independent glycerol oxidation pathway through ALE. During a culture experiment that lasted over seven days, the tetracycline concentration in the medium was gradually increased to evolve the strain towards higher 3-HP production. The genomes of two isolates with greatly enhanced 3-HP titres and yields with respect to the parental strain were sequenced to reveal the mutations underlying their superior production characteristics. The strains, named EV1 and EV2, acquired 34 and 29 mutations in their genome, respectively. Among these, EV1 exhibited a mutation in *cyaA*, encoding adenylyl cyclase, while EV2 showed a mutation in *crp*, encoding cyclic adenosine monophosphate (cAMP) receptor protein. These mutations seem connected, as adenylyl cyclase catal-

yses the conversion of adenosine triphosphate (ATP) to cAMP, the effector of MRTF CRP (Zubay *et al.*, 1970; Seok *et al.*, 2021).

To further study the effect of these mutations, they were separately introduced into the parental strain. The mutation affecting *cyaA* (S254R, situated in the catalytic domain of adenylyl cyclase (Crasnier *et al.*, 1994)) led to a drop in intracellular cAMP accumulation through a decreased catalytic efficiency. On the other hand, the mutated *crp* (T168P, situated in the DBD (McKay & Steitz, 1981)) led to decreased binding of CRP to the promoter region of *gapA*, encoding glyceraldehyde-3-phosphate dehydrogenase A, one of the regulon members of CRP. Because of the importance of CRP as a global regulator and adenylyl cyclase as the enzyme producing its effector (Gottesman, 1984), the authors conducted transcriptome analysis to identify differentially expressed genes in both mutant strains. 78 and 359 genes, of which 34 were overlapping, were differentially expressed in the *cyaA* and *crp* mutants, respectively. In both cases, *gapA* expression was decreased to 25% compared to the parental strain. It was hypothesised that the downregulation of *gapA* redistributed carbon flux to 3-HP production and decreased acetate accumulation from overflow metabolism. Besides, the *crp* mutation led to differential expression of several other genes involved in central carbon metabolism and transport. Finally, both mutations were combined, and the constructed strain was examined for increased production of 3-HP from glycerol supplemented with 1 g/L yeast extract during fed-batch fermentation. 55 g/L 3-HP was produced at a yield of 0.91 g/g, 93% of the theoretical maximum. No acetate accumulation could be registered throughout the entire bioprocess. In conclusion, this study applied ALE to reveal non-intuitive targets for metabolic engineering that led to an appreciable increase in 3-HP production. Importantly, as biosensor-based ALE mediates cell growth to product formation, the engineered strains did not display severe growth defects, in contrast to classical metabolic engineering endeavours (Seok *et al.*, 2021).

2.1.2.4 Metabolic flux regulation

As already stated earlier, high-level production of a metabolite can significantly constrain the growth rate of the producer strain. The fact is that traditional metabolic engineering efforts have often increased product titres, yields and productivities by constitutively overexpressing the enzymes of the pathway of interest in addition to knocking

out genes encoding competitive pathways (Song *et al.*, 2016). However, this can lead to metabolic imbalances such as competition for cellular resources between the production pathway and biomass synthesis and to accumulation of pathway intermediates (Glick, 1995; Martin *et al.*, 2003). This observation has incited metabolic engineers to tune the expression level of pathway enzymes relative to each other in order to optimise the pathway flux (Anthony *et al.*, 2009; Ajikumar *et al.*, 2010; Curran *et al.*, 2013; Nowroozi *et al.*, 2014). Still, such static pathway control is optimised to specific laboratory conditions, while real-life bioprocesses cannot guarantee the applicability of these conditions uniformly throughout the fermentor (Delvigne *et al.*, 2014; Pigou & Morchain, 2015).

These problems have inspired synthetic biologists to regulate metabolic flux in a more dynamic way using whole-cell biosensors. In brief, biosensors are used to regulate the expression of key enzymes in response to cellular or environmental cues. The resulting dynamic pathway regulation is a promising avenue to optimise pathway production, as evidenced by the recent surge of publications in this field (Liu & Zhang, 2018; Maury *et al.*, 2018; Moser *et al.*, 2018; Wang *et al.*, 2018; Otero-Muras *et al.*, 2019; Boada *et al.*, 2020; Liang *et al.*, 2020; Lv *et al.*, 2020; Qin *et al.*, 2020; Wu, Chen, Liu, Tian, Lv, Li, Du, Chen, Ledesma-Amaro & Liu, 2020; Xu *et al.*, 2020; Zhang *et al.*, 2020; Hao *et al.*, 2021; Lai *et al.*, 2021; Mannan & Bates, 2021; Torres-Bacete *et al.*, 2021; Zhu *et al.*, 2021; Ferrer *et al.*, 2022; Ni *et al.*, 2022; Verma *et al.*, 2022). Dynamic regulation has also been applied to increase 3-HP production through the malonyl-CoA pathway. Being heterologous to *S. cerevisiae*, this pathway is evidently not regulated by the host and biosensors allow to implement such regulation artificially.

One application has focused to implement a two-stage metabolic control system in response to the concentration of glucose. Separating a bioprocess in distinct growth and production phases relieves the competition for cellular resources between these two objectives, and often leads to enhanced product titres (see also Section 2.1.1.2). One possible approach to implement a two-stage process is to use promoter(s) activated under glucose limitation to control the expression of the heterologous pathway. Maury *et al.* (2018) identified promoters with glucose-dependent expression profiles by reanalysis of genome-wide transcription data in *S. cerevisiae* under glucose limitation and excess. 34 promoters corresponding to genes with interesting expression profiles were further studied by cloning these upstream of a quickly maturing and rapidly degrading GFP gene

variant. The resulting biosensors allowed to study the dynamic behaviour of the identified promoters by measuring fluorescence of cell cultures during microscale batch and fed-batch cultivation with glucose as sole carbon source and revealed distinct expression profiles. Three promoters showing significant activation under glucose limitation were used to drive the expression of *mcr*, the key enzyme of the malonyl-CoA pathway. Of these, regulation by the *ICL1* promoter led to a 70% increase in 3-HP titre as compared to the constitutive *PGK1* promoter during microscale fed-batch fermentation. Intriguingly, P_{ICL1} is characterised by a short but strong burst of activation upon glucose deprivation, whereas P_{PGK1} shows a longer window of activation. In conclusion, endogenous promoters of *S. cerevisiae* were identified using transcriptomics and reporter studies and were subsequently used for dynamic regulation of 3-HP production. Although this approach does not reveal the underlying regulatory mechanisms and is therefore less flexible, an increase in 3-HP titre could be reported. Of course, the reported promoters should in principle be applicable for the production of any product that merits two-stage production.

Another study has combined a two-stage metabolic control system with continuous control of *mcr* expression. For the latter, a malonyl-CoA biosensor based on *B. subtilis* FapR was developed by inserting three *fapO* TFBSs in the consitutive *TEF1* promoter driving expression of *mcr*. This way, expression of *mcr* is tailored to the intracellular malonyl-CoA availability mediated by FapR. In theory, the resulting continuous control should allow to balance metabolic fluxes automatically while reducing intermediate metabolite accumulation as well as metabolic burden from enzyme expression. However, implementation of the continuous control led to a decrease in 3-HP titre compared to constitutive expression of *mcr* when tested during microscale fed-batch cultivation of *S. cerevisiae* CEN.PK113-11C. This was attributed to insufficiently high intracellular malonyl-CoA concentrations to induce *mcr* expression adequately. A second layer of metabolic flux regulation was implemented by placing the fatty acid synthase 1 gene (*FAS1*) under the control of P_{HXT1}, a promoter repressed by low glucose availability (Ozcan & Johnston, 1995). This allows toe downregulate fatty acid biosynthesis upon glucose limitation leading to accumulation of malonyl-CoA. When both control layers were implemented simultaneously, production of 3-HP was increased tenfold, to 1.0 g/L. In conclusion, 3-HP production in *S. cerevisiae* was dynamically regulated through a com-

bination of two-stage and continuous metabolic control. During the initial growth phase, biomass could be produced with metabolic fluxes similar to those experienced under native conditions. Deprivation of glucose subsequently lead to accumulation of malonyl-CoA. This finally induced *mcr* expression, initiating the production phase (David *et al.*, 2016).

Many other control topologies have been implemented in literature for a variety of production pathways (Liu *et al.*, 2018; Xu, 2018; Han & Zhang, 2020; Jiang *et al.*, 2020; Hartline *et al.*, 2021). However, these two examples should suffice to show the main strategies employed and increased titres, yields and productivities obtained through the dynamic regulation of metabolic flux by whole-cell biosensors.

2.2 Modelling biosensors

Mathematical modelling has received increased attention in synthetic biology since the start of this century. Key strengths are its potential to predict behaviour of complex systems or reveal inconsistencies between the current model and the observed experimental data. Model simulations can guide experimental design and allow the scientist to allocate his expensive resources more efficiently (Ingalls, 2018). The focus of this section is the use of modelling for biosensor design. We start with a discussion of the Hill equation, an important formalism for biosensor modelling (Section 2.2.1). We then illustrate the use of the Hill function to fit dose-response curves directly or as an element in more complex models (Sections 2.2.3 and 2.2.4).

2.2.1 The Hill equation and ultrasensitivity

The Hill equation was introduced in 1910 to fit the binding curves of oxygen to haemoglobin that show a clear sigmoidal shape, in contrast to the hyperbolic binding curves that are normally observed. Archibald Hill found that an equation of the form

$$y = 100\frac{Kx^n}{1 + Kx^n} \tag{2.1}$$

where y be the percentage saturation of haemoglobin with oxygen, x the partial ten-

sion of oxygen and K and n empirical constants, fits the experimental data satisfactorily for *Hill coefficients* n ranging from 1.0 to 3.2 (Hill, 1910). Later, it was found that haemoglobin functions as a homotetramer and that it binds oxygen in a positively cooperative manner (Perutz, 1980). The latter means that binding a single molecule of oxygen increases the apparent affinity for further binding, giving rise to the S-shaped appearance of the binding curves.

For $n > 1$, the Hill function experiences a larger slope at intermediate input levels, a phenomenon called *ultrasensitivity* (Figure 2.4). Negative cooperativity can be modelled for $0 < n < 1$. Negatively cooperative binding curves show a subsensitive response to variations in ligand concentration (Zhang *et al.*, 2013). $n = 1$ returns the hyperbolic Langmuir equation (the binding equivalent of the Michaelis-Menten equation for enzyme kinetics).

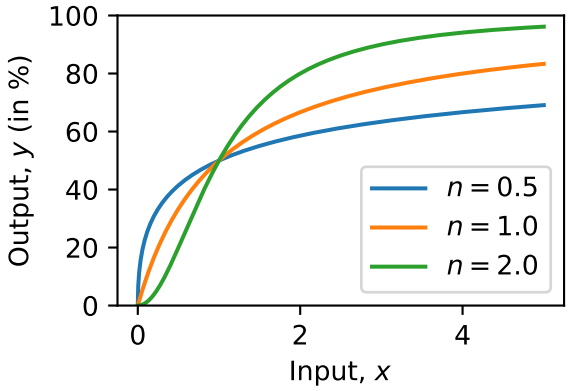

Figure 2.4: Simulations of the Hill function (Equation 2.1). In each case, K is set to 1.

Besides positive cooperative binding, signal transduction pathways and gene regulatory circuits can combine discrete mechanisms that could give rise to ultrasensitivity. These can be grouped by homo-multimerisation, multistep signalling, molecular titration, covalent modification cycles and positive feedback and can be fitted by adapted forms of the Hill equation (Zhang *et al.*, 2013).

Instead of considering distinct mechanisms that lead to a Hill equation pattern, Frank (2013) proposed common themes that could underly ultrasensitivity. Briefly, by including a stochastic term in the Michaelis-Menten equation to account for failure of the system to detect or transmit the input signal, sigmoid input-output relations were retrieved.

The sigmoid character was increasingly evident for more reactions in the signal cascade or greater failure rates of signal transduction. More strikingly, ultrasensitivity may arise from departure from the spatial homogeneity assumption of the *law of mass action*. The law of mass action claims that the rate of an elementary chemical reaction varies as the product of the concentrations of the participating reagents. As an example, consider the irreversible reaction of two molecular species A and B to yield P:

$$\mathrm{A} + \mathrm{B} \xrightarrow{k} \mathrm{P} \tag{2.2}$$

The rate v (Ms^{-1}) of reaction 2.2 then becomes

$$v = \frac{d[\mathrm{P}]}{dt} = k[\mathrm{A}][\mathrm{B}] \tag{2.3}$$

where k denotes the *mass action rate constant* of reaction 2.2 ($M^{-1}s^{-1}$), t denotes time (s) and the concentration of each chemical species X is denoted as [X] (M). Equation (2.3) can also be read as "the rate of a reaction is proportional to the probability by which the reactants collide with one another" (Ingalls, 2018). Amongst others, the well-known Michaelis-Menten equation is based on mass action (Michaelis & Menten, 1913).

However, application of the law of mass action requires two assumptions to be made. *Spatial homogeneity* requires that the concentrations of all chemical species be independent of the position in the reaction volume. Therefore, the reaction rate will also be independent of the positional coordinate. Secondly, the *continuum hypothesis* assumes that the species concentrations are sufficiently high that discrete changes in molecule count can be approximated as continuous changes in concentration. On the level of individual cells, the latter requirement is likely not to be satisfied, as the expression level of a particular protein is often low (for instance, the *lac* repressor is present at about 5 tetramers per cell) (Gilbert & Müller-Hill, 1966; Halling, 1989). This inconsistency can be overcome by considering the average behaviour of a cell culture (Ingalls, 2018). Still, diffusion limitations can arise in the cell through molecular crowding, compromising the assumption of a "well-mixed" reaction volume (Elowitz *et al.*, 1999; Kuthan, 2001). As an alternative to mass action, spatially explicit biochemical dynamics can

be simulated using specialised software (Andrews & Bray, 2004). By spatially explicit modelling of a simple reaction scheme under diffusion-limited conditions, input-output relations with a Hill coefficient as high as 2.4 could be retrieved (Frank, 2013).

In conclusion, while initially proposed as an empirical description of cooperative binding, the Hill equation can be used to model a variety of biological mechanisms. More interestingly, sigmoid input-output relations can arise in absence of a clear mechanistic description because of stochastic failure of signal transmission or limited diffusion. The arguments in this section aimed to prove that a hill coefficient $n > 1$ does not necessarily point to positive cooperativity.

2.2.2 Characteristics of a biosensor

Biosensor dose-response curves have an S-shaped appearance; hence, it should be of no surprise that the Hill equation is often used to fit this relationship:

$$P(M) = \begin{cases} b + a\frac{M^n}{\theta^n + M^n} & \text{for ligand-inducible biosensors,} \\ b + a\frac{\theta^n}{\theta^n + M^n} & \text{for ligand-repressible biosensors.} \end{cases} \tag{2.4}$$

In these equations, P refers to the expression level of the target protein and M to the applied metabolite concentration (mM). A fit to the Hill function thus summarises the response curves to four *dose-response parameters*: b, a, θ and n (Figure 2.5). The *basal output* b is defined as the minimum value of the dose-response curve. The *maximum increase in output* a is defined as the difference between the maximum and minimum values of the dose-response curve. The ratio of both defines the *dynamic range* μ:

$$\mu = \frac{a}{b} \tag{2.5}$$

The *threshold* θ is the metabolite concentration for which 50% of the maximum output is attained relative to the basal level. The Hill coefficient n, on the other hand, relates to the overall steepness of the dose-response curve (Mannan *et al.*, 2017). θ and n together determine the *operational range* of the biosensor. Biosensors with small Hill coefficients exhibit large operational ranges. Such *analog-like behaviour* is typically useful for monitoring metabolites over a large range of concentrations and HTS, ALE and metabolic

flux regulation, where a differential response is required for increasing concentrations. On the other hand, biosensors with large Hill coefficients display *digital-like behaviour* resembling on/off switches. They can be used to monitor the concentration of a metabolite in a more qualitative way or to turn target gene expression on or off in response to effector supplementation (Dietrich *et al.*, 2010).

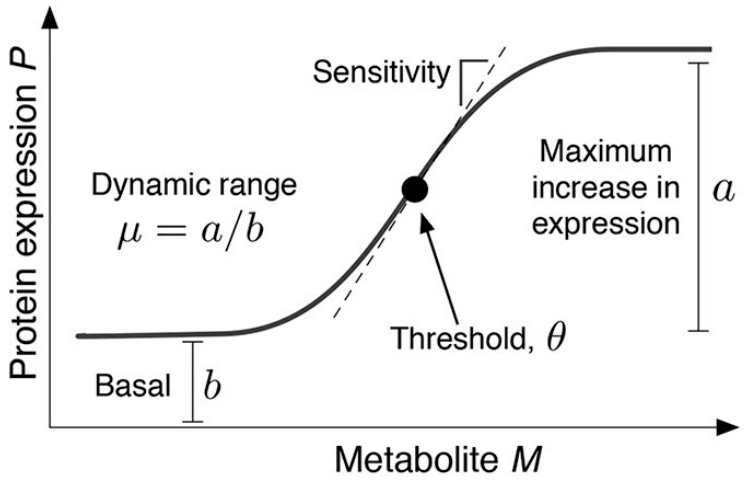

Figure 2.5: Biosensor dose-response curve with indication of the dose-response parameters. See main text for definition of these parameters. Adapted from Mannan *et al.* (2017).

Besides the response curve's shape, its molecular specificity is critical for the proper functioning of the biosensor (Rogers *et al.*, 2016). Only if the biosensor is responsive to a single metabolite or if the presence of any additional ligands can be excluded, can the dose-dependent output be solely attributed to the metabolite of interest (De Paepe *et al.*, 2017).

2.2.3 Dynamic modelling of gene expression

Modelling of gene regulatory networks is often addressed using *dynamic models*. A dynamical model describes how the abundances of the included components vary in time, in the form of a set of differential equations. As a simplification, interactions between components can be modelled by the law of mass action (cfr. Section 2.2.1). In the current section, this approach is illustrated with a simple model of gene expression

(Ingalls, 2018).

Consider the expression of a gene in *Bacteria* (Figure 2.6). Modelling all distinct reactions of transcription and translation would lead to far too complex models; instead, transcription and translation are modelled as zeroth- and first-order reactions, respectively. The degradation of mRNA and proteins balances transcription and translation. In addition, cell growth causes dilution of the mRNA and protein pools. Decay and dilution are collectively represented by the parameter δ_i (where $i = m$ for the transcript and $i = p$ for the protein). It is assumed that the "background expression machinery", containing nucleic acids, RNAP enzymes and ribosomes, is constant. Applying mass action then gives

$$\begin{aligned} \frac{d}{dt}m(t) &= v_m - \delta_m m(t) \\ \frac{d}{dt}p(t) &= k_1 m(t) - \delta_p p(t) \end{aligned} \tag{2.6}$$

where m and p signify the concentration of the considered mRNA transcript and protein product, respectively. v_m represents the rate of transcription of the gene of interest, while k_1 represents the per-mRNA translation rate. In the case of constitutive expression, v_m can be represented by a constant value k_0 (Equation 2.7). However, the majority of bacterial genes is regulated by transcription factors. Without knowledge of the precise mechanism, repressible transcription can be modelled by an inhibitory hill function, while activated transcription can be modelled by an activatory hill function. For simplicity, a ligand-insensitive transcription factor is considered.

$$v_m = \begin{cases} k_0 & \text{for constitutive expression} \\ \alpha_0 + \alpha \frac{1}{1+([\mathrm{TF}]/K)^n} & \text{for a repressor} \\ \alpha_0 + \alpha \frac{([\mathrm{TF}]/K)^n}{1+([\mathrm{TF}]/K)^n} & \text{for an activator} \end{cases} \tag{2.7}$$

In Equation (2.7), α_0 and α indicate the basal and maximum increase in transcription rate, while K designates the *apparent dissociation constant*. These equations assume that the transcription factor (TF) is in equilibrium with its TFBS(s). This assumption is justified, as transcription factor binding/unbinding occurs on a much faster time scale

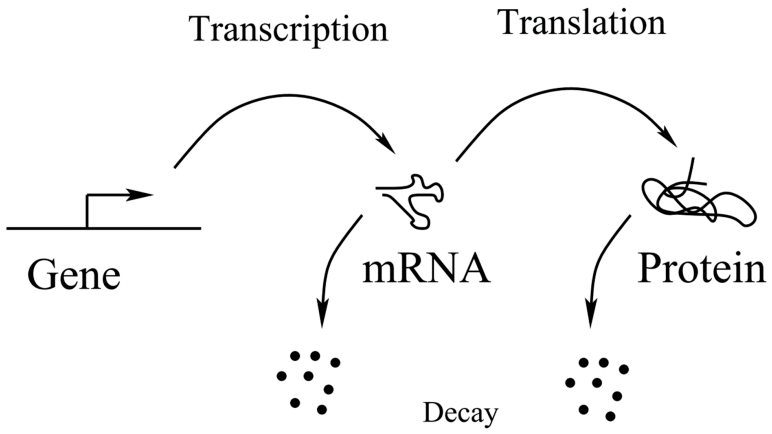

Figure 2.6: Central dogma of molecular biology. A gene (DNA) is transcribed to messenger RNA (mRNA) that can subsequently be translated to polypeptides (protein). Production of mRNA and protein are balanced by their decay. From Ingalls (2018).

than transcription.

In Equation (2.6), the mRNA concentration can be assumed in quasi-steady state as mRNA decay is usually much faster than protein decay. This assumption reduces Equation (2.6) to

$$m^{qss}(t) = \frac{v_m}{\delta_m} \tag{2.8}$$

$$\frac{d}{dt}p(t) = k_1 m^{qss}(t) - \delta_p p(t) \tag{2.9}$$

where the quasi-steady state concentration of mRNA transcripts was denoted as m^{qss} for clarity. As an example, consider the case of a repressor. Upon the substitution of Equation (2.7) in Equation (2.8), we obtain Equation (2.10).

$$m^{qss}(t) = \frac{\alpha_0}{\delta_m} + \frac{\alpha}{\delta_m}\frac{1}{1 + ([\mathrm{TF}]/K)^n} \tag{2.10}$$

Substituting Equation (2.10) once again in Equation (2.9), Equation (2.11) is obtained, which describes the time dependency of the expression of a protein under the control of

a constitutive repressor.

$$\frac{d}{dt}p(t) = \frac{\alpha_0 k_1}{\delta_m} + \frac{\alpha k_1}{\delta_m}\frac{1}{1 + ([\mathrm{TF}]/K)^n} - \delta_p p(t) \tag{2.11}$$

The behaviour of the system can be analysed by solving Equation (2.11) for certain initial conditions. As more mechanistic details of this hypothetical genetic circuit are unveiled, the model can be further refined. Conversely, if the model does not agree satisfactorily with experimental findings, the assumptions of the model have to be refined. In the long term, the concentrations often approach a constant value: the steady state. Steady states are of particular interest to the modeller and can be calculated by setting Equation (2.11) to zero:

$$\begin{aligned} p^{ss} &= \frac{\alpha_0 k_1}{\delta_m \delta_p} + \frac{\alpha k_1}{\delta_m \delta_p}\frac{1}{1 + ([\mathrm{TF}]/K)^n} \\ &= \beta_0 + \frac{\beta}{1 + ([\mathrm{TF}]/K)^n} = \beta_0 + \frac{\beta \cdot K^n}{K^n + [\mathrm{TF}]} \end{aligned} \tag{2.12}$$

where p^{ss} designates the steady-state concentration of the target protein, $\beta_0 = \frac{\alpha_0 k_1}{\delta_m \delta_p}$ and $\beta = \frac{\alpha k_1}{\delta_m \delta_p}$.

Recently, a dynamic model development workflow was published named *generation and analysis of models for exploring synthetic systems* (GAMES). It encompasses all steps from model formulation to parameter fitting and comparison of multiple candidate models, along with remediation strategies in case of failure in one of the steps. The researchers made their code freely accessible to improve the rigour and reproducibility of model development. Such efforts are precious to lower the barrier for biologists to start modelling and contribute to the progress of synthetic biology (Dray *et al.*, 2022).

2.2.4 The phenomenological model of Mannan for biosensor response

The dose-response curve of a particular biosensor most likely has to be engineered in order for the biosensor to be harnessed for different applications (see Sections 2.1.2 and 2.2.2). Indeed, native MRTFs have evolved to function optimally in their biological

context. Synthetic biologists generally isolate these MRTFs for an application discrete from their natural function and hence, their dose-response parameters are presumably suboptimal for the intended application (Mahr & Frunzke, 2016). Dose-response engineering, however, remains challenging due to several reasons. Firstly, a lot of biosensor components can be targeted for engineering, both in the sensing and regulation modules (see Section 2.1.1.1), but the expected outcome has not always been properly understood (De Paepe *et al.*, 2017). Secondly, changes to one component often affect multiple dose-response parameters at the same time (Wang *et al.*, 2015). Together, these aspects complicate rational biosensor design and therefore, response curve engineering has often been undertaken through trial-and-error methodologies (Mannan *et al.*, 2017; Koch *et al.*, 2019).

In 2017, a notable paper was published that aimed to better characterise the interplay between the tunable parameters and dose-response curves of MRTF-based biosensors. Mannan and colleagues formulated a phenomenological model to describe the dose-response relationships of all biosensor patterns in Figure 2.1B. This model (referred to as "the Mannan model for biosensor response" throughout this thesis) represents the input-output relations of the sensing and regulation modules as two separate Hill functions. That is, the concentration of active transcription factor TF is modelled as a function of metabolite concentration M, and the expression level of the target gene P is again modelled as a function of TF. P as well as the corresponding mRNA transcripts are assumed to be in steady state. Since both modules can hold negative or positive interactions, inhibitory or activatory Hill functions can once again be used (see also Equations 2.4 and 2.7):

$$f_i(x) = \begin{cases} b_i + \frac{a_i}{1+(K_i \cdot x)^{n_i}} & \text{(inhibitory Hill function),} \\ b_i + \frac{a_i \cdot (K_i \cdot x)^{n_i}}{1+(K_i \cdot x)^{n_i}} & \text{(activatory Hill function),} \end{cases} \tag{2.13}$$

for $i = \{1, 2\}$. $i = 1$ refers to the sensing module, whereas $i = 2$ refers to the regulation module. x refers to the inputs (M or TF, respectively) and $f_i(x)$ to the outputs (TF or P, respectively). Further, b_1 signifies the basal level and a_1 the maximum increase of transcription factor activity. K_1 is the transcription factor-metabolite affinity, while n_1 betokens the sensitivity of transcription factor-metabolite binding. On the other hand, b_2

models the basal level of promoter expression. a_2 is the maximum increase in promoter expression. K_2 represents the affinity of the active transcription factor for its TFBS and n_2 refers to the sensitivity of transcription factor-TFBS binding. The target gene expression can then be written as the composite function $P = f_2(f_1(M))$. As an example, consider pattern 3 in Figure 2.1B, the so-called *repressed-repressor*. As this pattern is composed of two negative interactions, both $f_1(M)$ and $f_2(TF)$ are represented by inhibitory Hill functions in Equation (2.13), yielding Equation (2.14) (Mannan *et al.*, 2017).

$$P = b_2 + \frac{a_2}{1 + (K_2 \cdot (b_1 + \frac{a_1}{1+(K_1 \cdot M)^{n_1}}))^{n_2}} \tag{2.14}$$

Equation (2.14) gives rise to a sigmoidal dose-response curve (Figure 2.7A); the same is true for the other patterns in Figure 2.1B. Next, Equation (2.14) was transformed to yield expressions for the dose-response parameters (i.e. b, a, θ and μ, see Section 2.2.2) as a function of the tunable parameters (i.e. b_1, a_1, K_1, n_1, b_2, a_2, K_2 and n_2). The Mannan model predicts a fundamental constraint in biosensor design: for increasing transcription factor-TFBS affinity (i.e. increasing K_2), both μ and θ increased. In other words, a broader dynamic range can be obtained by tuning of the TFBS, but at the same time, the concentration of metabolite required to elicit that response will also grow (Figure 2.7B). These predictions could be reproduced *in vivo*, by mutating the TFBS of the *E. coli lac* repressor (LacI, a repressed-repressor). Simulations indicated that, although the relationship between μ and θ is monotonically increasing for $b_1 = 0$, for increasing values of b_1, the dynamic range reaches a maximum value that cannot be exceeded by tuning K_2. Nevertheless, the affinity of an MRTF for its TFBS is one of the most commonly employed tunable parameters in biosensor design. Similar constraints between μ and θ were observed for the other patterns in Figure 2.1B (Mannan *et al.*, 2017).

Interestingly, the Mannan model also predicts two tunable parameters that affect μ and θ orthogonally: the promoter dynamic range (i.e. $\mu_2 = \frac{a_2}{b_2}$) and transcription factor-metabolite affinity (K_1), respectively (Figure 2.7C). This orthogonal control could be validated *in vivo* for *E. coli* LacI. μ_2 was increased by replacing the -35 and -10 regions of the *lacUV5* promoter with the corresponding regions of the stronger P_{A1} promoter of phage T7, while K_1 was tuned by switching to a lower affinity effector (methyl-1-

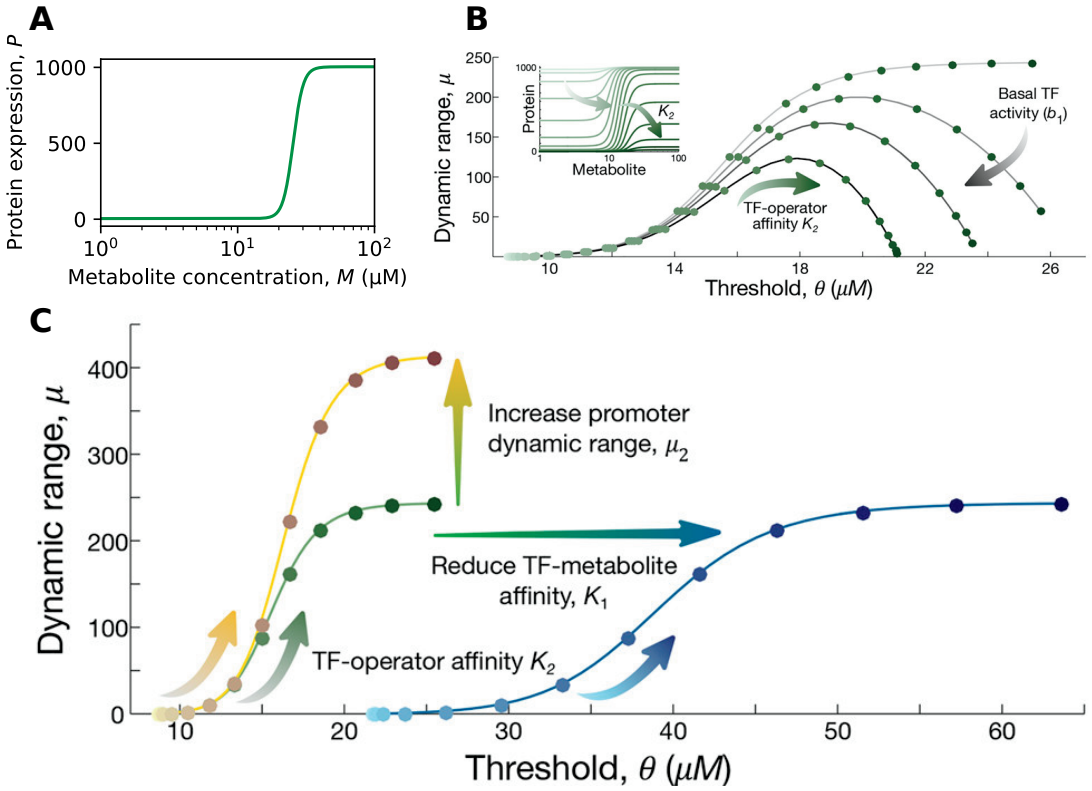

Figure 2.7: The Mannan model for biosensor response. **(A)** A dose-response curve obtained by simulation of Equation (2.14) with parameters $b_1 = 0.01$, $a_1 = 300$, $K_1 = 0.1$, $n_1 = 6$, $b_2 = 4.1$, $a_2 = 1000$, $K_2 = 0.9$ and $n_2 = 2$. **(B)** Simulations of the fundamental constraint in biosensor design for for varying values of b_1. The inset shows dose-response curves for varying K_2 at a high basal transcription factor activity. **(C)** Dynamic range and threshold can be controlled orthogonally by promoter dynamic range and transcription factor-metabolite affinity. Panels B-C were adapted from Mannan *et al.* (2017).

thio-β-D-galactopyranoside (TMG) instead of IPTG). However, good fits could not be obtained by solely refitting μ_2 or K_1. For the first experiment, a_2, b_2, and also n_2 were refitted; for the second experiment, K_1 had to be refitted in combination with b_1. These couplings cannot be fully explained by the model. Moreover, the K_1 value for TMG was 100-fold lower than the fitted value for IPTG, whereas earlier competition binding assays demonstrated that the affinity for TMG is approximately 10-fold lower (Gilbert & Müller-Hill, 1966). According to the model, μ_2 and K_1 should also provide orthogonal control of dynamic range and threshold in the other three biosensor patterns in Figure 2.1B (Mannan *et al.*, 2017).

Importantly, the authors claim that their phenomenological model is broadly applicable since it is constructed without any specific mechanistic assumptions. Hence, the reported design principles should also be of wide applicability. The authors finally suggest that a large portion of the biosensor design space can be accessed by tuning the promoter strength in combination with mutated TFBSs. For MRTFs with large basal activities (”leaky MRTFs”), perturbing the affinity for its cognate metabolite by protein engineering would be more appropriate (Mannan *et al.*, 2017).

2.3 β-Alanine-responsive transcription factors

BA (3-aminopropanoic acid) is the only β-type amino acid found in nature. Although nonproteinogenic, it plays an important role in metabolism as a direct precursor for pantothenic acid (vitamin B5) which is, in turn, a component of CoA and the acyl-carrier protein (Maas, 1952; Alberts & Vagelos, 1966). *Escherichia coli* mainly synthesises BA through the α-decarboxylation of L-aspartic acid catalysed by PanD (Cronan Jr, 1980). Alternatively, the degradation of uracil and possibly the hydrolysis of carnosine supply BA to the cell (Slotnick & Weinfeld, 1957; Campbell, 1960; Klein *et al.*, 1986). BA can also be actively taken up by the action of CycA, an H^+ symporter (Schneider *et al.*, 2004). The degradation of spermine and spermidine could also constitute a source of BA (Razin *et al.*, 1958). The degradation of BA has been characterised in *Pseudomonas fluorescens*. It proceeds in two enzymatic steps: in the first step, a transaminase catalyses the transamination reaction of BA with pyruvic acid yielding malonate semialdehyde and L-alanine. Malonate semialdehyde is a toxic intermediate that is oxidatively decar-

boxylated to acetyl-CoA and CO_2 by the action of the acylating malonate-semialdehyde dehydrogenase (Yamada & Jakoby, 1960; Hayaishi *et al.*, 1961; Dalwadi & King, 2020). Notably, a malonate-semialdehyde dehydrogenase purified from *Pseudomonas aeruginosa* was shown to form malonic acid instead of acetyl-CoA (Nakamura & Bernheim, 1961).

Given its importance as a central metabolite, it should be of no surprise that several MRTFs bind BA and regulate its metabolism in response. The first BA-responsive regulator to be described, BarR, was discovered in the *Thermoproteota Sulfolobus acidocaldarius* (Sa-BarR) and *Sulfurisphaera tokodaii* (St-BarR) (Liu *et al.*, 2014). BarR is a member of the FFRP family, one of the largest transcription factor families in *Archaea* (Pérez-Rueda & Janga, 2010). This transcription factor family was named after the role of *E. coli* leucine-responsive regulatory protein (Lrp), which is the archetype of the family, to coordinate metabolism under feast or famine conditions (that is, high or low nutritional availability) (Anderson *et al.*, 1976; Calvo & Matthews, 1994). The Lrp/AsnC family of transcriptional regulators is another frequently used designation in literature. FFRPs are small proteins (for instance, Sa-BarR only consists of 152 amino acid residues) that consist of two domains: an N-terminal DBD and a C-terminal *regulation of amino acid metabolism* (RAM) domain separated by a linker region (Platko & Calvo, 1993; Ettema *et al.*, 2002). The latter typically encompasses the effector binding site(s) and enables oligomerisation. Besides, FFRPs typically contain short unstructured N- and C-terminal tails that are essential in *E. coli* Lrp for higher-order oligomerisation and DNA binding (Chen *et al.*, 2001; de los Rios & Perona, 2007; Ziegler & Freddolino, 2021).

Purified Sa-BarR and St-BarR form octamers *in vitro* in the presence and absence of BA, notwithstanding that a small fraction of St-BarR eluted as lower-order complexes during gel filtration experiments in the presence of 5 mM BA. Both proteins are closely related to *S. tokodaii* glutamine receptor protein (Grp, 69% sequence identity), an FFRP family member for which crystal structures have been elucidated. In these structures, Grp makes up a tetramer of dimers arranged in a closed ring, the RAM domains interacting centrally and the DBDs projecting outward. In this common quaternary structure across FFRPs, the DNA is assumed to wrap around the contour of the assembly. Its effector glutamine binds to two different binding sites: one situated at the RAM domain between

two adjacent dimers (the canonical Type I binding site of FFRPs) and a second one near the DBD (Type IV binding site) (Kumarevel *et al.*, 2008; Shrivastava *et al.*, 2009). These crystal structures were used to build a structural homology model of St-BarR. BA could be modelled in a cleft corresponding to the former binding site of Grp, pointing out several residues potentially involved in BA binding in BarR (Liu *et al.*, 2014).

The *barR* gene is situated in a divergent orientation with respect to a putative BA—pyruvate aminotransferase gene (*SACI_RS10335* and *STK_RS06165* in *S. acidocaldarius* and *S. tokodaii*, respectively). Sa-BarR binds to a semi-palindromic binding motif found in the *barR/SACI_RS10335* intergenic region and coding sequence of *SACI_RS10335*. It was shown to bind to this region *in vivo* both in the presence and absence of exogenous BA. In contrast to most FFRPs that function as autorepressors, Sa-BarR activates transcription of its own structural gene in a BA-independent fashion (Kölling & Lother, 1985; Lin *et al.*, 1992; Cordone *et al.*, 2005; McFarland & Dorman, 2008). In addition, Sa-BarR activates *SACI_RS10335* expression in response to BA. It was hypothesised that this regulatory mechanism induces BA degradation under excess BA in order to recover energy and building blocks under the form of acetyl-CoA. Finally, it also functions as an activator of the *SACI_RS11230/SACI_RS11235* operon encoding glutamate synthase in the presence of BA. The upstream region of the glutamate synthase gene is a common binding target of FFRPs, and BarR could function here to connect BA and α-amino acid metabolism in *S. acidocaldarius* (Ernsting *et al.*, 1992; Nguyen-Duc *et al.*, 2013; Song *et al.*, 2013; Vassart *et al.*, 2013; Lu *et al.*, 2019). Our lab is currently characterising a novel BA-responsive FFRP, AHOS_RS02205, a homologue of BarR and Grp from the *Thermoproteota* member "*Acidianus hospitalis*".

On the other hand, OapR is a BA-responsive transcription factor that belongs to the GntR family (briefly introduced in Section 2.1.1.1). The GntR family is a large family of transcription factors that contain a conserved N-terminal DBD and a C-terminal oligomerisation and/or effector-binding domain connected by a peptide linker of variable length (Haydon & Guest, 1991). Family members can be classified into seven subfamilies according to the architecture of the C-terminal domain. OapR belongs to the MocR subfamily, which is widespread across *Bacteria* but absent in *Archaea* (Suvorova & Rodionov, 2016). It was named after *Ensifer meliloti* MocR, a transcription factor putatively involved in the regulation of rhizopine catabolism (Rossbach *et al.*,

1994). MocR subfamily transcription factors possess a large C-terminal domain homologous to the *class-I pyridoxal 5'-phosphate* (PLP)-*dependent aminotransferase* family, an ancient enzyme family represented by aspartate aminotransferase (AAT) (Mehta *et al.*, 1993). They typically regulate metabolic pathways involving PLP and amino acids, but their regulatory targets can be as divergent as peptidoglycan biosynthesis, virulence genes in phytopathogenic bacteria and membrane proteins (Belitsky & Sonenshein, 2002; Savidor *et al.*, 2014; Takenaka *et al.*, 2015; Milano *et al.*, 2016).

Very few MocR members have been structurally or functionally characterised. For instance, *B. subtilis* GabR is the only MocR transcription factor for which three-dimensional structures have been determined. These structures reveal a homodimeric assembly with the two monomers swapping domains in a head-to-tail manner. The C-terminal domains interface centrally, forming two binding pockets for PLP. GabR binds to PLP covalently through its active site lysine residue (K312). Upon addition of its effector γ-aminobutyric acid (GABA), the active site PLP dislodges from this lysine and instead forms a Schiff base with GABA through a transaldimination reaction. Hence, GabR can only bind GABA in the presence of cofactor PLP. Notwithstanding their striking similarity to class-I aminotransferases, no MocR transcription factor has been found to catalyse a chemical reaction to date (Edayathumangalam *et al.*, 2013; Okuda *et al.*, 2015).

oapR is transcribed in divergent orientation of the *oapTD* operon encoding putative BA—pyruvate transaminase and malonate-semialdehyde dehydrogenase. In *C. necator* H16, it activates transcription of *oapTD* in response to BA (Hanko *et al.*, 2020). Its TFBSs have been predicted through comparative genomics analysis and are present as two direct repeats (D1 and D2) and one inverted repeat (I3) in the *oapR/oapTD* intergenic region (Suvorova & Rodionov, 2016). This genetic arrangement closely resembles that of GabR, although the precise topology is not equivalent. A mechanism for GabR-mediated regulation has been proposed based on experimental findings in *B. subtilis*. In the absence of GABA, GabR binds to D1 and I3 with its inversely oriented DBDs and to the sequence in between through electrostatic interactions with a positively charged groove along its AAT-like domain (Edayathumangalam *et al.*, 2013; Amidani *et al.*, 2017). This arrangement occludes the -35 core promoter sequence for RNAP, and hence, prevents initiation of transcription. Binding of GABA would then induce a conformational change, displacing the DBD bound from I3 to D2, from where GabR can

recruit RNAP and activate transcription. At the same time, GabR represses transcription from its structural gene irrespective of effector binding (Belitsky & Sonenshein, 2002; Nardella *et al.*, 2020).

Escherichia coli does not encode any homologs to OapR and expression from a heterologous *oapTD* promoter was not mobilised by endogenous transcription factors in response to BA. Thus, OapR allows orthogonal gene expression control in *E coli*, making it an interesting MRTF for use as BA-responsive biosensor in *E coli*. Such biosensors have been constructed by the Minton group and our group, which revealed a dual function in *E coli* in response to BA. In the absence of BA, increasing the expression level of *oapR* results in a decrease in fluorescent reporter gene expression from the *oapTD* promoter, indicating that OapR represses transcription in the absence of BA. Upon BA supplementation, OapR functions as an activator, as the fluorescence measured in the presence of OapR and BA exceeds the fluorescence in the absence of OapR (Hanko *et al.*, 2020; Magnus, 2021).

Chapter 3

Aims and objectives

Whole-cell biosensors are emerging synthetic biology tools that allow to differentially regulate gene expression in response to changes in physicochemical parameters experienced by the cell, such as metabolite concentration. This makes biosensors, and particularly biosensors based on metabolite-responsive transcription factors (MRTFs), excellently suited for metabolic engineering applications. For instance, dynamic pathway regulation has developed a true hype among metabolic engineers. Dynamic pathway regulation allows to dynamically and autonomously control metabolic flux by regulating key pathway enzymes in response to changes in intracellular metabolite concentration, and its implementation has led to increased production titers, rates and yields in a range of applications.

In order to be widely applicable, biosensors need to be tunable. However, a lack of quantitative understanding of the effect of different interventions on its response curve resulted in a high reliance on trial-and-error methodologies during biosensor development. In this context, a simple phenomenological model has been published, which revealed fundamental constraints for the design of MRTF-based biosensors and exposed tunable parameters that allow orthogonal control of the dynamic range and threshold. Because the model was formulated without imposing mechanistic assumptions, it was claimed that the model should be broadly applicable.

The first part of this thesis aims to obtain further insight in the effect of perturbations to

MRTF-based biosensors on their response curve. For this purpose, model simulations were compared to published results of real-life biosensors. In addition, a parameter estimation method to fit biosensor characterisation data to the model was implemented in Python and evaluated using simulated data sets.

The second part of this thesis aims to engineer and characterise *Escherichia coli* biosensors responsive to β-alanine, an intermediate compound for the production of industrially important platform molecules such as 3-hydroxypropionic acid. Mutations were introduced in the predicted transcription factor binding sites of *Cupriavidus necator* OapR through site-directed mutagenesis and their effect on the dose-response characteristics was characterised *in vivo*. In addition, experimental dose-response curves of biosensors based on AHOS_RS02205 from "*Acidianus hospitalis*" were analysed and fitted to a Hill function.

Together, these endeavours fit in with the goal to improve the production of 3-hydroxypropionic acid by implementing biosensor-mediated dynamic control at the β-alanine node in *Escherichia coli*.

Chapter 4

Materials and methods

4.1 Mathematical modelling

Simulations and fitting were performed in Python 3.9.12 (Van Rossum & Drake, 2009) along with the packages `NumPy 1.21.5`, `Matplotlib 3.5.1`, `pandas 1.4.3`, `seaborn 0.11.2`, `lmfit 1.0.3`, `SciPy 1.7.3` and `SALib 1.4.5`.

Throughout this thesis the common logarithm $\log_{10}(x)$ is denoted $\lg(x)$ and the natural logarithm $\log_e(x)$ is denoted $\ln(x)$ for brevity.

4.1.1 Dose-response parameter estimation

Experimental dose-response curves obtained through the procedure described in Section 4.2.2.2 were fitted to a Hill function as described before (Landry *et al.*, 2018). Briefly, a data set comprising measurements of the normalised fluorescence (FL/OD_{600}) across biological replicates was fit to Equation (2.4) using the `fit` function of the `lmfit Model` class. Data points corresponding to a supplementation with 100 mM β-alanine (BA) were omitted for fitting, as their concomitant decreased FL/OD_{600} were not expected to be directly related to the function of the biosensor. Default options were used, except for the `weights` option that was used to multiply each residual with the inverse of the mean FL/OD_{600} at that data point, which improves the fitting at low BA concentrations. 95% confidence intervals on b, a, θ and n were calculated using

the `conf_interval` function, which executes the F-test (Newville *et al.*, 2016).

4.1.2 Development of a parameter estimation method for the Mannan model for biosensor response

4.1.2.1 The parameter estimation method

The parameter estimation method (PEM) implemented in MATLAB by Mannan *et al.* (2017) was adapted to run on Python. This method departs from eight transcription-factor binding site (TFBS) mutants that hence exhibit a different transcription factor-TFBS affinity (K_2). Briefly, for each strain i, all biological replicates were fitted separately to a Hill function (Equation 2.4), yielding $\mu_{i,E}$, $\theta_{i,E}$ and $b_{i,E}$ where the subscript E denotes that the parameters were experimentally determined. These dose-response parameters were globally fitted to the Mannan model by forcing all model parameters except K_2 (that is, $p = \{b_1, a_1, K_1, n_1, b_2, a_2, n_2\}$) to be equal for all strains. For each strain i individually, all model parameters (including K_2 were fixed across biological replicates. Concretely, dose-response parameters were fitted to the Mannan model by minimising the objective function

$$C(K_{2,1}, K_{2,2}, ..., K_{2,n}, p) = \sum_{i=1}^{n} \left(\frac{\mu_{i,E} - \mu_{i,M}(K_{2,i}, p)}{\hat{\mu}_E} \right)^2 + \left(\frac{\theta_{i,E} - \theta_{i,M}(K_{2,i}, p)}{\hat{\theta}_E} \right)^2 + \left(\frac{\lg(b_{i,E}) - \lg(b_{i,M}(K_{2,i}, p))}{\lg(\hat{b}_E)} \right)^2 \tag{4.1}$$

where n denotes the number of strains; $\mu_{i,M}$, $\theta_{i,M}$ and $b_{i,M}$ designate the i^{th} model predictions of the dose-response parameters; $\hat{\mu}_E$, $\hat{\theta}_E$ and $\hat{b}_E$ designate the maximum measured values of the dose-response parameters and $K_{2,i}$ is the fitted value of K_2 for strain i. The repressed-repressor pattern was chosen as case study, for which the model predictions can be computed as

$$\mu = \mu_2 \frac{((b_1 + a_1)^{n_2} - b_1^{n_2})K_2^{n_2}}{(1 + (K_2(b_1 + a_1))^{n_2} + \mu_2)(1 + (b_1 K_2)^{n_2})} \tag{4.2}$$

$$\theta = \frac{1}{K_1} \sqrt[n_1]{\frac{a_1 K_2}{\sqrt[n_2]{A - 1} - b_1 K_2} - 1} \tag{4.3}$$

$$b = b_2 + \frac{a_2}{1 + (K_2(b_1 + a_1))^{n_2}} \tag{4.4}$$

Note that the M subscript was omitted for brevity in Equations (4.2) to (4.4). In Equation (4.3), A is a function of tunable parameters as:

$$A = 2 \cdot \frac{(1 + (K_2(b_1 + a_1))^{n_2}) \cdot (1 + (b_1 K_2)^{n_2})}{2 + ((b_1 + a_1)^{n_2} + b_1^{n_2})K_2^{n_2}} \tag{4.5}$$

The optimisation problem in Equation (4.1) can be solved *via* nonlinear least-squares. The choice of *covariates*, including the optimisation algorithm, imposed bounds and scaling, impacts the obtained results (Kreutz, 2019). Mannan *et al.* (2017) used *multistart optimisation* to minimise $C(K_{2,1}, K_{2,2}, ..., K_{2,n}, p)$ but not all covariates were clearly defined. Therefore, a strategy based on Dray *et al.* (2022) is proposed below.

First, bounds should be imposed on the model parameters in order to prevent mathematical errors (such as division by zero), biologically irrelevant results and to improve convergence. Bounds were applied as in Mannan *et al.* (2017), except for n_1 (the sensitivity of metabolite-transcription factor binding) and K_2 that were further constrained and n_2 (the sensitivity of transcription factor-TFBS binding) that was extended (Table 4.1).

Table 4.1: Bounds applied for parameter estimation for the Mannan model for biosensor response.

Parameter	Lower bound	Upper bound
b_1	0	0.1
a_1	0	500
K_1	$1 \cdot 10^{-4}$	1

Continued on next page

Table 4.1: Bounds applied for parameter estimation for the Mannan model for biosensor response. (Continued)

Parameter	Lower bound	Upper bound
n_1	1.0	4.0
b_2	0	20
a_2	0	$3 \cdot 10^3$
K_2	$5 \cdot 10^{-3}$	1
n_2	0.5	3.0

A PEM should be able to identify the global optimum of the objective function in the presence of multiple local minima. Often, *deterministic optimisation algorithms* are applied for fitting. Such methods start from an initial guess and seek to decrease the objective function over multiple iterations using derivative information. However, these algorithms can get stuck in local minima depending on the choice of initial guess. This limitation can be overcome by initiating multiple independent optimisation runs from a series of random initial guesses. If implemented correctly, a portion of the runs end in local minima while another portion ends in the global optimum. *Latin hypercube sampling* (LHS) was used to generate the initial guesses (Figure 4.1). For illustration, drawing N samples from a 2-dimensional parameter space through LHS can be thought of as dividing the parameter space in N^2 boxes (Figure 4.1B). For the first sample, a box in the first column of the parameter space is selected; the sample is subsequently drawn from a uniform distribution within this box. This rationale is repeated for the second sample, where any box from the second column can be chosen that is situated in a row that has not yet been sampled in the previous step. Therefore, LHS allows for better sampling of the parameter space, where purely uniformLy drawn samples can accidentally lie close to each other (Figure 4.1C-D) (McKay *et al.*, 1979).

The LHS implementation of the `SALib` package was applied to generate initial guesses within the parameter bounds (Herman & Usher, 2017). K_1, n_1, K_2 and n_2 were sampled over a logarithmic scale by sampling the exponent space and subsequently transforming the sample m as 10^m. All other parameters were directly sampled on a linear scale. Further, eight samples were drawn for K_2 in total, one for each strain. This way, 500 pa-

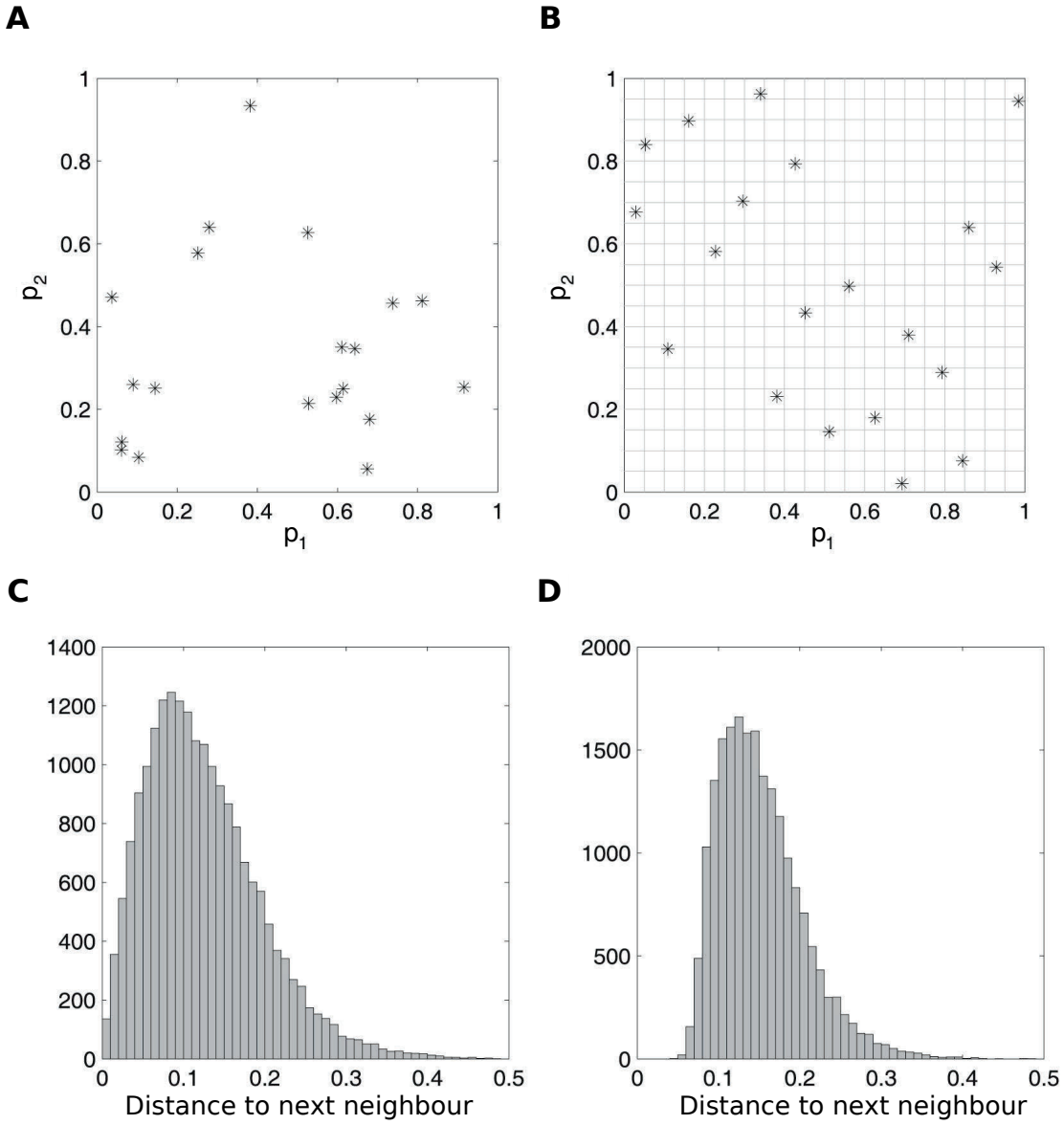

Figure 4.1: Comparison of random sampling of initial parameter guesses from a uniform distribution and using Latin hypercube sampling (LHS). **(A-B)** Generation of 20 initial guesses in a two-dimensional parameter space (with hypothetical parameters p_1 and p_2) from a uniform distribution **(A)** and using LHS **(B)**. **(C-D)** Histogram showing the distribution of the Euclidean distances to the nearest neighbour in parameter space for 1000 repetitions of the sampling of 20 initial guesses in a two-dimensional parameter space from a uniform distribution **(C)** and using LHS **(D)**. Adapted from Raue *et al.* (2013).

rameter sets were generated and subsequently optimised by minimising Equation (4.1) using the *Levenberg-Marquardt* algorithm implemented using the `minimize` function from the `lmfit` package (Levenberg, 1944; Marquardt, 1963; Moré, 1978; Newville *et al.*, 2016). Default parameters were applied, except for the option `nan_policy` that was set to 'propagate'. Levenberg-Marquardt is a deterministic optimisation algorithm that interpolates between the *Gauss-Newton* and *steepest descent* methods, and it is a reasonable first choice for nonlinear least-squares minimisation problems (Ashyraliyev *et al.*, 2009; Transtrum & Qiu, 2012; Dray *et al.*, 2022). The parameter set yielding the lowest objective value is then reported as the best fit.

4.1.2.2 Evaluation of the parameter estimation method

The proper implementation of the PEM set out in Section 4.1.2.1 was evaluated according to Dray *et al.* (2022). This method departs from simulated data for which the model parameters are known in advance. It is then evaluated whether the PEM succeeds in finding global optima and whether the parameter values corresponding to the optimal solution are similar to the true values.

As we were not yet able to construct sufficient amounts of TFBS mutants for OapR and AHOS_RS02205 to fit the Mannan model adequately, we departed from LacI (a repressed-repressor, see Figure 2.1B) data from literature ('LacExp1' from Mannan *et al.* (2017)). This data includes experimental values of μ, θ and b for a varying number of biological replicates across eight TFBS mutants. For each strain i, the standard deviations of the dose-response parameters were calculated (s_{μ_i} and s_{θ_i} and s_{b_i}). LHS was then performed as described in Section 4.1.2.1 to generate 10^4 parameter sets and their objective function value was calculated with respect to the LacI data using Equations (4.1) to (4.5). The nine parameter sets yielding the lowest objective function values (i.e. best describing the LacI data) were considered as true values of nine simulated data sets. For each data set, the true dose-response parameters ($\mu_{i,True}$, $\theta_{i,True}$, $b_{i,True}$) were computed for each strain i from Equations (4.2) to (4.4) and simulated (μ, θ, b) triplets were generated by drawing the same number of replicates for a given strain as reported by Mannan *et al.* (2017) from ($\mathcal{N}(\mu_{i,True}, s^2_{\mu_i})$, $\mathcal{N}(\theta_{i,True}, s^2_{\theta_i})$, $\mathcal{N}(b_{i,True}, s^2_{b_i})$). If negative or zero values were generated in the process, they were arbitrarily set to 1.0 to exclude biological and mathematical infeasibilities.

The procedure described above yields nine simulated data sets akin to experimental data for which the true model parameters are known. The data sets show a variety of plausible behaviours with various objective function properties that might challenge the PEM in different ways. Each data set was individually fitted to the Mannan model according to the strategy expounded in Section 4.1.2.1. Note that the same initial guesses generated through LHS were used for all data sets in order to save on computational power. The obtained objective function values were then compared to the objective computed for the data set with and without noise (C_{pass}) and the best fit was compared to the ideal noise-free solution.

4.2 Molecular biology work

4.2.1 Materials

4.2.1.1 Buffers, media and solutions

This section lists the buffers, media and solutions used throughout this thesis project. Chemicals used for their preparation were obtained from Sigma Aldrich (St. Louis, MO, USA), unless indicated otherwise (Table 4.2).

Table 4.2: Composition of the different media, buffers and solutions used throughout this thesis project.

Medium/buffer/solution[a]	Composition[b]
Media	
Solid LB medium	25.0 g/L LB broth base (Labconsult, Brussels, Belgium), 15.0 g/L micro agar (Labconsult).
Liquid LB medium	25.0 g/L LB broth base (Labconsult).
InBio basis	83.3 mM NH_4Cl, 47.3 mM $(NH_4)_2SO_4$, 27.5 mM KH_2PO_4, 52.5 mM K_2HPO_4, 50.0 mM MOPS, 17.9 mM NaCl. pH was adjusted to 7.0 with NaOH/HCl. Autoclaved.

Continued on next page

Table 4.2: Composition of the different media, buffers and solutions used throughout this thesis project. (Continued)

Medium/buffer/solution[a]	Composition[b]
InBio glucose & $MgSO_4$ solution	458 mM glucose, 10.1 mM $MgSO_4$. Autoclaved.
InBio vitamin & trace element solution	23.7 mM $FeCl_2$, 50.0 mM $CaCl_2$; 11.3 mM $MnCl_2$, 3.10 mM $CuCl_2$, 3.00 mM $CoCl_2$, 10.5 mM $ZnCl_2$, 0.400 mM H_3BO_4, 1.10 mM disodium EDTA, 3.60 mM thiamine.HCl. Filter-sterilised.
InBio molybdate solution	4.00 mM Na_2MoO_4. Filter-sterilised.
InBio minimal medium	4 volumes of InBio basis + 1 volume InBio glucose & $MgSO_4$ solution + 1/1000 InBio vitamin & trace element solution + 1/10000 InBio molybdate solution.
10X MOPS Buffer (Teknova, Hollister, CA, USA)	400 mM MOPS, 40.0 mM tricine, 0.100 mM $FeSO_4$, 95.0 mM NH_4Cl, 2.76 mM K_2SO_4, 5.00 µM $CaCl_2$, 5.25 mM $MgCl_2$, 500 mM NaCl, $2.92 \cdot 10^{-6}$ mM $(NH_4)_2MoO_4$, $4.00 \cdot 10^{-4}$ mM H_3BO_3, $3.02 \cdot 10^{-5}$ mM $CoCl_2$, $9.62 \cdot 10^{-6}$ mM $CuSO_4$, $8.08 \cdot 10^{-5}$ mM $MnCl_2$, $9.74 \cdot 10^{-6}$ mM $ZnSO_4$.
Potassium phosphate dibasic solution (Teknova)	132 mM K_2HPO_4.
10X ACGU Solution (Teknova)	1.99 mM adenine, 1.99 mM cytosine, 1.99 mM guanine, 1.99 mM uracil
MOPS EZ Rich Defined Medium[c]	For the preparation of 1000 mL of medium: 100 mL of 10X MOPS Buffer, 10 mL potassium phophate dibasic solution, 100 mL 10X ACGU Solution, 10 mL of 20% glucose (Teknova), 780 mL sterile H_2O (Teknova). Filter-sterilised.
Buffers/solutions	

Continued on next page

Table 4.2: Composition of the different media, buffers and solutions used throughout this thesis project. (Continued)

Medium/buffer/solution[a]	Composition[b]
1000X chloramphenicol solution	25 mg/mL chloramphenicol in ethanol.
1000X kanamycin solution	60 mg/mL kanamycin.
5X TBE buffer	446 mM Tris, 445 mM H_3BO_3, 10 mM disodium EDTA.
6X DNA loading dye	25% Ficoll 400, 0.1% xylene cyanol, 0.1% bromophenol blue, 0.5% SDS, 0.1% disodium EDTA.
FastAP (Thermo Scientific, Waltham, MA, USA)	1 U/µL FastAP Thermosensitive Alkaline Phosphatase.
2X GoTaq® Green Master Mix (Promega, Madison, WI, USA)	2.5 U/µL GoTaq® DNA Polymerase, 400 µM dATP, 400 µM dGTP, 400 µM dCTP, 400 µM dTTP, 3.00 mM $MgCl_2$ in Green GoTaq® Reaction Buffer (pH 8.5).
2X KAPA HiFi DNA Polymerase master mix (Roche, Basel, Switzerland)	1 U/µL KAPA HiFi DNA Polymerase, 600 µM dATP, 600 µM dGTP, 600 µM dCTP, 600 µM dTTP, 2.00 mM $MgCl_2$ in KAPA HiFi Fidelity Buffer.
Naringenin solution[d]	Naringenin in ethanol.

[a] **Media abbreviations:** LB, lysogeny broth.

[b] **Chemical abbreviations:** MOPS, 3-(morpholin-4-ium-4-yl)-1-propanesulfonate; EDTA, ethylenediaminetetraacetic acid; SDS, sodium dodecyl sulphate; dATP, deoxyadenosine triphosphate; dGTP, deoxyguanosine triphosphate; dCTP, deoxycytidine triphosphate; dTTP, deoxythymidine triphosphate.

[c] The manufacturer suggests an additional supplement ('5X EZ Supplement') containing the twenty canonical α-amino acids and calcium pantothenate. This was omitted as it could influence the assay of interest in our work.

[d] Concentration varied depending on the experiment.

4.2.1.2 Kits

Several commercially available kits were used to facilitate cloning (Table 4.3). All kits were obtained from Promega.

Table 4.3: Overview of the kits used throughout this thesis project.

Kit	Use[a]
PureYield™ Plasmid Miniprep System	Plasmid DNA purification.
PureYield™ Plasmid Midiprep System	Plasmid DNA purification.
Wizard® SV Gel and PCR Clean-Up System	Purification of restricted DNA fragments from agarose gels and PCR products.

[a] **Abbreviations:** PCR, polymerase chain reaction.

4.2.1.3 Bacterial strains

This section describes the different *Escherichia coli* strains used throughout this thesis project (Table 4.4).

Table 4.4: Overview of *E. coli* strains used troughout this thesis project.

Strain	Genotype	Use	Supplier
DH5α	F^- λ^- Φ80*lacZ*ΔM15 Δ(*lacZYA-argF*)U169 *recA1 endA1 hsdR17*(r_K^-, m_K^+) *phoA glnV44 thi-1 gyrA96 relA1*	Cloning	Invitrogen (Carlsbad, CA, USA)
DH10B	F^- λ^- Φ80*lacZ*ΔM15 *mcrA* Δ(*mrr-hsdRMS-mcrBC*) Δ*lacX74 recA1 endA1 araD139* Δ(*ara-leu*)7697 *galU galK rpsL*(Str^R) *nupG*	Cloning	Thermo Scientific

Continued on next page

Table 4.4: Overview of *E. coli* strains used troughout this thesis project. (Continued)

Strain	Genotype	Use	Supplier
MG1655	F^- λ^- *rph-1 rfb-50 ilvG*$^-$	*In vivo* characterisation of biosensors	Netherlands Culture Collection of Bacteria (Utrecht, The Netherlands)

4.2.1.4 Plasmid vectors

This section describes the different plasmid vectors used throughout this thesis project for the construction of whole-cell biosensors (Table 4.5).

Table 4.5: Overview of plasmid vectors used throughout this thesis project.

Plasmid	Features[a]	Source
Inducible transcription factor constructs (pITC)		
pITC	pSC101 *ori*, chloramphenicol resistance gene and naringenin-inducible system, consisting of the *Herbaspirillum seropedicae fdeR* transcription factor gene and corresponding *fdeAR* promoter (De Paepe *et al.*, 2018).	Drs. ir. Amber Bernauw & Dr. ir. Indra Bervoets (MICR VUB)
pITC-AHOS_RS02205	pITC containing the *AHOS_RS02205* gene from "*Acidianus hospitalis*" W1 (codon-optimised for *E. coli*) under control of *H. seropedicae* P_{fdeAR}	Drs. ir. Amber Bernauw (MICR VUB)
pITC-OapR	pITC containing the *oapR* gene from *Cupriavidus necator* H16 (codon-optimised for *E. coli*) under control of *H. seropedicae* P_{fdeAR}	Drs. ir. Amber Bernauw (MICR VUB)
Promoter reporter constructs (pPRC)		

Continued on next page

Table 4.5: Overview of plasmid vectors used throughout this thesis project. (Continued)

Plasmid	Features[a]	Source
pPRC	pBR322 *ori*, Tn*5* (containing a kanamycin resistance gene) and a promoterless *mkate2* reporter gene.	Drs. ir. Amber Bernauw & Dr. ir. Indra Bervoets (MICR VUB)
pPRC-1S	pPRC containing the "*A. hospitalis*" W1 *AHOS_RS02205* promoter upstream of the *mkate2* reporter gene.	Drs. ir. Amber Bernauw (MICR VUB)
pPRC-2S	pPRC containing the "*A. hospitalis*" W1 *AHOS_RS02210* promoter upstream of the *mkate2* reporter gene.	Drs. ir. Amber Bernauw (MICR VUB)
pPRC-49S	pPRC containing the *C. necator* H16 *oapTD* promoter upstream of the *mkate2* reporter gene.	Drs. ir. Amber Bernauw (MICR VUB)
pPRC-mut1.1	pPRC-49S containing the 8C>G substitution in the D1 site.	Constructed during this project.
pPRC-mut1.2	pPRC-49S containing the 7A>T substitution in the D1 site.	Constructed during this project.
pPRC-mut2	pPRC-49S containing the 8C>G substitution in the D2 site.	Constructed during this project.
pPRC-mut3.1	pPRC-49S containing the 10G>C substitution in the I3 site.	Constructed during this project.
pPRC-mut3.2	pPRC-49S containing the 8C>G substitution in the I3 site.	Constructed during this project.
pPRC-mut3.3	pPRC-49S containing the 7A>T substitution in the I3 site.	Constructed during this project.

Continued on next page

Table 4.5: Overview of plasmid vectors used throughout this thesis project. (Continued)

Plasmid	Features[a]	Source
pPRC-54S	pPRC containing the *Sulfolobus acidocaldarius barR* promoter upstream of the *mkate2* reporter gene.	Drs. ir. Amber Bernauw (MICR VUB)
pPRC-55S	pPRC containing the *S. acidocaldarius SACI_RS10335* promoter upstream of the *mkate2* reporter gene.	Drs. ir. Amber Bernauw (MICR VUB)
Single-plasmid biosensor constructs (pBS)		
pBS	pITC containing the promoterless *mkate2* reporter gene inserted in MCS2.	Drs. ir. Amber Bernauw & Dr. ir. Indra Bervoets (MICR VUB).
pBS-OapR+_	pITC-OapR containing the promoterless *mkate2* reporter gene.	Constructed during this project.
pBS-_+49S	pITC containing the *mkate2* reporter gene under control of P_{oapTD} from *C. necator* H16.	Constructed during this project.
pBS-OapR+49S	pITC-OapR containing the *mkate2* reporter gene under control of P_{oapTD} from *C. necator* H16.	Constructed during this project.

[a] **Abbreviations:** *ori*, origin of replication; MCS, multiple cloning site.

4.2.1.5 Oligonucleotide primers

This section describes the different oligonucleotide primers used throughout this thesis project (Table 4.6). Oligonucleotides were synthesised by Sigma Aldrich (St. Louis, MO, USA) or Integrated DNA Technologies (Leuven, Belgium). All primers were used at 10 µM stock concentration.

Table 4.6: Overview of oligonucleotide primers used throughout this thesis project.

Name	Sequence (5' → 3')	Purpose[a]
IB0560	1 GTCCAGATAG CCCAGTAGCT 21 GACATTC	Fw primer for colony PCR of pPRC and pBS constructs and sequencing primer for pPRC and pBS constructs.
IB0313	1 GTGGTTATTC ACGGTGCCTT 21 CC	Rv primer for colony PCR of pPRC constructs.
IB0264	1 CCTTCGTAAA TCTGGCGAGT 21 GG	Rv primer for colony PCR of pBS constructs.
VC001	1 GTGCAGTTTG AACACGCCAA 21 ACCCTACAGG CAAGCGTCAT 41 C	Fw primer for the construction of pPRC-mut1.1 through SDM.
VC002	1 GATGACGCTT GCCTGTAGGG 21 TTTGGCGTGT TCAAACTGCA 41 C	Rv primer for the construction of pPRC-mut1.1 through SDM.
VC003	1 GTGCAGTTTG AACACGCCAA 21 ACCGAACAGG CAAGCGTCAT 41 C	Fw primer for the construction of pPRC-mut1.2 through SDM.
VC004	1 GATGACGCTT GCCTGTTCGG 21 TTTGGCGTGT TCAAACTGCA 41 C	Rv primer for the construction of pPRC-mut1.2 through SDM.
VC011	1 GTCCGGACAA GACTGCCTTA 21 CCCTGCAGTT TGAACACG	Fw primer for the construction of pPRC-mut2 through SDM.
VC012	1 CGTGTTCAAA CTGCAGGGTA 21 AGGCAGTCTT GTCCGGAC	Rv primer for the construction of pPRC-mut2 through SDM.

Continued on next page

Table 4.6: Overview of oligonucleotide primers used throughout this thesis project. (Continued)

Name	Sequence (5' → 3')	Purpose[a]
VC019	1 GTCTGGCCGA CAAAATATAC 21 ACTCCGGACA AGACTGCCTT 41 AC	Fw primer for the construction of pPRC-mut3.1 through SDM.
VC020	1 GGTAAGGCAG TCTTGTCCGG 21 AGTGTATATT TTGTCGGCCA 41 GACG	Rv primer for the construction of pPRC-mut3.1 through SDM.
VC021	1 CGTCTGGCCG ACAAAATATA 21 GAGTCCGGAC AAGACTGCCT 41 TACC	Fw primer for the construction of pPRC-mut3.2 through SDM.
VC022	1 GGTAAGGCAG TCTTGTCCGG 21 ACTCTATATT TTGTCGGCCA 41 GACG	Rv primer for the construction of pPRC-mut3.2 through SDM.
VC023	1 CGTCTGGCCG ACAAAATATT 21 CAGTCCGGAC AAGACTGCCT 41 TACC	Fw primer for the construction of pPRC-mut3.3 through SDM.
VC024	1 GGTAAGGCAG TCTTGTCCGG 21 ACTGAATATT TTGTCGGCCA 41 GACG	Rv primer for the construction of pPRC-mut3.3 through SDM.
AB245	1 GACGATGGTA GTCAGCTGCC 21 TACTAAG	Sequencing of *mkate2* on pPRC-mut constructs.
VC031	1 TCTGATGCCA CTTGTCCACC 21 TCTC	Sequencing of *mkate2* on pPRC-mut constructs.
VC032	1 AGAGGTGGAC AAGTGGCATC	Sequencing of *ori* on pPRC-mut constructs.
VC033	1 AAAGGTTGGG CTTCGGAATC	Sequencing of *ori* on pPRC-mut constructs.

Continued on next page

Table 4.6: Overview of oligonucleotide primers used throughout this thesis project. (Continued)

Name	Sequence (5' → 3')	Purpose[a]
VC034	1 AGACTGGGCG GTTTTATGG	Sequencing of Tn*5* on pPRC-mut constructs.
IB0036	1 GCGGAGCCTA TGGAAAAACG	Sequencing of Tn*5* on pPRC-mut constructs.

[a] **Abbreviations:** Fw, forward; Rv, reverse; PCR, polymerase chain reaction; SDM, site-directed mutagenesis; *ori*, origin of replication.

4.2.1.6 Restriction enzymes

This section describes the different restriction enzymes used throughout this thesis project (Table 4.7). FastDigest restriction enzymes were used, obtained from Thermo Scientific.

Table 4.7: Overview of restriction enzymes used throughout this thesis project.

Restriction enzyme	Recognition site (5' → 3') and restriction sites (↓)	Purpose
BamHI	G↓G A T T C C C T A A↑G	Restriction of pITC and pPRC.
KpnI	G G T A C↓C C↑C A T G G	Restriction of pITC and pPRC.
DpnI	G A^{m6}↓T C C T ↑A^{m6}G	Digest parent DNA after site-directed mutagenesis.

4.2.2 Methods

4.2.2.1 Cloning

For all cloning experiments, *E. coli* DH5α or DH10B was used depending on the availability. Strains were grown at 37 °C, either on solid or in liquid LB medium supplemented with the appropriate antibiotics. DNA concentrations were determined on a NanoDrop spectrophotometer (Thermo Scientific).

Construction of pBS constructs. pBS plasmids were constructed using restriction enzyme digestion and ligation. For this purpose, pITC and pPRC constructs were digested with BamHI and KpnI. Reaction mixtures were prepared according to the manufacturer's instructions, except for the final volume that was doubled to 40 µL. The reaction mixtures were subsequently incubated for 30 min at 37 °C. Next, 1 µL of FastAP was added to the pITC reaction mixtures to prevent self-ligation. Both reaction mixtures were then incubated for 30 more minutes at 37 °C, after which the enzymes were inactivated through incubation at 70 °C (15 min). Restricted plasmids were resolved by agarose gel electrophoresis (0.8% UltraPure™ Agarose from Invitrogen, Carlsbad, CA, USA). The desired fragments were subsequently purified from gel and loaded on 0.8% agarose gels to estimate their concentrations.

Each pPRC fragment was subsequently ligated in the appropriate linearised pITC vector using T4 DNA Ligase (Thermo Scientific) according to the manufacturer's instructions with adaptations. Briefly, the insert DNA was added in threefold molar excess to the linearised vector. Subsequently, 2 µL of 10X T4 DNA Ligase buffer and 5 U of T4 DNA Ligase were added and nuclease-free H_2O was added to 20 µL. The reaction mixtures were incubated overnight at room temperature (RT), after which they were used to transform $CaCl_2$-competent *E. coli* DH10B. 100 µL aliquots of competent cells were thawed on ice for at least 10 min and, subsequently, 5 µL of reaction mixture was added to each aliquot. The cells were gently mixed and incubated on ice for 15 min. Next, a heat shock was applied by incubating the cells for 3 min at 37 °C. The cells were then incubated 10 min at RT and transferred to 1 mL of liquid LB medium. Cultures were incubated for 2 h at 37 °C in a shaking incubator and plated on solid LB medium supplemented with chloramphenicol to obtain colonies.

Colony polymerase chain reaction (PCR) was performed to confirm successful cloning. For this purpose, single colonies were picked with a sterile toothpick and transferred to both a replica plate and a microcentrifuge tube containing 30 µL of nuclease-free H_2O. The replica plate was then incubated overnight to obtain colonies and the microcentrifuge tubes were incubated for 10 min at 100 °C. They were subsequently allowed to cool down, after which the cell debris was spun down by centrifugation. The resulting supernatant was used as template for PCR using 2X GoTaq® Green Master Mix according to manufacturer's instructions. Each PCR tube contained 6 µL of nuclease-free

H_2O, 2 µL of both primers (IB0560 and IB0264), 12.5 µL of master mix and 1.5 µL of template. The applied PCR program consisted of (1) inital denaturation at 95 °C for 2 min, (2) 35 cycles of denaturation (30 s at 95 °C), annealing (30 s at 56 °C) and extension (1.5 min at 72 °C) and (3) final extension at 72 °C for 3 min. Next, the obtained PCR products were analysed by agarose gel electrophoresis (0.8% agarose). Plasmids for which the PCR products showed the desired migration were purified and sent for Sanger sequencing using IB0560 as sequencing primer (Eurofins Genomics, Ebersberg, Germany). Transformants containing the intended pBS constructs were stored in 50% glycerol at -80 °C.

Construction of pPRC-mut constructs. pPRC-mut plasmids were constructed from pPRC-49S by site-directed mutagenesis according to Edelheit *et al.* (2009) with adaptations. For each desired mutation, two primers were designed that encompass said mutation (VC001 up to VC024 in Table 4.6). These primers were used to amplify pPRC-49S in two separate PCR reactions. PCR was performed using the 2X KAPA HiFi DNA Polymerase master mix. Reaction mixtures contained 500 ng of pPRC-49D, 4 µL of primer (forward or reverse), 12.5 µL of master mix and nuclease-free H_2O added to a final volume of 25 µL. The applied PCR program consisted of (1) inital denaturation at 95 °C for 3 min, (2) 35 cycles of denaturation (20 s at 98 °C), annealing (15 s at 65 °C) and extension (3.5 min at 72 °C) and (3) final extension for 10 min at 72 °C. The two single-primer PCR reaction products introducing the same mutation were then combined and denatured at 95 °C to separate the newly synthesised DNA strands from the template. The mixtures were subsequently cooled to 37 °C to allow reannealing of the mutated strands. Finally, 3 µL of DpnI was added and the reaction mixtures were incubated overnight at 37 °C to digest the methylated, non-mutated template DNA (pPRC-49S).

4 µL of mutated plasmids was then transformed in $CaCl_2$-competent *E. coli*. Transformation by heat shock was performed as described before, with the minor difference that a heat shock of 40 s at 42 °C was applied instead of 3 min at 37 °C. Colony PCR was performed as described before except for the primers (IB0560 in combination with IB0313), number of cycles (35) and extension time (35 s and 2 min for the final extension). A band at 600 bp on an (1.0%) agarose gel indicates the presence of a pPRC-49S-derived plasmid. PCR products showing such bands were sent for sequencing using

IB0560 as sequencing primer. For each intended mutation, one colony corresponding to a PCR product containing the intended mutation was used to isolate the mutated plasmid, yielding the six different pPRC-mut constructs (see Table 4.5). For each mutated plasmid, all plasmid features were sequenced using primers AB245, VC031 up to VC034 and IB0036 (see Table 4.6); no secondary mutations were found.

4.2.2.2 *In vivo* characterisation of biosensors

E. coli MG1655 was used for all *in vivo* experiments. To assemble a whole-cell biosensor, 1 µL of the desired pITC and pPRC plasmid vectors were co-transformed in $CaCl_2$-competent *E. coli* MG1655 according to the heat shock method as described before.

For the actual *in vivo* biosensor characterisation, biosensor strains were cultured in InBio minimal medium or MOPS EZ Rich Defined Medium. For each strain, a preculture was first set up in 200 µL of the desired culture medium supplemented with chloramphenicol and kanamycin using transparent 96-well microplates (Greiner Bio-One, Frickenhausen, Germany). Individual colonies were inoculated using a sterile toothpick and the plate was sealed using a Breathe-Easy® sealing membrane (Sigma-Aldrich) and grown overnight at 30 °C with shaking (600 rpm). A black 96-well plate (Greiner Bio-One) was prepared by adding the required amount of naringenin stock solution to each well. Stock concentrations were chosen such that no more than 40 µL of stock had to be added to the wells. Following the addition, the solvent was evaporated for 10 min at 75 °C followed by further evaporation at RT. The selected culture medium was subsequently added to each well and β-alanine (BA) was supplemented. Finally, the overnight preculture was used to inoculate each well at a 300-fold dilution. This way, the final volume in each well totalled 200 µL. These operations were supported by the visual pipetting aid Pipette Show (Falk *et al.*, 2022). Optical density at 600 nm (OD_{600}) and mKate2 fluorescence (FL; excitation wavelength: 588 nm, emission wavelength: 633 nm) were measured every 20 min on a Synergy H1 microplate reader (BioTek Instruments Inc, Winooski, VT, USA) over the course of a 60 h culture experiment. Plates were incubated at 30 °C with shaking (425 rpm). The gain factor was set to 110 unless indicated otherwise.

Data analysis was performed using the programming language R 4.1.3 (R Core Team, 2022). For each time point, mean OD_{600} and FL values across four replicates of blank medium were subtracted from all measurements. Growth curves (OD_{600} plotted in func-

tion of time) and the time-dependency of FL were studied to select three consecutive time points where all cultures are in stationary phase and the FL profiles are sufficiently stable. For each well, the mean OD_{600} and FL was taken across these three time points. Subsequently $\frac{\mathrm{FL}}{\mathrm{OD}_{600}}$ could be computed (reported in *relative fluorescence units* or RFU).

Chapter 5

Results and Discussion

5.1 Mathematical modelling

5.1.1 Systematic analysis of the Mannan model for biosensor response for the repressed-repressor architecture

In order to better comprehend the model and the behaviour it encodes, extensive simulations of the Mannan model for biosensor response were performed. Briefly, each tunable parameter was perturbed and the effect on the dose-response relationships and the basal output b, threshold θ and dynamic range μ was assessed. The simulations were restricted to the repressed-repressor architecture for brevity. In each case, a biological explanation was provided for the observed behaviour. Where possible, the results were compared to published results of real-life biosensors based on the repressed-repressors LacI (*lac* repressor) and ArsR from *Escherichia coli*, *tet* repressor (TetR) from Tn10 and FapR from *Bacillus subtilis*.

In the case of a repressed-repressor, the basal transcription factor activity, b_1, corresponds to the concentration of active transcription factor (TF) that is able to bind to the transcription-factor binding site (TFBS) and repress transcription at full induction. One can indeed think of a transcription factor locked in the active conformation with no effector bound to its binding site. Perturbations in b_1 lead to very small shifts of TF that coincide with a large effect on the expression level of the output gene P (Figure 5.1). b

is largely unaffected, but θ and μ reveal interesting nonlinearities. It was already noted by Mannan *et al.* (2017) that μ decreases for increasing b_1, but not that θ reaches a maximum value. The large effect on μ can be partly ascribed to the high transcription factor-TFBS affinity K_2 used during the simulations: for high values of both b_1 and K_2, repressors will be able to bind to the TFBS even in the presence of effector, thus lowering the maximum increase in expression a and hence μ (see Equation 2.5) (Mannan *et al.*, 2017).

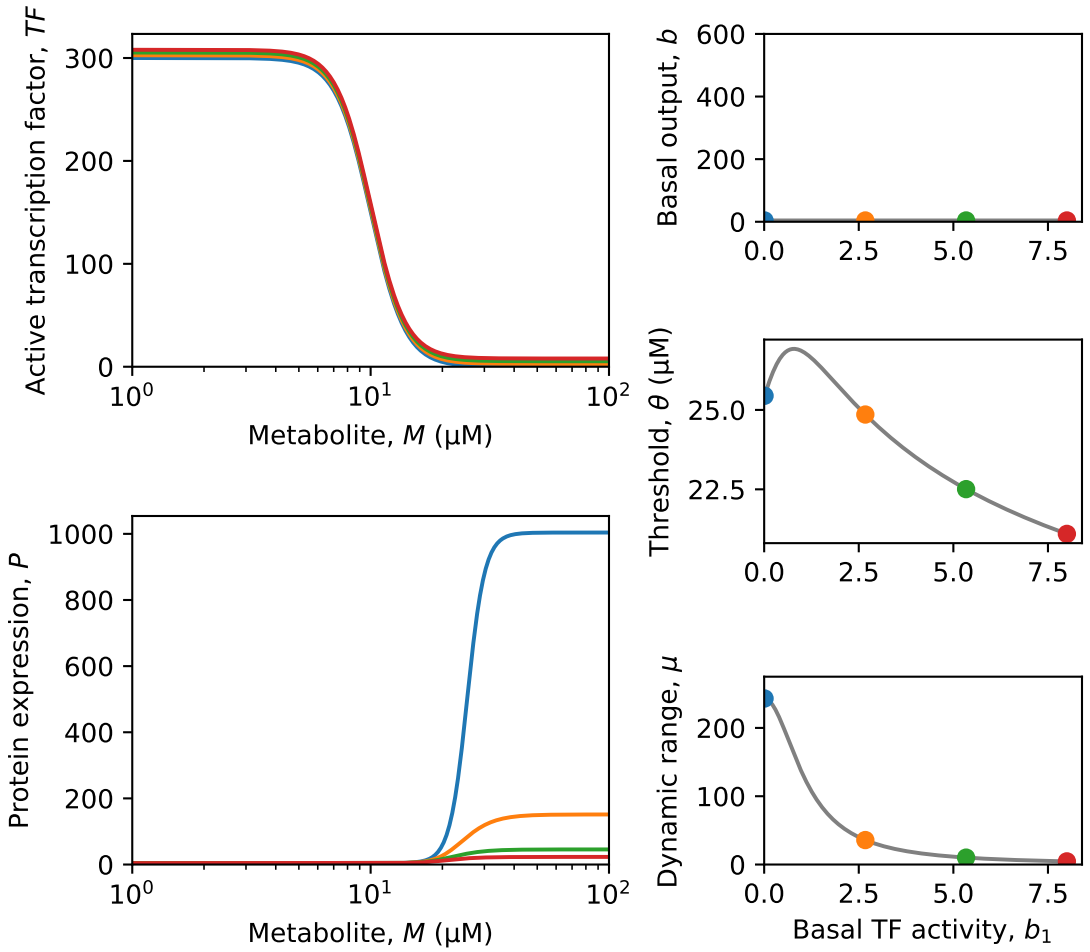

Figure 5.1: Model simulations perturbing the basal transcription factor activity (b_1) for the repressed-repressor architecture. Data were computed from Equations (2.13), (2.14) and (4.2) to (4.4) with parameters $a_1 = 300$, $K_1 = 0.1$, $n_1 = 6.0$, $b_2 = 4.1$, $a_2 = 1000$, $n_2 = 2.0$ and b_1 values span the range $b_1 = 0.01$ to $b_1 = 8$. Metabolite concentrations (M) are plotted on a logarithmic scale.

The maximum increase in transcription factor activity a_1 represents the portion of repressor that can be deactivated by the effector. Increasing a_1 corresponds to increases in the expression level of the transcription factor. Perturbations in a_1 alter b, θ and μ simultaneously (Figure 5.2). For small values of a_1, large values of b are reported that rapidly decrease for elevating a_1. For low-level transcription factor expression, there is indeed not enough repressor present to fully repress transcription of the target gene, consistent with the elevated basal output. It is therefore recommended to choose an adequately strong promoter for the transcription factor gene, as low levels of b are frequently desired for applications of whole-cell biosensors (Chen *et al.*, 2018; Hicks *et al.*, 2020). For intermediate levels of a_1, μ increases steeply, mainly due to the very small values of b in the denominator of Equation (2.5). However, the engineer should be aware that θ increases at the same time and that high-level expression of the transcription factor may impair growth. Higher transcription factor concentrations require higher metabolite concentrations to elicit the same effect, consistent with these results.

Although an intuitive explanation can be provided for the observed results for perturbations in a_1, literature does not corroborate these results unanimously. For FapR-based biosensors studied in *Saccharomyces cerevisiae*, increasing the metabolite-responsive transcription factor (MRTF) expression level by employing a higher *fapR* copy number led to a lower basal output and broader dynamic range (Li *et al.*, 2015; Chen *et al.*, 2018). However, an effect on θ can hardly be descried in these studies, as few data points were measured and fits to a Hill function of the form of Equation (2.4) were omitted. Furthermore, the threshold of a TetR-based biosensor could be markably decreased in *Salmonella enterica* by random mutagenesis of the promoter expressing *tetR*. a was slightly increased, but b unaffected. Western blot analysis confirmed that TetR expression was reduced for these mutated promoters (Georgi *et al.*, 2012). Wang *et al.* (2015) could also confirm that increased expression of TetR and ArsR led to increased thresholds of the corresponding *E. coli* biosensors. However, this was consistently accompanied by a significant decrease in μ, which appears to be in contradiction to the Mannan model. Increasing the MRTF expression level could perhaps lead to simultaneous increases in a_1 and b_1 in a system-dependent manner, possibly increasing or decreasing the dynamic range. More intricate couplings with other parameters can neither be excluded. Although its effect on μ seems to be difficult to predict, tuning the

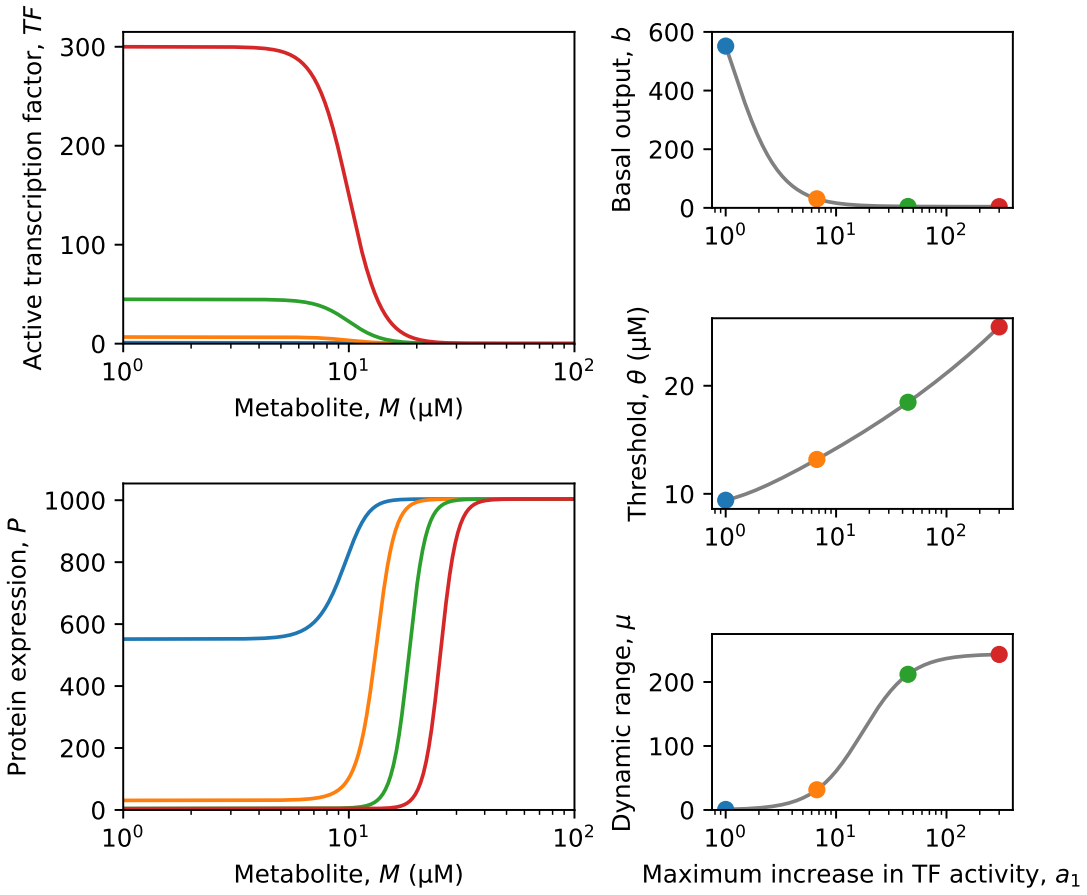

Figure 5.2: Model simulations perturbing the maximum increase in transcription factor activity (a_1) for the repressed-repressor architecture. Data were computed from Equations (2.13), (2.14) and (4.2) to (4.4) with parameters $b_1 = 0.01$, $K_1 = 0.1$, $n_1 = 6.0$, $b_2 = 4.1$, $a_2 = 1000$, $n_2 = 2.0$, and a_1 values span the range $a_1 = 1$ to $a_1 = 300$. Metabolite concentrations (M) and a_1 are plotted on a logarithmic scale.

MRTF expression level is very straightforward, making it an interesting parameter for response curve engineering.

Perturbations in the metabolite-transcription factor affinity K_1 confirm the orthogonal control of the threshold already discussed in Section 2.2.4 (Figure 5.3). The observation that $\theta \propto K_1^{-1}$ can easily be rationalised. For lower affinity (i.e. lower K_1, see Equation 2.13), higher metabolite concentrations are required to inactivate the transcription factor. Therefore, more active repressor is available at higher metabolite concentrations, leading to an enhanced overall threshold. An analogous reasoning can be followed for higher K_1.

The sensitivity of metabolite-transcription factor binding n_1 determines the steepness of the transcription factor activation curve (Figure 5.4). Higher values of n_1 correspond to steeper activation curves, in turn leading to steeper relationships between protein expression P and M. Transcription factors with a higher n_1 also exhibit lower thresholds because they are more responsive to their cognate metabolite, as a result of which the transcription factor becomes inactivated at lower concentrations. Although n_1 does not affect b and μ, one can hardly consider n_1 to provide orthogonal control of θ as cooperative binding is complex and so is engineering cooperativity (Falson *et al.*, 2004; Setny & Wiśniewska, 2018; Anashkin *et al.*, 2019).

The portion of output gene expression unaffected by repressor binding is represented by the basal level of promoter expression b_2. Perturbations in b_2 lead to vertical displacement of the relationship between P on the one hand and M or TF on the other hand (Figure 5.5). b_2 affects the basal output as $b \propto b_2$; its relationship to μ is more intricate. Hence, the basal output can be reduced by diminishing the leaky promoter expression. For repressors, this is determined to a considerable extent by the position of the TFBS in the promoter region. The strongest repression is generally achieved for a TFBS residing between the -35 and -10 core promoter elements, followed by TFBSs residing in proximal (downstream of the -10 sequence) and distal (upstream of the -35 sequence) locations, respectively (Lanzer & Bujard, 1988; Elledge & Davis, 1989). Multiple TFBSs have also been introduced in promoter regions to reduce b, to varying success (Li *et al.*, 2015; David *et al.*, 2016). For LacI and TetR, it has been shown that one TFBS suffices to achieve maximal repression in *E. coli*. However, this study used TFBSs exhibiting a

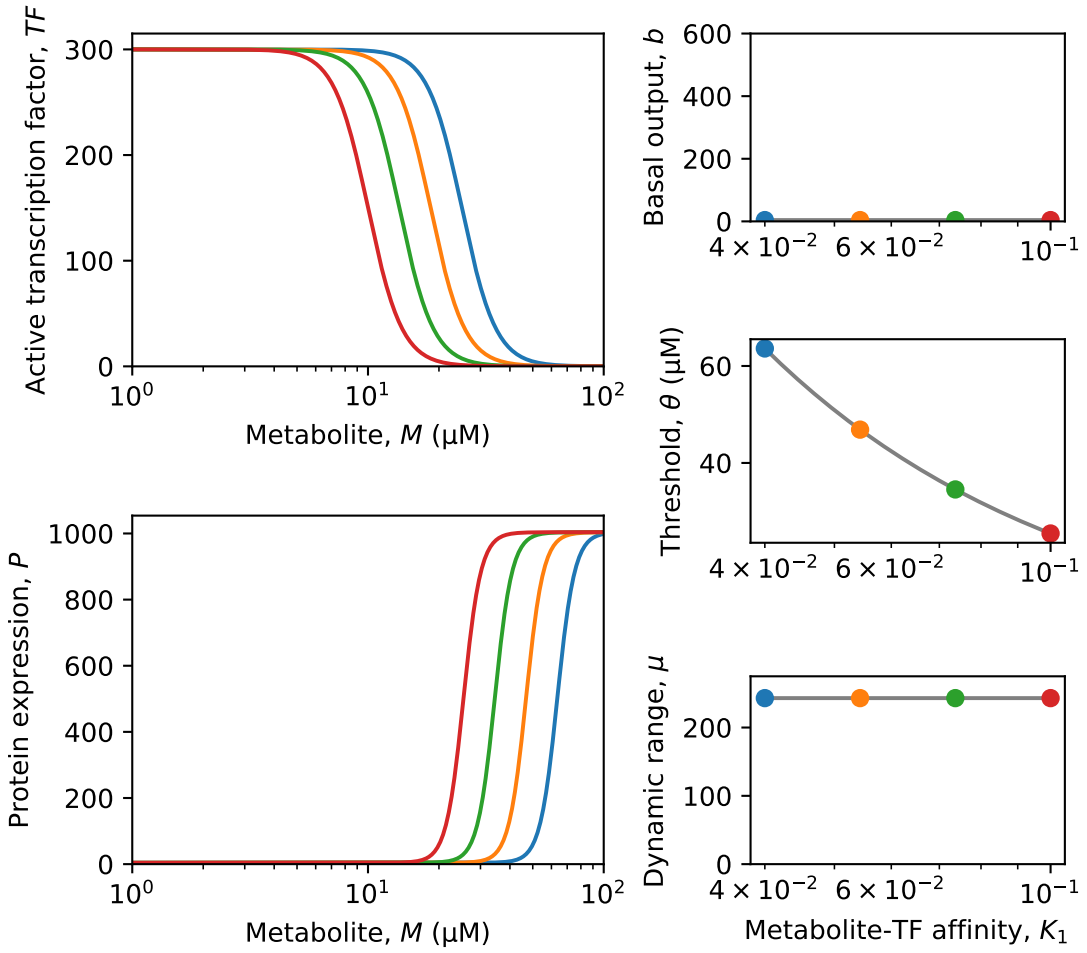

Figure 5.3: Model simulations perturbing the metabolite-transcription factor affinity (K_1) for the repressed-repressor architecture. Data were computed from Equations (2.13), (2.14) and (4.2) to (4.4) with parameters $b_1 = 0.01$, $a_1 = 300$, $n_1 = 6.0$, $b_2 = 4.1$, $a_2 = 1000$, $K_2 = 0.9$, $n_2 = 2.0$, and a_1 values span the range $K_1 = 4 \cdot 10^{-2}$ to $K_1 = 1 \cdot 10^{-1}$. Metabolite concentrations (M) and K_1 are plotted on a logarithmic scale.

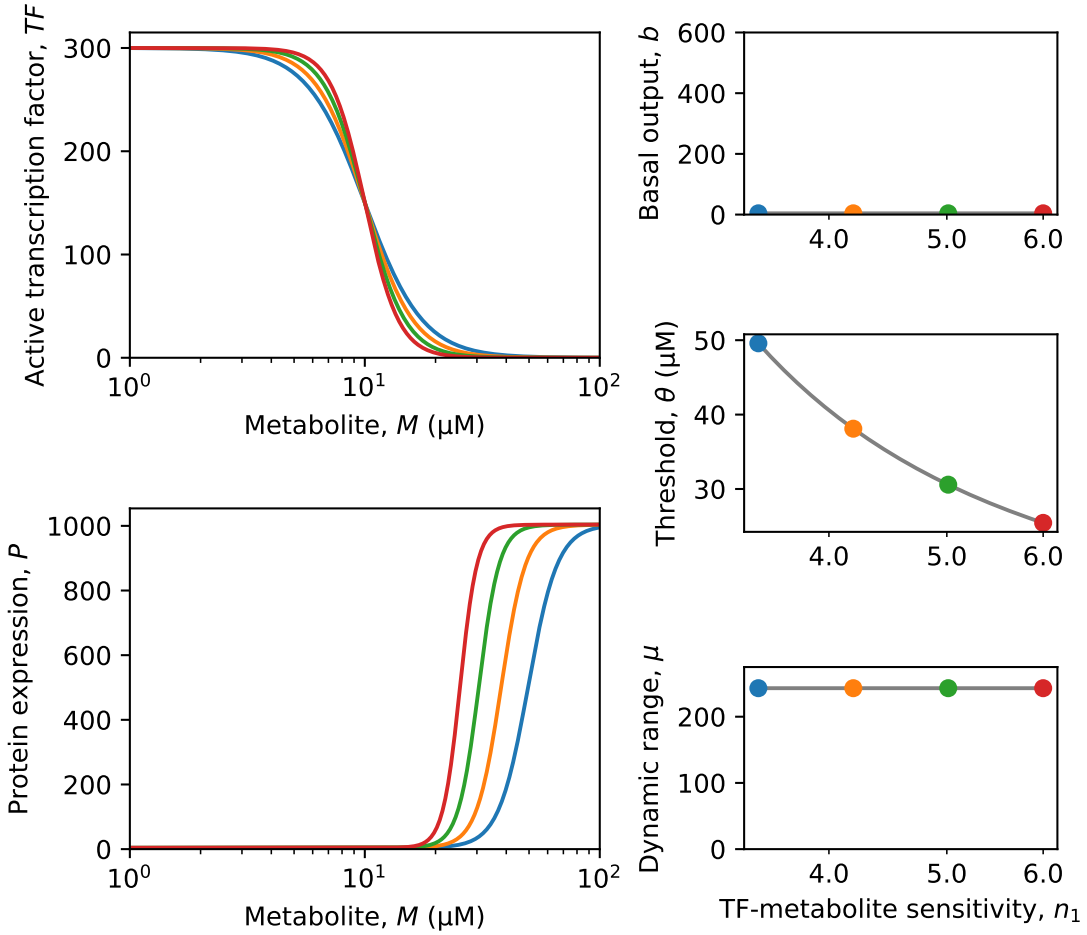

Figure 5.4: Model simulations perturbing the sensitivity of transcription factor-metabolite binding (n_1) for the repressed-repressor architecture. Data were computed from Equations (2.13), (2.14) and (4.2) to (4.4) with parameters $b_1 = 0.01$, $a_1 = 300$, $K_1 = 0.1$, $b_2 = 4.1$, $a_2 = 1000$, $K_2 = 0.9$, $n_2 = 2.0$, and n_1 values span the range $n_1 = 3.5$ to $n_1 = 6.0$. Metabolite concentrations (M) and n_1 are plotted on a logarithmic scale.

high affinity for the MRTF and a high MRTF expression level (Cox III *et al.*, 2007).

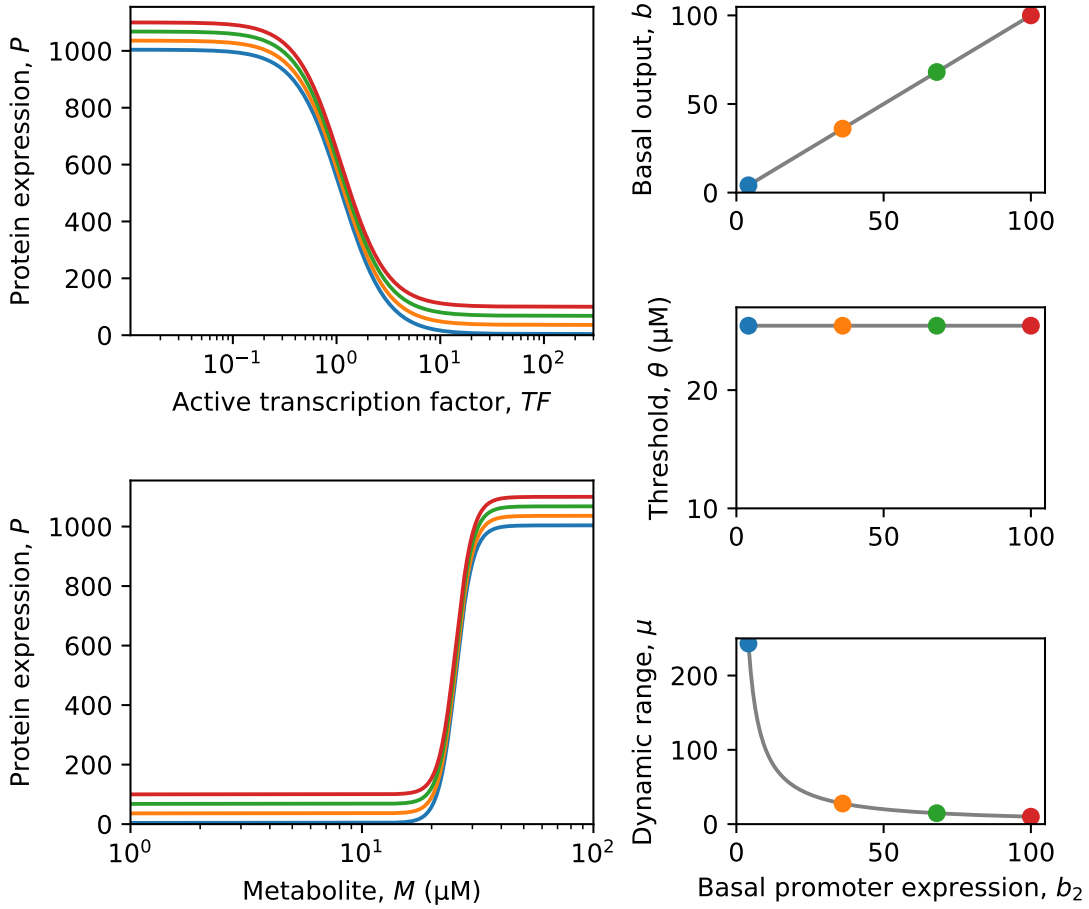

Figure 5.5: Model simulations perturbing the basal level of promoter expression (b_2) for the repressed-repressor architecture. Data were computed from Equations (2.13), (2.14) and (4.2) to (4.4) with parameters $b_1 = 0.01$, $a_1 = 300$, $K_1 = 0.1$, $n_1 = 6.0$, $a_2 = 1000$, $K_2 = 0.9$, $n_2 = 2.0$, and b_2 values span the range $b_2 = 4.10$ to $b_2 = 100$. Metabolite concentrations (M) are plotted on a logarithmic scale.

The maximum increase in promoter expression a_2 pertains to the output expression level at full induction and relative to the basal level. Increases in a_2 can be brought about by increasing the promoter strength (Mannan *et al.*, 2017). The orthogonal control of the dynamic range by the promoter dynamic range $\mu_2 = \frac{a_2}{b_2}$ was already discussed in Section 2.2.4 and is supported by the simulations in Figure 5.6. However, careful analysis of Equation (4.4) reveals that not only μ, but also b is a function of a_2. However, the

fraction in Equation (4.4) is dominated by the denominator under the chosen parameter settings. Therefore, this term only increases with 0.01 at most, which is not visibly discernible on Figure 5.6. However, this term does become significant for lower values of a_1, K_2 and/or n_2 (shown in Figure I.1 for a_1). Under such conditions, the increase in dynamic range is limited by collateral damage of a_2 on b, as $b \propto a_2$. This suggests that, under conditions of low MRTF expression or low MRTF-TFBS affinity, not all TFBSs will be occupied and will be expressed to a promoter strength-dependent extent. This trend can be observed for some biosensors; for others only the increase in dynamic range is observed (Cox III *et al.*, 2007).

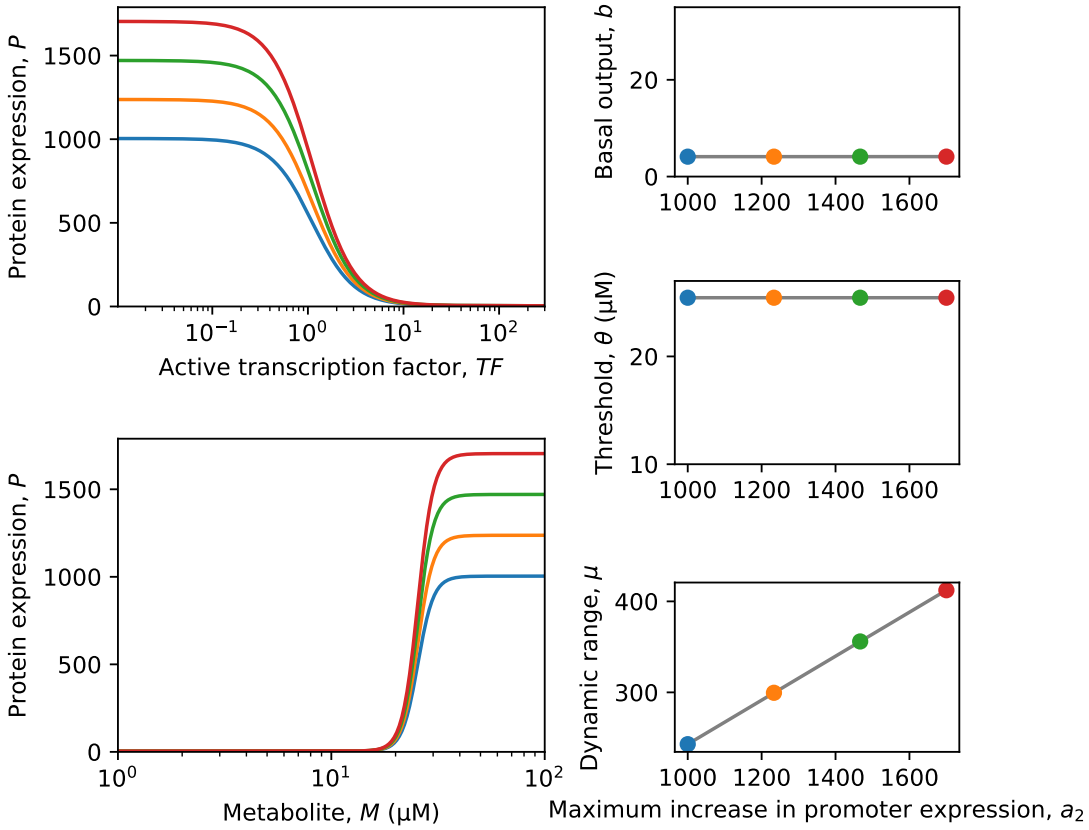

Figure 5.6: Model simulations perturbing the maximum increase in promoter expression (a_2) for the repressed-repressor architecture. Data were computed from Equations (2.13), (2.14) and (4.2) to (4.4) with parameters $b_1 = 0.01$, $a_1 = 300$, $K_1 = 0.1$, $n_1 = 6.0$, $b_2 = 4.1$, $K_2 = 0.9$, $n_2 = 2.0$, and a_2 values span the range $a_2 = 1000$ to $a_2 = 1700$. Metabolite concentrations (M) are plotted on a logarithmic scale.

The effect of tuning K_2 has also been thoroughly discussed in Section 2.2.4. The simulations in Figure 5.7 are in accordance with (Mannan *et al.*, 2017). For higher values of K_2, lower amounts of active transcription factor are required to fully repress transcription of the output gene. Therefore, higher metabolite concentrations are required for derepression, explaining the increased threshold. For low values of K_2, the basal output increases steeply. In this regime, the maximum level of active transcription factor ($b_1 + a_1$) does no longer suffice to fully repress transcription of the target gene even in the absence of metabolite, leading to an enhanced basal output. Note also the striking agreement between the effect of tuning a_1 and K_2. Indeed, Equation (2.14) can easily be rewritten to encompass the product $K_2 \cdot a_1$, which explains why a_1 and K_2 are partly interchangeable. This complementarity can lead to *unidentifiability*, a problem during parameter estimation (Waterfall *et al.*, 2006; Dray *et al.*, 2022). The behaviour becomes more dissimilar for increased b_1 values (see Figure I.2).

By analogy with n_1, increasing the sensitivity of transcription factor-TFBS binding n_2 leads to steeper P versus M or TF curves. However, in contrary to n_1, θ is not notably affected but b and μ are (Figure 5.8). Consider the blue curve in Figure 5.8: due to its subsensitive response to varying concentrations of active transcription factor, P only increases with 1.5% and 97% of a_2 relative to b_2 for $M = 0$ and $M \to +\infty$, respectively.

5.1.2 Development of a parameter estimation method for the Mannan model for biosensor response

Parameter estimation method (PEM) evaluation data sets were generated based on experimental data of Mannan *et al.* (2017) and show a range of distinct behaviours (see Figure 5.9). For instance, μ, θ and b varied widely for data set 6, whereas the change in μ and b for increasing θ was only limited for data set 7. A constant, however, across all data sets is the large amount of noise. This poses a significant challenge to the PEM. For each data set, the data points were drawn from a normal distribution with standard deviations of (μ, θ, b) computed from real experimental data from Mannan *et al.* (2017) (see Section 4.1.2.2). The variation is therefore expected to be biologically relevant, but remind once more of a common issue in molecular biology (Lillacci & Khammash,

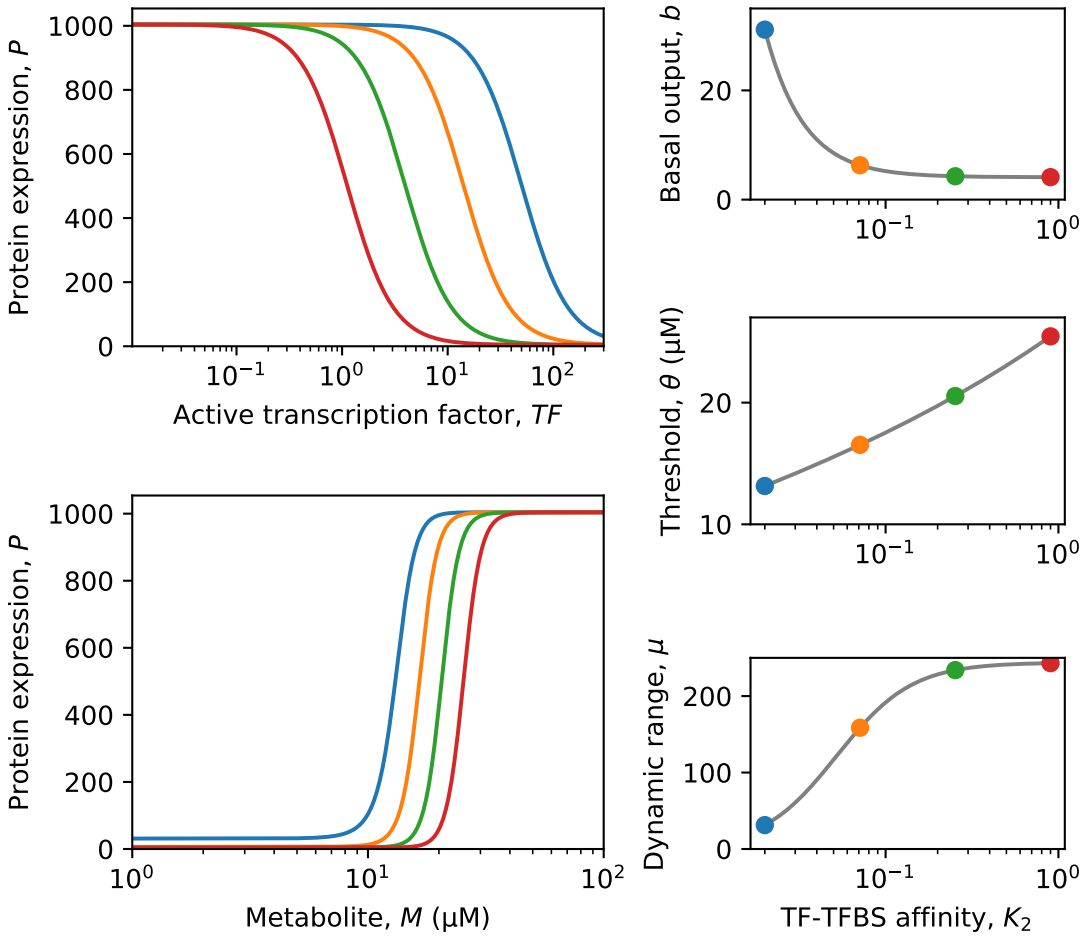

Figure 5.7: Model simulations perturbing the transcription factor-TFBS affinity (K_2) for the repressed-repressor architecture. Data were computed from Equations (2.13), (2.14) and (4.2) to (4.4) with parameters $b_1 = 0.01$, $a_1 = 300$, $K_1 = 0.1$, $n_1 = 6.0$, $b_2 = 4.1$, $a_2 = 1000$, $n_2 = 2.0$, and K_2 values span the range $K_2 = 0.02$ to $K_2 = 0.9$. Metabolite concentrations (M) and K_2 are plotted on a logarithmic scale.

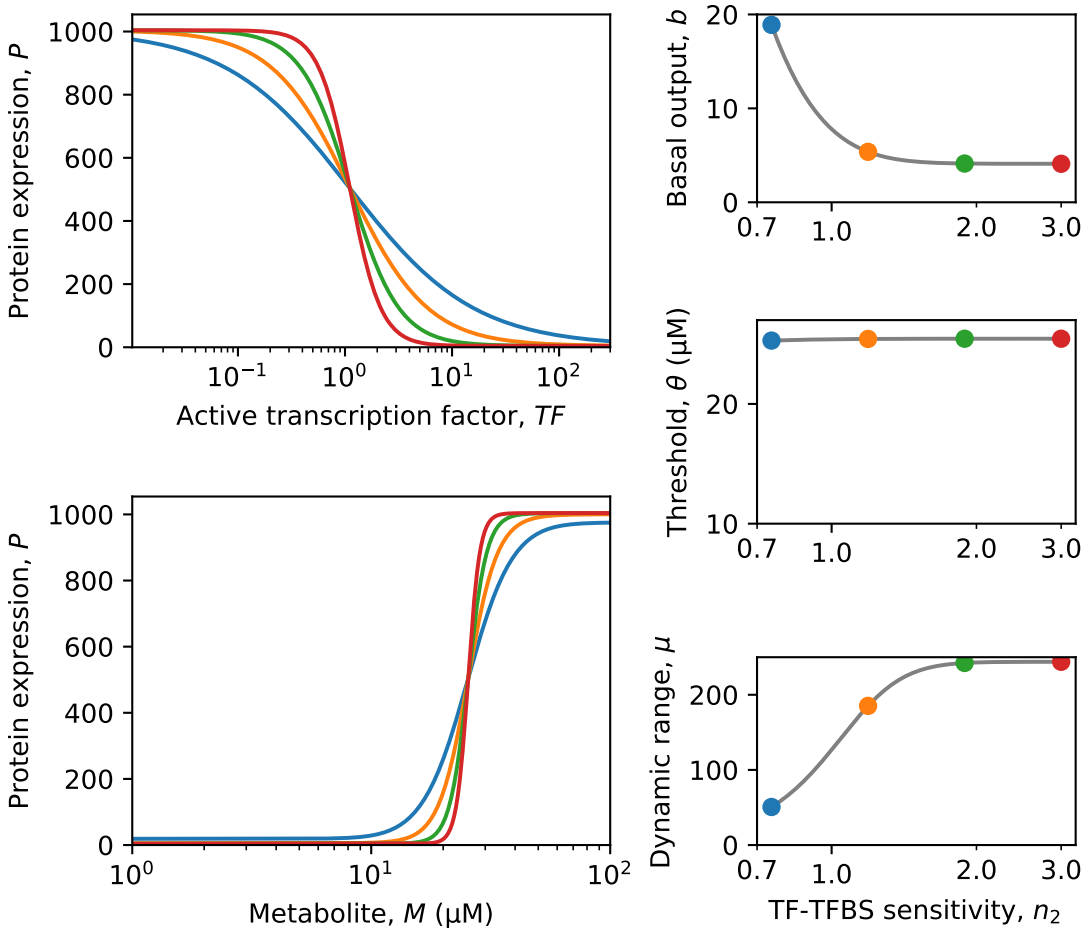

Figure 5.8: Model simulations perturbing the sensitivity of transcription factor-TFBS binding (n_2) for the repressed-repressor architecture. Data were computed from Equations (2.13), (2.14) and (4.2) to (4.4) with parameters $b_1 = 0.01$, $a_1 = 300$, $K_1 = 0.1$, $n_1 = 6.0$, $b_2 = 4.1$, $a_2 = 1000$, $K_2 = 0.9$, and n_2 values span the range $n_2 = 0.75$ to $n_2 = 3.0$. Metabolite concentrations (M) and n_2 are plotted on a logarithmic scale.

2010). The value of C_{pass}, the objective function value comparing a PEM evaluation data set before and after noise addition (see Section 4.1.2.2), can be used as a measure for the amount of noise in a given data set. This indicated that PEM evaluation data set 1 was least noisy; data set 7 showed the most noisy.

The entire PEM evaluation procedure took approximately 39 h to run on a single 3.20 GHz processor of a Dell XPS 8930 computer. All fits could reproduce the overall behaviour of the data sets, except for one (data set 7). Moreover, the objective function values reached by the best 290 optimisation runs surpassed C_{pass} for all data sets. These observations suggest that the PEM is implemented properly, and that the behaviour of the simulated data sets can be reproduced by minimising Equation (4.1). Below, overall trends across the results are indicated by means of three representative PEM data sets. The remainder of the results can be consulted in Appendix I.

For each data set, the best fit (yielding the lowest objective function value upon optimisation) is reported along with the remainder of the fits. This reveals that a significant portion of the optimisation runs return model parameters that yield bad fits. This might occur if the objective function exhibits local minima. In this case, the algorithm could then get stuck in such a minimum and exit as it cannot further minimise Equation (4.1). However, the performance visualisations do not support the presence of local minima as the main cause of this observation. If local minima were the main bottleneck, the performance visualisation would show a step-by-step decrease of the objective function with multiple runs reproducibly returning the same objective, corresponding to a local minimum. Our results displayed a rather continuously decreasing profile of the objective function, steep for high C but gradually levelling out to reach a global optimum. The experimental noise is neither a likely cause, as the same phenomena were observed when fitting a simulated data set with a smaller amount of noise (results not shown).

It was noted that, among the ”worst” fits (those displaying the highest objective function values upon optimisation) across the data sets, most optimal parameter values resided near the bounds (results not shown). This was not true for the best fits (Figures I.9 to I.13), except for b_1, that appeared to be more attracted to the boundaries for the best fits as compared to the worst fits. It has been put forward that when the algorithm meets a boundary of parameter space, it might get stuck there and return a suboptimal fit, unable

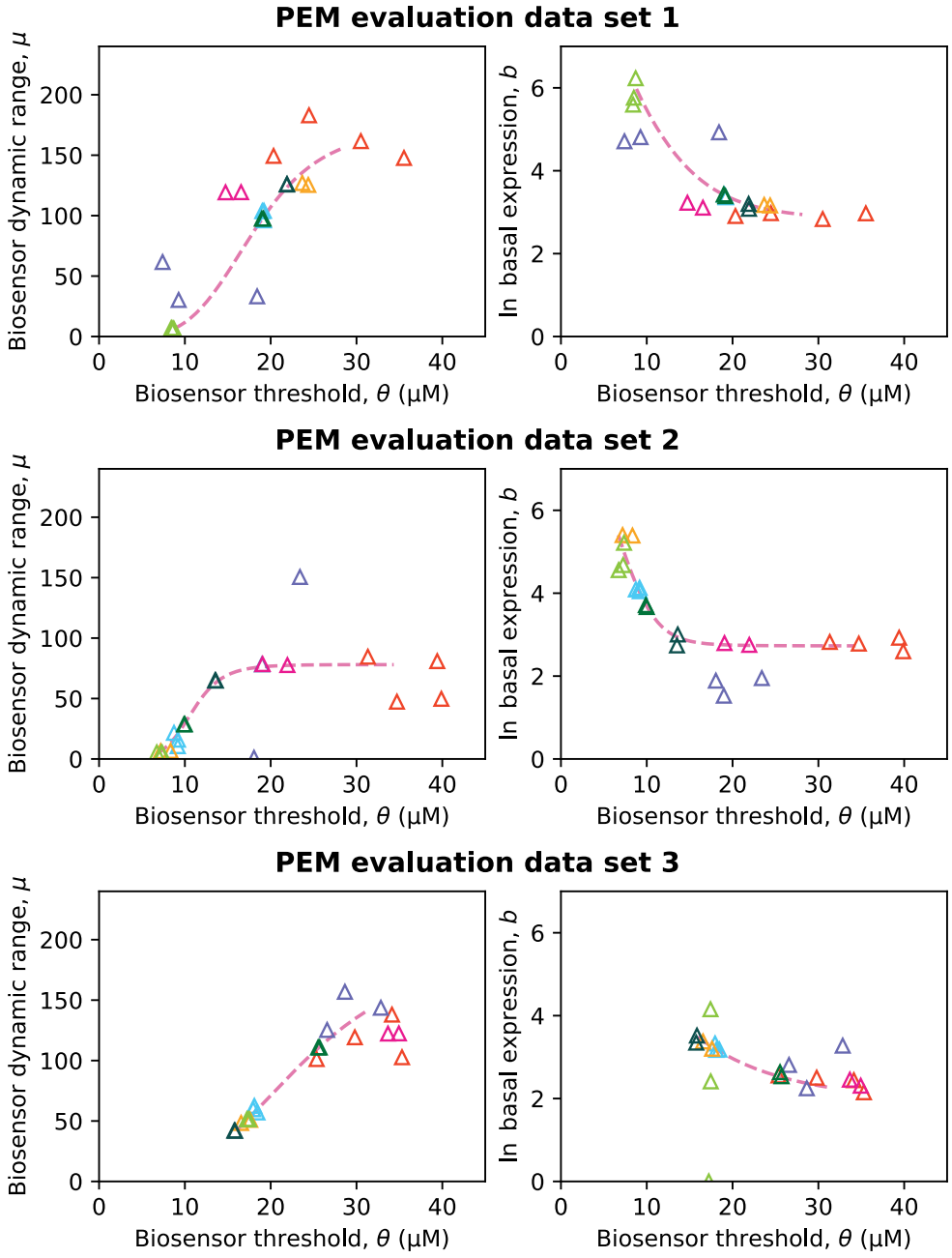

Figure 5.9: Simulated data sets used for parameter estimation method (PEM) evaluation. For each data set, the pink dashed line corresponds to an interpolation of ($\mu_{i,True}$, $\theta_{i,True}$, $b_{i,True}$), the dose-response parameters prior to noise addition. The dose-response parameters obtained after noise addition (μ_i, θ_i, b_i) make up the PEM evaluation data and are represented as triangles, colour-coded according to their strain i. (Continued on next page)

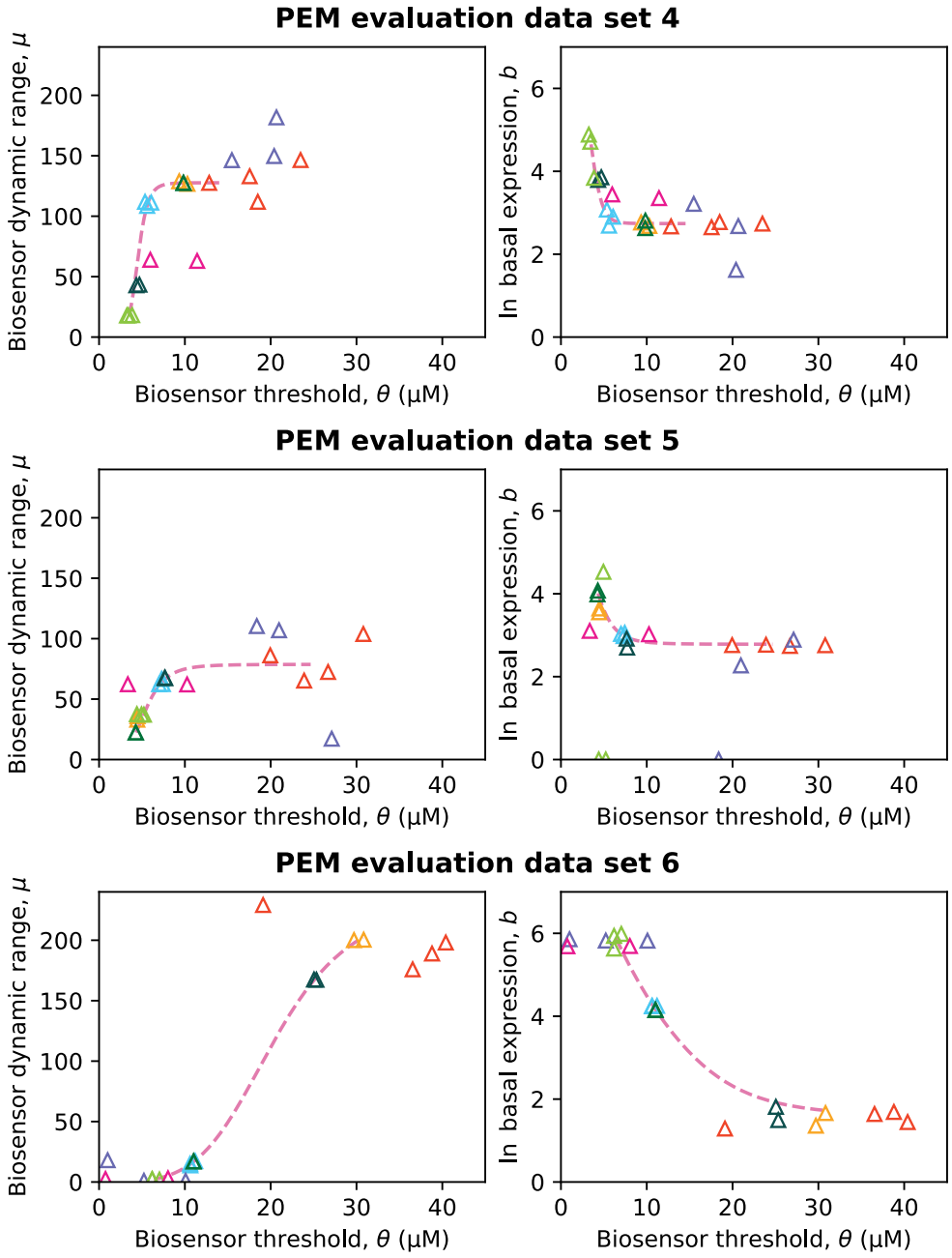

Figure 5.9: Simulated data sets used for parameter estimation method (PEM) evaluation. For each data set, the pink dashed line corresponds to an interpolation of ($\mu_{i,True}$, $\theta_{i,True}$, $b_{i,True}$), the dose-response parameters prior to noise addition. The dose-response parameters obtained after noise addition (μ_i, θ_i, b_i) make up the PEM evaluation data and are represented as triangles, colour-coded according to their strain i. (Continued on next page)

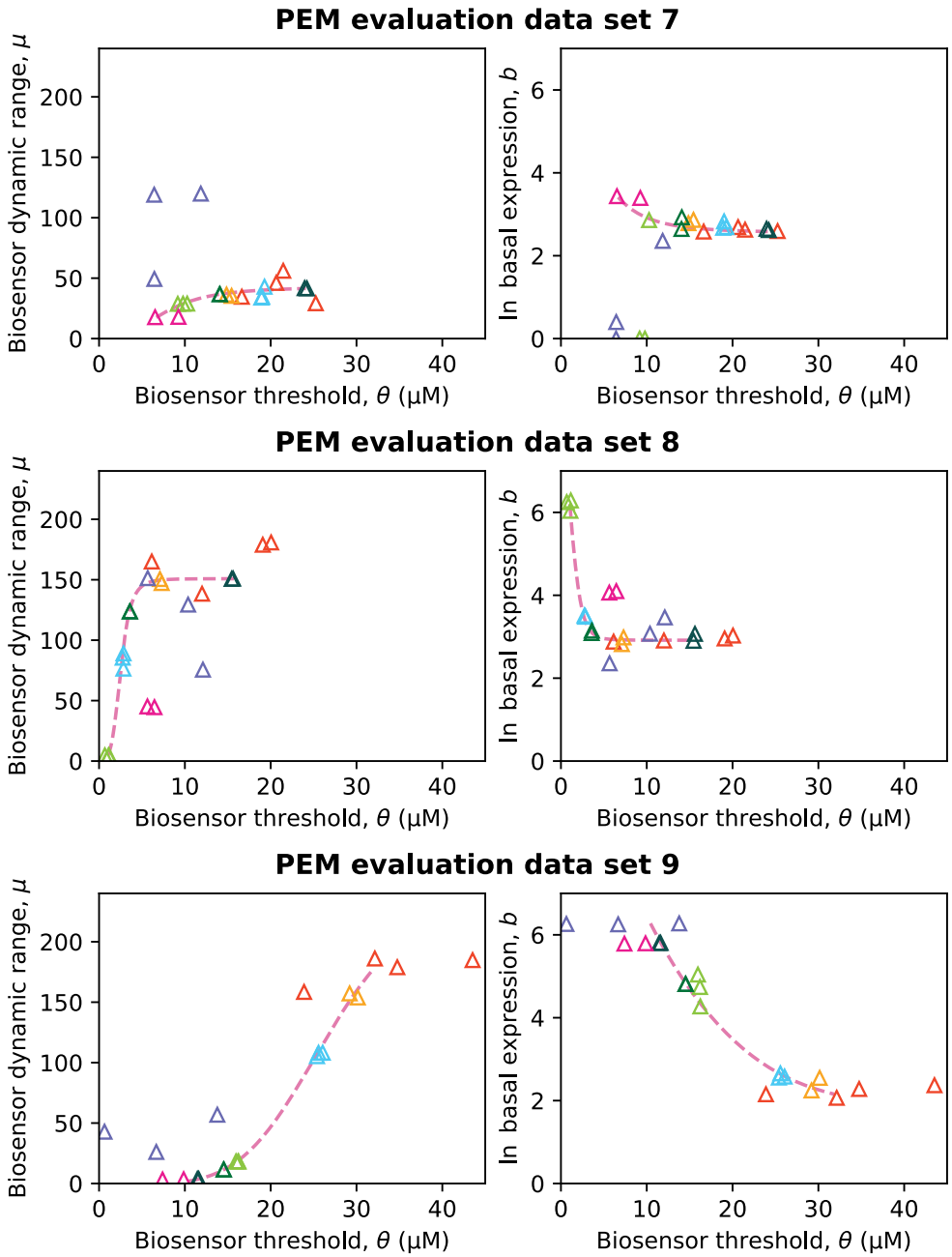

Figure 5.9: Simulated data sets used for parameter estimation method (PEM) evaluation. For each data set, the pink dashed line corresponds to an interpolation of ($\mu_{i,True}$, $\theta_{i,True}$, $b_{i,True}$), the dose-response parameters prior to noise addition. The dose-response parameters obtained after noise addition (μ_i, θ_i, b_i) make up the PEM evaluation data and are represented as triangles, colour-coded according to their strain i. (Continued)

to move away from the boundary towards the global optimum (Transtrum *et al.*, 2010). Our approach based on Latin hypercube sampling (LHS) ensures in principle that some initial guesses are generated in proximity to the boundaries of parameter space. However, Mannan *et al.* (2017) reported that they consistently recovered the global optimum across 500 optimisation runs. Furthermore, their resulting fits only deviated slightly from the best fit. This implies that they implemented the PEM differently, for instance, by generating initial guesses through a method distinct of LHS or using a different optimisation algorithm with superior convergence.

Notwithstanding the noise in the data sets and the suboptimal performance of the developed PEM, the best fit nearly overlapped the ideal noise-free solution for data set 1 (Figure 5.10B). It was pulled to slightly higher μ and lower b, fitting the noise to a limited extent. This is also reflected by the objective function value (Figure 5.10C). The algorithm succeeded to minimise Equation (4.1) further than the objective function value of the ideal noise-free solution, *overfitting* the data. PEM runs corresponding to data sets 2, 6 and 9 reached similarly high quality of fitting (Figures I.3, I.6 and I.8). Notably, PEM evaluation data sets 1, 6 and 9 are the least noisy of the nine considered data sets, as reflected by their C_{pass}. Data set 2 on the other hand has the third highest value of C_{pass} of all data sets.

Data set 3 suits as an example case of fits that reproduced the main trend in the data, but lack some features of the ideal noise-free solution. The best fit of data set 3 did not adequately reproduce the curvature of the noise-free solution, especially for the relationship between μ and θ (Figure 5.11B). This is due to overfitting and a lack of data points for intermediate θ (between 20 µM and 25 µM). Data sets 4, 5 and 8 displayed similar faults in fitting (Figures I.4, I.5 and I.7). Notably, these data sets showed superior convergence as compared to the data sets yielding better fits, represented by data set 1. In case of data set 3, the objective function value returned by the 400 best runs was essentially the same (Figure 5.11C). Although a slight upward tendency can be noticed starting from run index 50, model parameters corresponding to the 100 best runs produced visibly overlapping fits (results not shown). This counterintuitive observation might be due to the higher occurrence of outliers in these data sets. These play a relatively large part in the value of the objective function, possibly increasing the *basin of attraction*. For multistart deterministic optimisation (as implemented in our PEM),

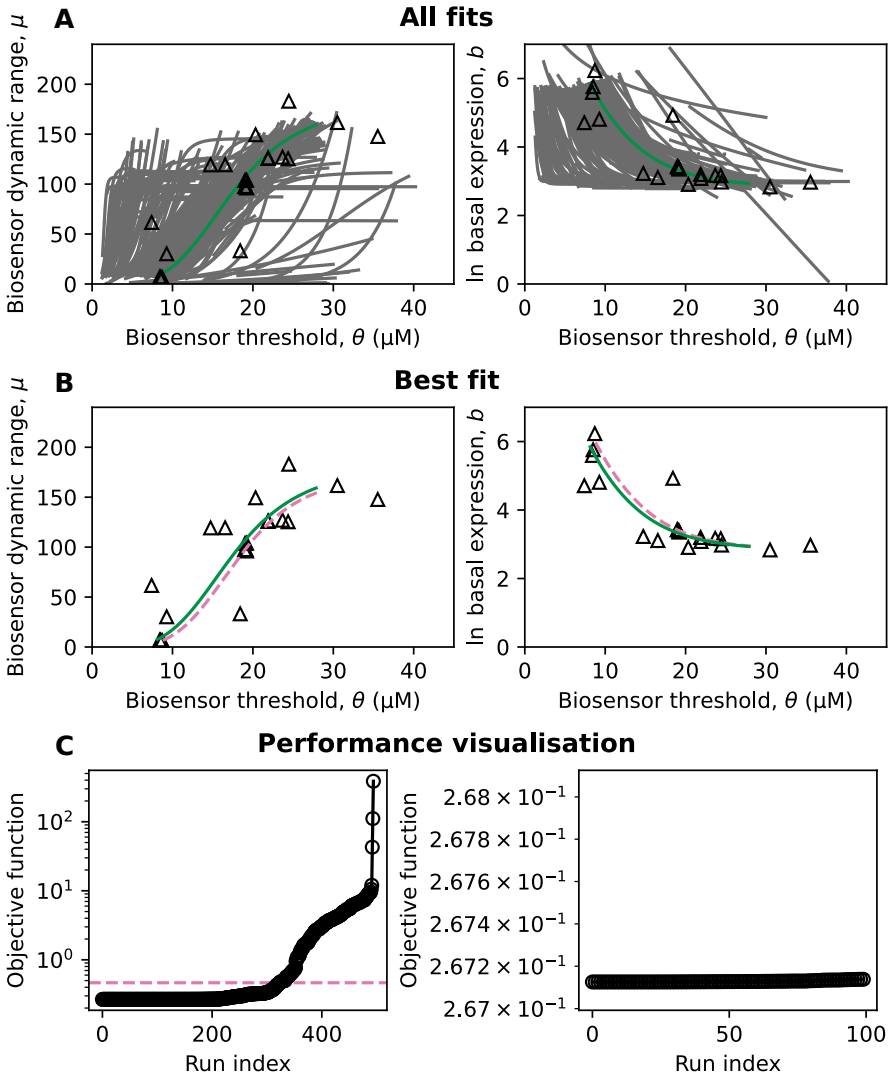

Figure 5.10: Results of fitting to parameter estimation method (PEM) evaluation data set 1. **(A)** All 500 fits to the data. The best fit (displaying the lowest objective function value) is plotted as in green, the remainder of the fits is shown in grey. **(B)** Only the best fit is shown in green, along with the ideal noise-free solution (an interpolation of ($\mu_{i,True}$, $\theta_{i,True}$, $b_{i,True}$)) plotted as a pink dashed line. In panels A and B, triangles represent the dose-response parameters of PEM evaluation data set 1 (the training data). **(C)** Performance visualisation for the PEM evaluation problem. Left panel: the objective function values obtained after 500 runs of least-squares optimisation (run index 0 up to 499) are sorted increasingly. The horizontal dashed line represents C_{pass}, the objective function value computed for the ideal noise-free solution. Right panel: a close-up on the 100 best runs suggests that a global optimum is reached. C_{pass} is beyond the axis limits.

the rate of convergence to the global optimum is correlated with the size of the basin of attraction, which could explain this observation (Raue *et al.*, 2013).

Finally, the best fit obtained for data set 7 did not capture the true trend at all (Figure 5.12B). Not surprisingly, this data set exhibited the largest level of noise of all data sets under consideration. Three outlier data points at higher dynamic range influence the objective function so much, that the most effective way to minimise it is by drawing the fit almost vertically (left panel in Figure 5.12B). Four outlying points have a similar effect on the b component of the fit (right panel). The objective function value of the best fit is 4 times smaller than C_{pass}, pointing to substantial overfitting (Figure 5.12C). However, we cannot assume that a global optimum is reached for this PEM evaluation problem, as the optimum with lowest objective function value was only returned 12 times.

With the exception of data set 7, the objective function seems to converge to the same, global optimum for each distinct PEM evaluation data set. This idea is strengthened by the fact that the 100 best fits of these data sets visibly overlap (for data sets 2, 6 and 9, the 82, 80 and 59 best runs were used, respectively, to exclude runs that converged to a distinct objective function value, results not shown). However, only data set 1 yielded consistent parameter estimates close to the true parameter values for the 100 best optimisation runs (Figures I.9 to I.13). Only b_1 was not reasonably close to the true value (Figure I.9A). Although all parameter values were estimated reproducibly for data sets 2 and 5, they were more dissimilar to the true values. The remaining PEM evaluation problems all yielded at least one variable parameter. Notably, a_1 and K_2 show variation across all of these data sets. Conversely, b_2 and a_2 could each time be uniquely estimated.

The observations raised above suggest that for some data sets, the model exhibits unidentifiability. Unidentifiability is a common issue in systems biology models that occurs when some regions of the objective function landscape are flat. This implies that the objective is insensitive to changes in some parameters, and distinct parameter sets can yield similarly good fits. The fact that this unidentifiability is not observed across all considered PEM evaluation data sets argues that the problem is related to the experimental data, and that is not an inherent property of the model (practical identifiability

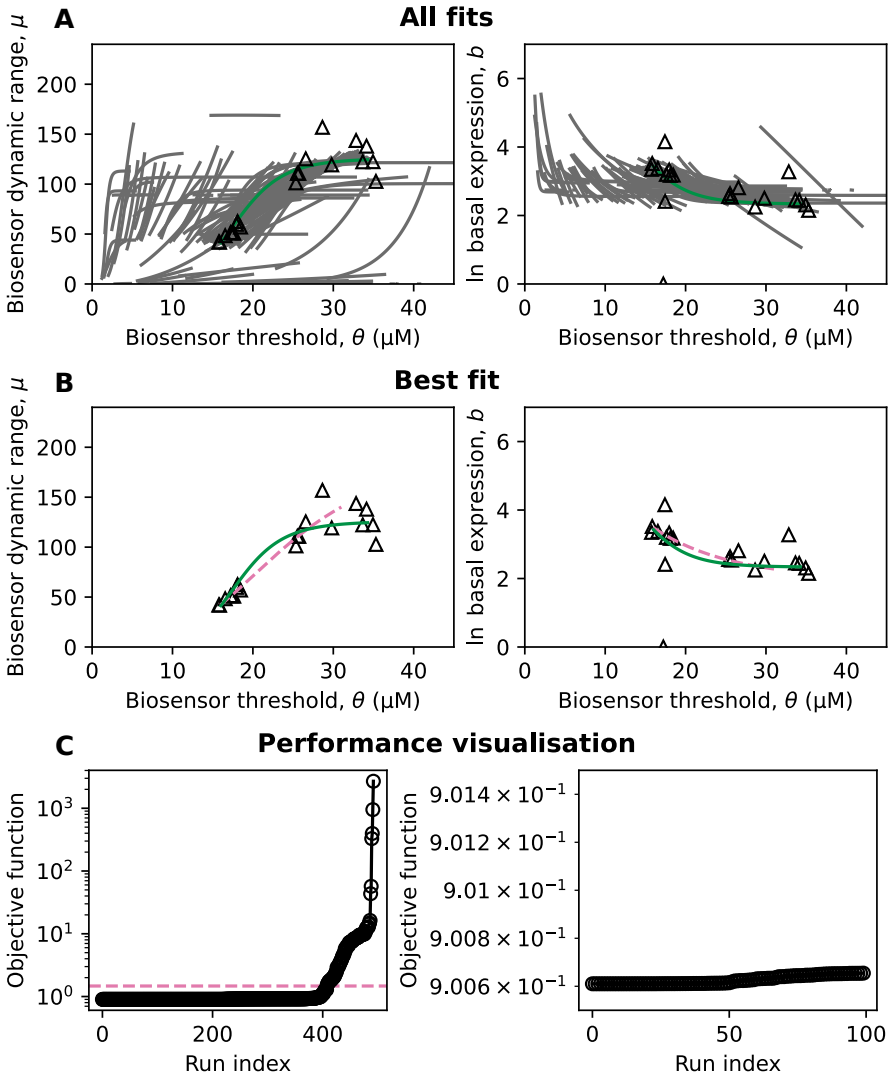

Figure 5.11: Results of fitting to parameter estimation method (PEM) evaluation data set 3. **(A)** All 500 fits to the data. The best fit (displaying the lowest objective function value) is plotted as in green, the remainder of the fits is shown in grey. **(B)** Only the best fit is shown in green, along with the ideal noise-free solution (an interpolation of ($\mu_{i,True}$, $\theta_{i,True}$, $b_{i,True}$)) plotted as a pink dashed line. In panels A and B, triangles represent the dose-response parameters of PEM evaluation data set 3 (the training data). **(C)** Performance visualisation for the PEM evaluation problem. Left panel: the objective function values obtained after 500 runs of least-squares optimisation (run index 0 up to 499) are sorted increasingly. The horizontal dashed line represents C_{pass}, the objective function value computed for the ideal noise-free solution. Right panel: a close-up on the 100 best runs suggests that a global optimum is reached. C_{pass} is beyond the axis limits.

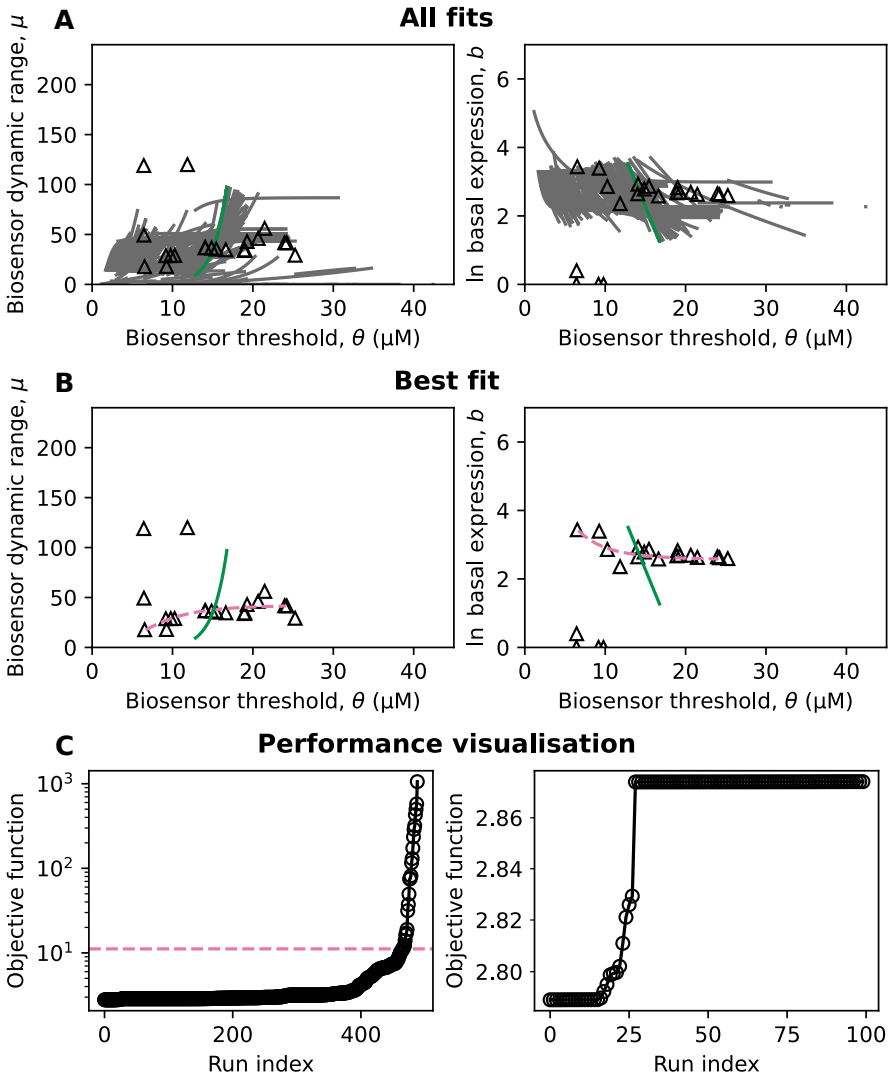

Figure 5.12: Results of fitting to parameter estimation method (PEM) evaluation data set 7. **(A)** All 500 fits to the data. The best fit (displaying the lowest objective function value) is plotted as in green, the remainder of the fits is shown in grey. **(B)** Only the best fit is shown in green, along with the ideal noise-free solution (an interpolation of $(\mu_{i,True}, \theta_{i,True}, b_{i,True})$) plotted as a pink dashed line. In panels A and B, triangles represent the dose-response parameters of PEM evaluation data set 7 (the training data). **(C)** Performance visualisation for the PEM evaluation problem. Left panel: the objective function values obtained after 500 runs of least-squares optimisation (run index 0 up to 499) are sorted increasingly. The horizontal dashed line represents C_{pass}, the objective function value computed for the ideal noise-free solution. Right panel: a close-up on the 100 best runs cannot assure that a global optimum is reached. C_{pass} is beyond the axis limits.

rather than structural). However, unambiguous identification of the type of identifiability and the unidentifiable parameters requires specialised methodologies, for example the analysis of the parameter *profile likelihood* (Raue *et al.*, 2009, 2010, 2011; Wieland *et al.*, 2021). Such analysis exceeds the scope of this thesis.

In conclusion, the PEM developed during this thesis and described in Section 4.1.2.1 can yield good fits, but some issues should be addressed before it can be used for real-life problems: suboptimal convergence, overfitting and potential unidentifiability. The latter two presumably originate from the quality of the experimental data. While based on real-life data, some data points are clearly outliers. In reality, one would probably consider to repeat these experiment; however, we did not omit these outliers to prevent any bias. Follow-up experiments and remediation strategies for these problems are briefly proposed below.

It has been suggested that the suboptimal convergence, conveyed by the relatively smooth appearance of the sorted objective function values upon optimisation, would be caused by interference of the parameter bounds. This can be alleviated by employing the *geodesic* Levenberg-Marquardt algorithm, an adapted version with improved convergence (Transtrum *et al.*, 2010, 2011). *Trust region methods* also exhibit faster convergence when far from the global optimum (Vanden Berghen, 2004).

One way to remedy overfitting is to reduce the experimental variation. Increasing the amount of biological replicates leads to a better estimation of the variation, however, the standard deviation of the experimental results does not necessarily decrease (Cumming *et al.*, 2007). A better solution would be to resort to robust regression, where the influence of outliers is diminished by including a loss function in the objective function (Huber, 1981). One could also consider to replace the μ term in Equation (4.1). b and a can be directly obtained from fitting, whereas μ is computed from Equation (2.5) and therefore incorporates both the error on b and a. Instead, a or $\lg(a)$ could be included in Equation (4.1). Further experiments should verify whether this improves the PEM.

If the encountered difficulties to estimate some parameter values (see Figures I.9 to I.10, I.12 and I.13) were truly caused by practical unidentifiability, they should also be alleviated by improving the quality of the experimental data. In addition, the amount of data could be increased by measuring the dose-response curves of more TFBS mutants (Raue

et al., 2009). If this would be infeasible, the final option is to remove the unidentifiable parameter from the model (Maiwald *et al.*, 2016).

5.2 *In vivo* characterisation of biosensors

Biosensors based on the MRTFs OapR from *Cupriavidus necator* H16 and AHOS_RS02205 from "*Acidianus hospitalis*" were studied in the industrially important bacterial species *E. coli*. In advance of the results presented in Sections 5.2.2 and 5.2.3, these MRTFs confer differential regulation in response to the supplementation of their effector β-alanine (BA), leading to functional biosensors. However, their mechanism and even their regulatory pattern might be different in their native context, as some accessory factors are likely to be absent in *E. coli*.

Where feasible, the dose-response curves were parameterised by fitting to a Hill function (see also Section 4.1.1). Some parameters are quoted throughout the discussion, but all best-fit parameters can be consulted in Appendix II.

5.2.1 Plasmid systems used for the assembly of biosensors

All studied biosensors were constructed using a modular two-plasmid system, consisting of a promoter reporter constructs (pPRC) and inducible transcription factor constructs (pITC) (Figure 5.13A). pPRC plasmid vectors harbour the *mkate2* fluorescent reporter gene. The promoter of interest can be conveniently introduced upstream of *mkate2* using *seamless ligation cloning extract* (SLiCE) cloning (Zhang *et al.*, 2012). In addition, the MRTF gene under study can be introduced in multiple cloning site 4 (MCS4) on the pITC vector. This way, it is placed under transcriptional control of the naringenin-inducible *fdeAR* promoter which is regulated by the FdeR transcription factor from *Herbaspirillum seropedicae* (De Paepe *et al.*, 2018). This allows facile tuning of the expression level of the MRTF of interest, as the supplementation of different concentrations of naringenin provides different levels of induction of the MRTF gene (Bernauw *et al.*, 2022).

Using the above two-plasmid system, a biosensor can be assembled by co-transforming the pITC-TF (pITC harbouring the transcription factor gene) and pPRC-Prom (pPRC harbouring the corresponding promoter upstream of the *mkate2* gene) in *E. coli* MG1655.

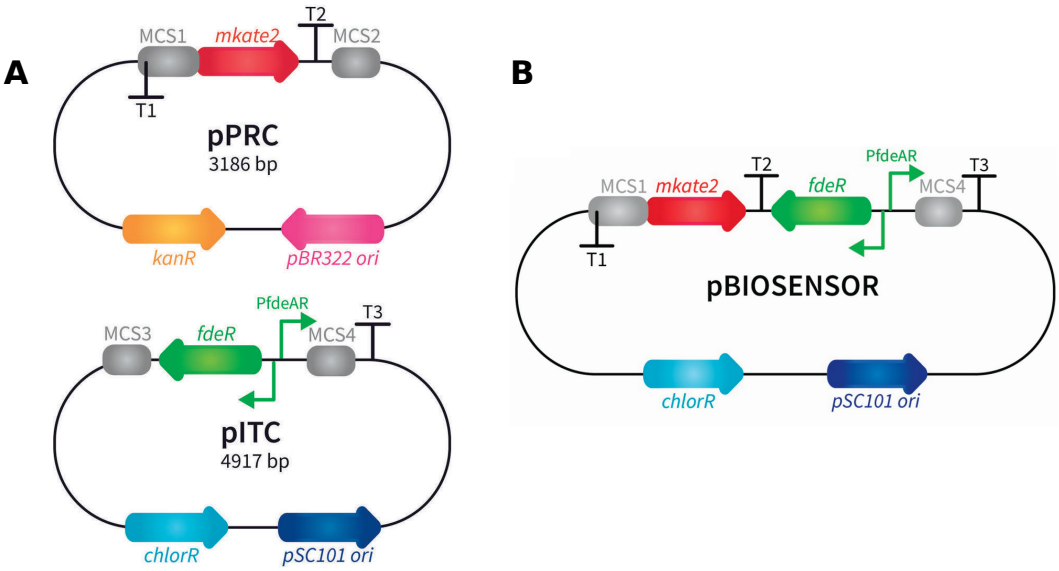

Figure 5.13: Overview of the plasmid systems used during this thesis project. **(A)** The two-plasmid system used for the *in vivo* characterisation of biosensors consists of a promoter reporter construct (pPRC) and inducible transcription factor construct (pITC). The studied promoter sequence is inserted in multiple cloning site (MCS) 1, the transcription factor gene is inserted in MCS4. **(B)** The single plasmid pBiosensor constructs (pBS for short) can be constructed from the pPRC and pITC constructs by inserting the module encompassing *mkate2* in pPRC in MCS3 on pITC. Further details on the used plasmid vectors can be consulted in Table 4.5. Plasmid maps were kindly provided by Drs. ir. Amber Bernauw.

The resulting biosensors were studied during *in vivo* experiments, where the normalised fluorescence (FL/OD_{600}) output in response to different BA concentrations was measured using a fluorescent plate reader (Section 4.2.2.2). The ”full” biosensor strains were accompanied during these experiments with control strains. ”no TF” strains lacking the transcription factor gene (pITC + pPRC-Prom) convey the response of the studied promoters in the absence of transcription factor. Finally, negative controls (NC) were used, harbouring the empty pITC and pPRC plasmids.

For industrial applications, plasmid-based systems are often avoided as they can be lost if the applied selective pressure is inadequate. Instead, the different components are inserted in the genome of the biosensor strain. To simulate these effects during biosensor characterisation, a single-plasmid biosensor construct (pBS) with low copy number can be used (Figure 5.13B). In this case, the copy number of the MRTF gene equals that of the output gene, resembling the situation on the genome. On the downside, pBS constructs provide less flexibility. pBS constructs were constructed for the wild-type (WT) OapR biosensor during this thesis project using restriction enzyme digestion and ligation (see also Section 4.2.2.1). However, their *in vivo* analysis was beyond the scope of this thesis.

5.2.2 OapR-based biosensors

5.2.2.1 Preliminary experiments to determine a suitable β-alanine concentration gradient and naringenin concentration

Based on previous experiments at our lab, a BA concentration gradient expected to span the operational range of a WT OapR biosensor (harbouring pITC-OapR + pPRC-49S) was prepared. In addition, four naringenin concentrations were selected: 30 mg/L, 10 mg/L, 2 mg/L and 0 mg/L. In the absence of naringenin, the BS strain is expected to give rise to similarly low levels of output fluorescence as a no TF strain, as the expression of OapR is not induced in the former. However, leaky expression from the *fdeAR* promoter might give rise to a background concentration of OapR present in the cell, which is undesirable for applications where the OapR expression needs to be tightly controlled. In addition, 2 mg/L naringenin is not expected to induce OapR expression to a large extent. Consequently, low *mkate2* expression levels are expected for these biosensors. In order to be able to descry small differences in output fluorescence between the BS

and no TF strains, the gain factor was increased to 150. The biosensors were studied in minimal medium to exclude the potential interaction of compounds structurally related to BA with OapR (Hanko *et al.*, 2020).

A dose-response curve measured at 30 mg/L naringenin reveals a previously unnoted feature for OapR biosensors: the output fluorescence levels out at about 0.50 mM BA until 2.0 mM BA and increases once again for BA concentrations exceeding 3.5 mM (Figure 5.14A). The first shift in output fluorescence seems to occur at a relatively low BA concentration. Due to the lack of data points in this region as well as for saturating BA concentrations, attempts to fit this dose-response curve to a Hill function failed. Therefore, the dynamic range can not be accurately estimated. Instead, the *fold induction* can be used in analogy with Equation (2.5):

$$Fold = \frac{P_{high} - P_{low}}{P_{low}} \tag{5.1}$$

where P_{high} and P_{low} signify the normalised fluorescence in the presence and absence of effector, respectively. At 2.0 mM BA, a fold induction of 97 was obtained that increased further to 211 for 40 mM BA.

In the presence of 10 mg/L naringenin, a good fit to Hill function could be obtained (Figure 5.14B). The decrease in MRTF concentration drastically changes the shape of the dose-response curve. Although no dose-response parameters were obtained in the presence of 30 mg/L naringenin, the basal output seems to be increased and the maximum increase in expression decreased. This observation is consistent with OapR's role as a dual-function transcription factor in *E. coli*, combining properties of a repressor and activator (see also Section 2.3). The dynamic range only approximates to 1.

At 2 mg/L and 0 mg/L naringenin, no meaningful fits could be obtained (Figure 5.15). Both dose-response curves reveal a larger influence of noise on these data sets. In the presence of 2 mg/L naringenin, the fold induction at 2.0 mM BA and 40 mM BA equals 0.6 and 1.3, respectively (Figure 5.15A). In the absence of naringenin, no induction was observed, as the output fluorescence of the BS and no TF strains overlap at 40 mM BA (Figure 5.15B). The upward trend of the normalised fluorescence output for increasing BA concentrations is most likely OapR-independent, as it is common to both the BS as the no TF strain. Instead, the metabolism of BA might lead to an increase in gene

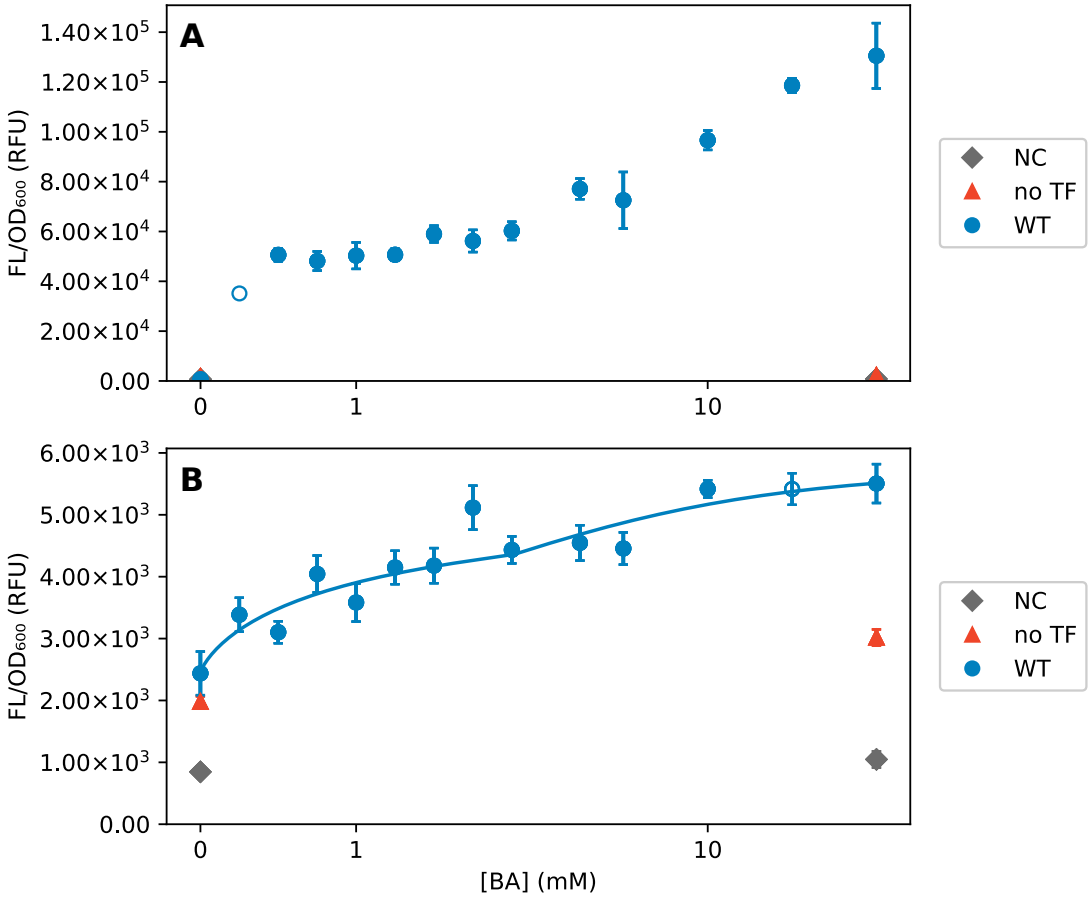

Figure 5.14: Response curve of the OapR/P_{oapTD} biosensor in the presence of 30 mg/L naringenin **(A)** and 10 mg/L naringenin **(B)**. **(A)** The relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 50 h incubation in minimal medium supplemented with different β-alanine (BA) concentrations and 30 mg/L naringenin. **(B)** FL/OD_{600} was measured after 40 h incubation in minimal medium supplemented with the same [BA] gradient and 10 mg/L naringenin. In each case, the gain factor was set to 150. [BA] is plotted on a symmetrical logarithmic scale. Grey diamonds represent the control strain (empty pITC + pPRC), red triangles the biosensor lacking the *oapR* gene (pITC + pPRC-49S) and blue circles the wild-type biosensor (pITC-OapR + pPRC-49S). All error bars represent the mean ± standard error of the mean across four biological replicates, except for the data points plotted as open circles (one biological replicate for the WT strain in the presence of 0.25 mM BA in panel (A) and three biological replicates for the WT strain in the presence of 20 mM BA in panel (B)) as some strains did not grow. In panel (B), the solid line was plotted using the best-fit parameters from a fit to Equation (2.4) (ligand-inducible version, parameters can be consulted in Table II.1. Experimental data for panel (B) was kindly provided by Dr. ir. Indra Bervoets.

expression, among which *mkate2* that is transcribed at a basal level in the absence of OapR.

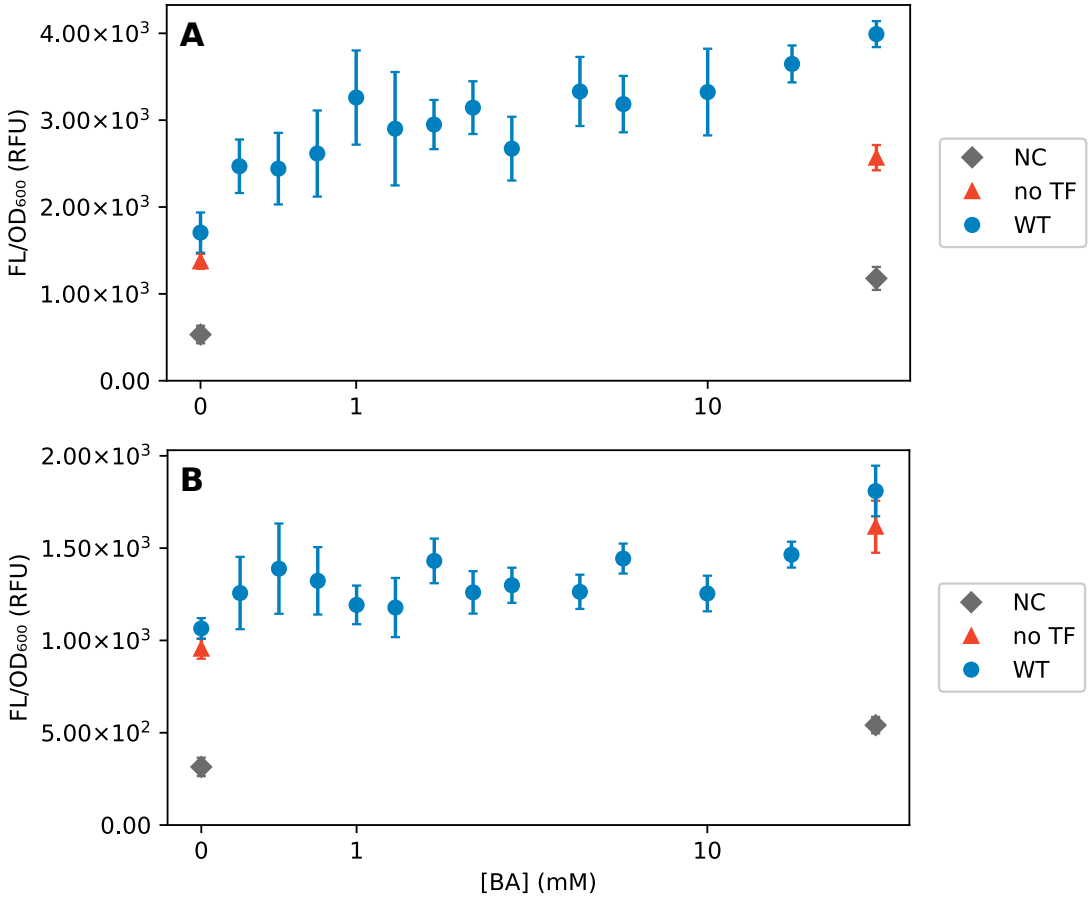

Figure 5.15: Response curve of the OapR/P_{oapTD} biosensor in the presence of 2 mg/L naringenin **(A)** and in the absence of naringenin naringenin **(B)**. In each case, the relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 45 h in minimal medium supplemented with different β-alanine (BA) concentrations. For panel (A), the gain factor was set to 150; for panel (B), the gain factor was adjusted to 140 due to a mistake. [BA] is plotted on a symmetrical logarithmic scale. Grey diamonds represent the control strain (empty pITC + pPRC), red triangles the biosensor lacking the *oapR* gene (pITC + pPRC-49S) and blue circles the wild-type biosensor (pITC-OapR + pPRC-49S). All error bars represent the mean ± standard error of the mean across four biological replicates, except for the NC strain in in panel (A) where three biological replicates were used. Experimental data for panel (B) was kindly provided by Dr. ir. Indra Bervoets.

In conclusion, these preliminary response curves suggest that OapR biosensors have a large dynamic range, provided that *oapR* is expressed to a high level. An interesting two-stage behaviour was observed in the presence of 30 mg/L naringenin, that requires further study. Leaky expression of *oapR* in the absence of naringenin appears unimportant.

Based on these considerations, all further experiments on OapR biosensors applied 30 mg/L naringenin. Based on the determined threshold value in the presence of 10 mg/L (1.3 mM, 95% confidence interval (CI) 0.6 mM - $+\infty$), the BA gradient was refined to include more data points at lower BA concentrations. Besides, we chose to continue with MOPS EZ Rich Defined Medium. This commercially available medium is expected to lead to more reproducible results. By leaving out the medium component containing amino acids and calcium pantothenate, it was intended to reduce the influence of structurally or metabolically related compounds on the OapR-mediated expression of *mkate2*. Finally, the gain was adjusted to 110.

Under these conditions, a good fit to the Hill function could be obtained. The second stage of output fluorescence is less apparent compared to the response curves obtained under previous conditions, but is suggested by the lack-of-fit of the data points between 10 mM and 40 mM BA. The fit also reveals a low basal output (58 RFU (95% CI 46-71 RFU)) and a broad dynamic range ($\mu \approx 135$). Besides, a low threshold is reported (0.27 mM (95% CI 0.21-0.36 mM)).

5.2.2.2 Tuning the transcription-factor binding site

Little is known of the mechanisms of transcriptional regulation by MocR subfamily transcription factors, posing a challenge upon their implementation in whole-cell biosensors. For instance, OapR has not yet been subjected to *in vitro* characterisation to date. Nevertheless, its TFBSs on the genome of *C. necator* H16 have been predicted through comparative genomics, in addition to ten additional genomes of *Betaproteobacteria*. Three binding sites were predicted, of which two direct repeats (D1 and D2) and one inverted repeat (I3) (Figure 5.17A). D1 and D2 are separated by about 11 nucleotides, corresponding to one full turn of B-DNA (10.5 nucleotides). D2 and I3 are separated by 15 nucleotides. All TFBSs are located upstream of the predicted core promoter sequence, which could explain why OapR primariliy functions as an activator, as this region is pri-

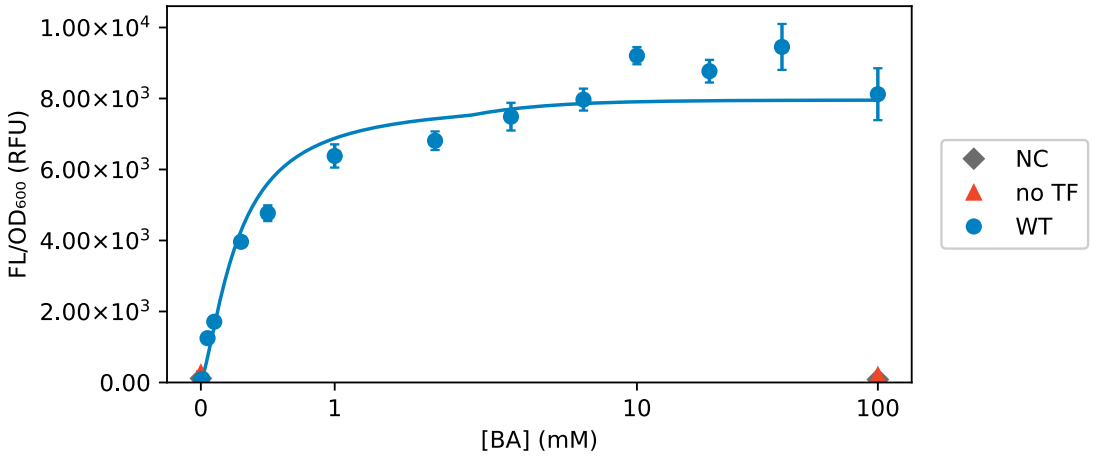

Figure 5.16: Response curve and fitted Hill function of the OapR/P_{oapTD} biosensor in the presence of 30 mg/L naringenin. The relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 35 h incubation in rich medium supplemented with different β-alanine (BA) concentrations. The gain factor was set to 110. [BA] is plotted on a symmetrical logarithmic scale. Grey diamonds represent the control strain (empty pITC + pPRC), red triangles the biosensor lacking the *oapR* gene (pITC + pPRC-49S) and blue circles the wild-type biosensor (pITC-OapR + pPRC-49S). All error bars represent the mean ± standard error of the mean across four biological replicates. The solid line was plotted using the best-fit parameters from a fit to Equation (2.4) (ligand-inducible version, parameters can be consulted in Table II.1.

marily located by activators in *E. coli*. The predicted I3 site shows a single nucleotide overlap to the predicted -35 sequence, a position most suited for class I activation. D1 and D2 reside further from the core promoter sequence and could be involved in class II activation (van Hijum *et al.*, 2009; Suvorova & Rodionov, 2016).

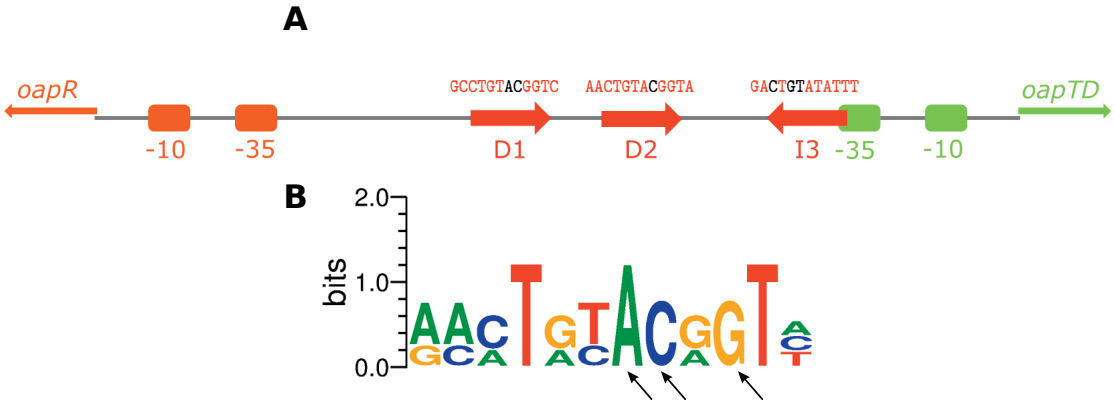

Figure 5.17: Predicted OapR DNA binding sites in the genome of *C. necator* H16. **(A)** Schematic overview of the *oapR*/*oapTD* intergenic region. The predicted transcription-factor binding sites (TFBSs) are indicated in red (D1, D2 and D3) and their native sequence is displayed. Nucleotides indicated in red were separately targeted for mutation. The core promoter sequences of *oapR* (orange) and *oapTD* (green) were predicted using `BPROM` (Softberry Inc., Mount Kisco, NY, USA). **(B)** Sequence logo showing the conservation of nucleotides across the D1, D2 and I3 sites of *C. necator* H16. Three positions are indicated (7A, 8C and 10G) that are conserved across all sites and were separately targeted targeted for mutation. TFBS predictions were taken from (Novichkov *et al.*, 2013). Sequence logo was created using WebLogo with the average GC content of *C. necator* H16 (66.3%) (Schneider & Stephens, 1990; Crooks *et al.*, 2004).

Six single-nucleotide mutations were introduced in the pPRC-49S plasmid through site-directed mutagenesis. In order to maximally affect OapR binding, nucleotides that were present in all three TFBSs predicted for *C. necator* H16 and sharing high-level conservation in the same type of site (that is, D1, D2 or I3) across *Betaproteobacteria* were targeted for mutation (Suvorova & Rodionov, 2016) (Figure 5.17B). This led to the mutant biosensors mut1.1 (8C>G substitution in the D1 site), mut1.2 (7A>T in the D1 site), mut2 (8C>G substitution in the D2 site), mut3.1 (10G>C substitution in the D1 site), mut3.2 (8C>G substitution in the D1 site), mut3.3 (7A>T substitution in the D1 site). Note that throughout this thesis, the nucleotide positions of I3 are defined based on the

sequence of the bottom strand in order to parallel the D1 and D2 sites.

All designed mutations could be successfully obtained, with efficiencies ranging from 20% to 75% across different mutations, only slightly lower than reported in the original publication of the applied method (Edelheit *et al.*, 2009). Therefore, all mutated plasmids were co-transformed with pITC-OapR in *E. coli* MG1655 to investigate the effect on the fluorescence output. The normalised fluorescence was measured in the absence of BA, as well as in the presence of low (0.50 mM) and high (20 mM) BA concentrations. In addition, a control strain was included harbouring pITC-OapR and the empty pPRC, whose response was only measured at 0 and 20 mM BA. All mut strains displayed decreased output fluorescence in the presence of BA compared to the control strains (Figure 5.18).

However, during the preparation of this manuscript, it was noted that the pPRC-mut3.1 and pPRC-mut3.3 plasmids exhibited a smaller overlapping peak on the Sanger sequencing chromatograms, located at the position of the designed mutation. In both cases, the overlapping peak corresponded to the WT sequence at that position, suggesting that the obtained pPRC-mut3.1 and pPRC-mut3.3 plasmids were not genotypically pure. To test this hypothesis, a dilution of the original colony harbouring pPRC-mut3.1 was plated out. Three separate colonies were isolated and cultivated overnight and their plasmids were purified and sent for sequencing. This revealed three distinct behaviours: one colony exhibited the designed mutation, another one the WT sequence and the third exhibited the same pattern of overlapping peaks noted before. Therefore, the results corresponding to mut3.1 and mut3.3 should be interpreted with caution, as the original plasmid (pPRC-49S) might be present in these mut strains and the ratio of original to mutated plasmids is unknown. Still, their reduced output fluorescence compared to the WT strain in the presence of BA marks these strains for further investigation (Figure 5.18A).

The inducibility of mut3.2 is most severely affected (Figure 5.18B). It is only induced 0.9-fold at 0.5 mM BA and 1.7-fold at 20 mM BA. Further, mut 1.1, mut 1.2 and mut2 show increased fluorescence at 0 mM BA compared to the WT, suggesting that their basal output might be affected. To study these effects in more detail as well as to determine the effect on θ and n, all mutant strains were studied in the presence of a larger

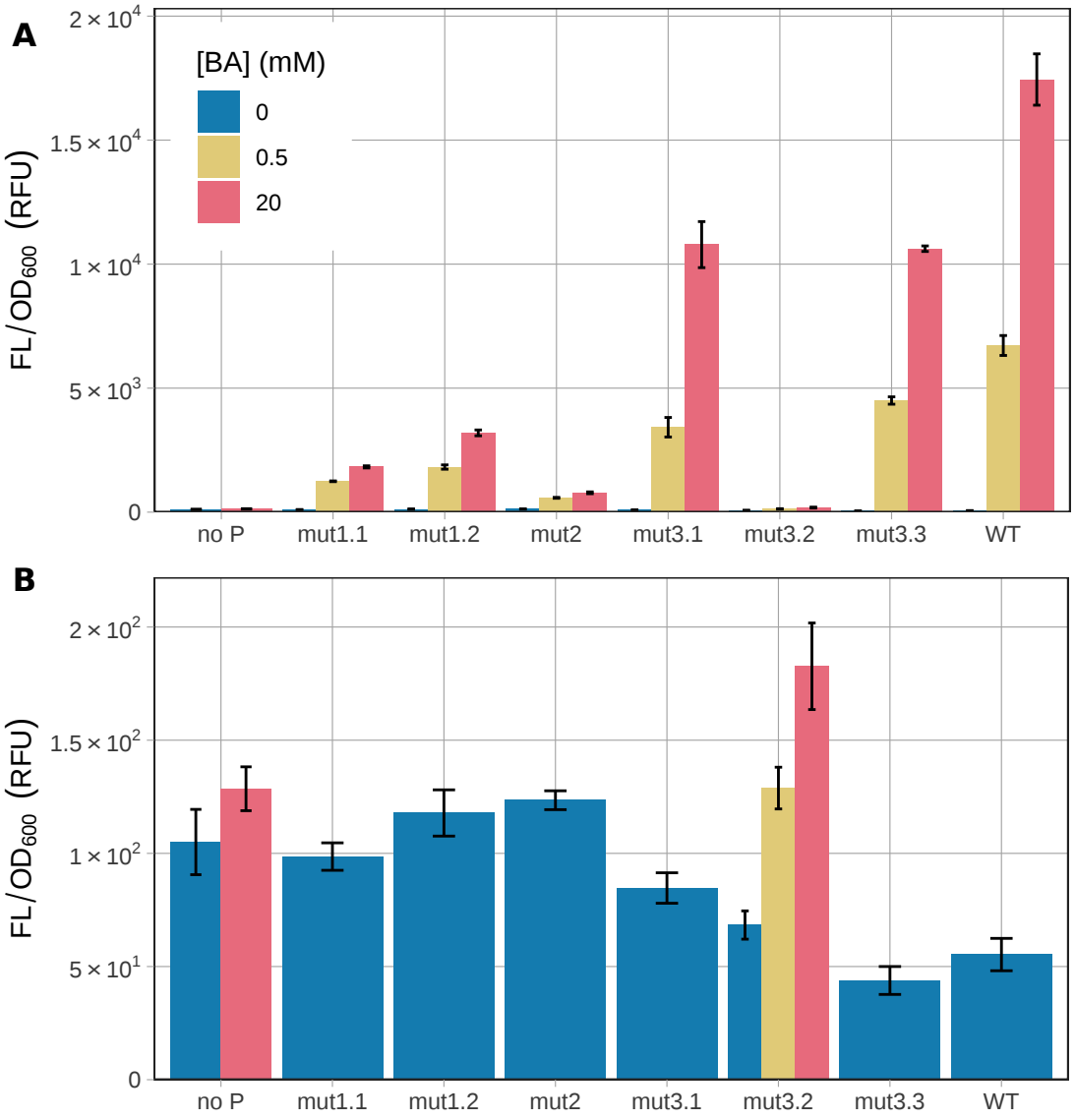

Figure 5.18: Quantitative evaluation of transcription-factor binding site (TFBS) mutations to the *oapTD* promoter. Biosensors harbouring a promoterless reporter gene (pITC-OapR + pPRC), the different TFBS mutations in the *oapTD* promoter region (pITC-OapR + pPRC-mut1.1 up to pPRC-mut3.3) and the wild-type biosensor (pITC-OapR + pPRC-49S) were grown in rich medium supplemented with 0 mM (blue bars), 0.5 mM (yellow bars) and 20 mM β-alanine (BA) (red bars) and 30 mg/L naringenin. The relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 39 h . The gain factor was set to 110. **(A)** Overview of the FL/OD_{600} values across all measured conditions. **(B)** Close-up on the lowest measured FL/OD_{600} values to display the basal output and the limited induction of the no P and mut3.2 strains. Reported values represent the mean ± standard error of the mean across four biological replicates, except for the no P and mut3.3 strains where three biological replicates were used. Experimental data was kindly provided by Dr. ir. Indra Bervoets.

range of BA concentrations.

For both D1 mutants, fits to a Hill function confirm the observed increase in b (mut1.1: 124 RFU (95% CI 95-154 RFU), mut1.2: 109 RFU (95% CI 94-124 RFU)) and decrease in a (mut1.1: 1228 RFU (95% CI 1104-1360 RFU), mut1.2: 1456 RFU (95% CI 1340-1583 RFU)) with respect to the WT (Figure 5.19). These effects translate to a decreased dynamic range ($\mu = 10$ and 13, respectively). Besides, the threshold is decreased for mut1.1 ($\theta = 0.07$ mM (95% CI 0.05-0.10 mM)). The threshold of mut1.2 lies in the same range as for the WT. This leads to a smaller operational range for the mut1.1 biosensor ($\approx$ 0-1 mM).

The mut2 biosensor also shows an increased basal output (124 RFU (95% CI 103-145 RFU)) (Figure 5.20). The value of a is even further reduced to 408 RFU (95% CI 348-487 RFU), reducing the dynamic range to 3.3. The data is fit best with a small value of $n = 0.8$ (95% CI 0.6-1.2), highlighting the shallow nature of the response.

For the mut3.1 biosensor, b and a values appear largely unaffected and its dynamic range is also close to the value calculated for the WT (117 and 135, respectively) (Figure 5.21A). However, it has a more shallow appearance than the WT dose-response curve, which is also reflected by their increased θ and higher n. The high degree of dissimilarity among biological replicates, reflected by the large standard error, could be due to the genotypic impurity of the pPRC-mut3.1 plasmid prep discussed earlier. Indeed, the transformation step of the purified mut plasmids in *E. coli* MG1655 that precedes the plate reader experiment could resolve the mutated and unmutated plasmids in distinct colonies and distinct replicates. This dissimilarity was also noted for the mut3.3 biosensor. However, in the latter case, this effect was caused by a single biological replicate. From 0.05 mM BA on, it consistently outfluoresced the other replicates, reaching sixfold higher FL/OD_{600} at 40.0 mM BA. Therefore, it was considered an outlier and removed from the data set. A fit of the resulting data set to a Hill function revealed a narrow dynamic range of 12, mainly due to a which was reduced with a factor 9 compared to the WT (Figure 5.21C). Besides, the fitting suggested a decreased θ.

Finally, the response curve of the mut3.2 biosensor evidenced a significantly affected a value ($a = 139$ RFU (95% CI 83-247 RFU)) (Figure 5.21B). In fact, the fluorescent output does not exceed that of the no TF strain. This suggests that the introduced muta-

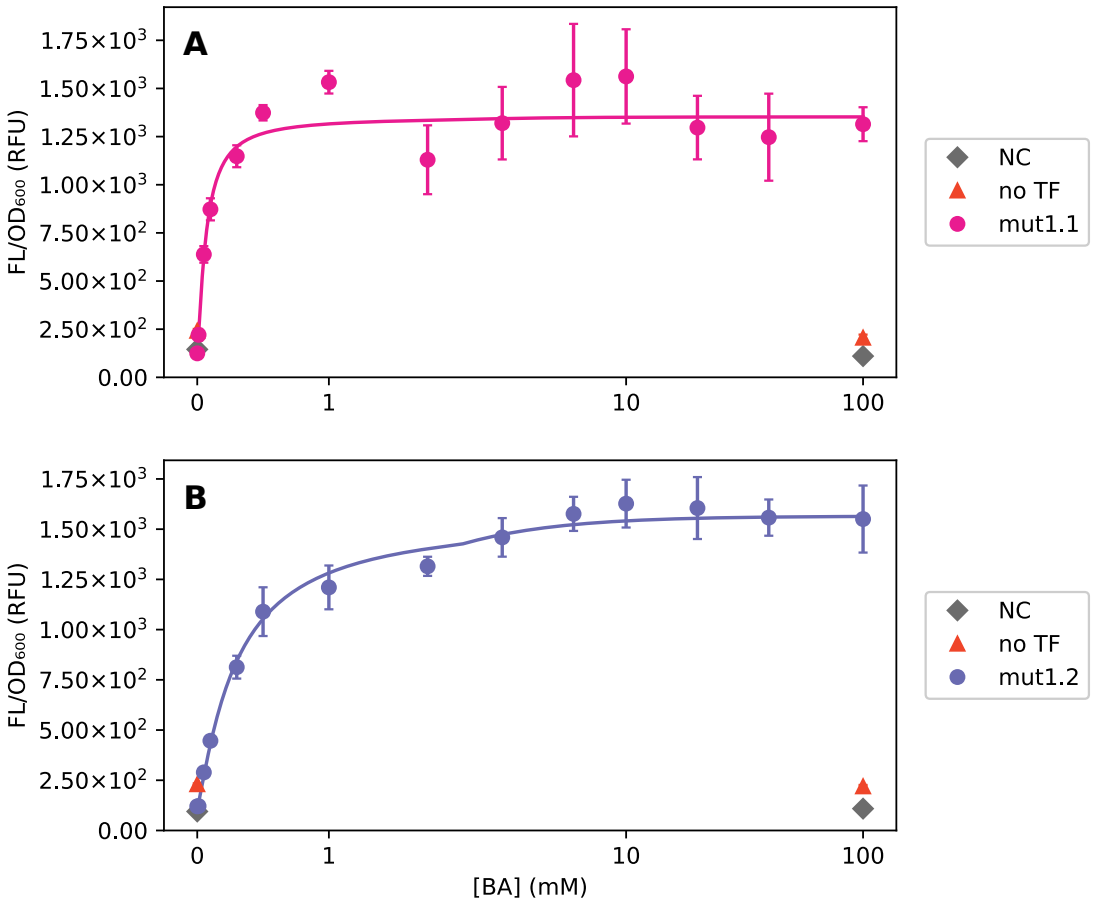

Figure 5.19: Response curve and fitted Hill functions of the mut1 biosensors in the presence of 30 mg/L naringenin. **(A)** The relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 40 h incubation in rich medium supplemented with different β-alanine (BA) concentrations. Grey diamonds represent the control strain (empty pITC + pPRC), red triangles the biosensor lacking the *oapR* gene (pITC + pPRC-mut1.1) and violet circles the full biosensor (pITC-OapR + pPRC-mut1.1). **(B)** FL/OD_{600} was measured after 30 h incubation in rich medium supplemented with the same [BA] gradient. The same encoding as for panel (A) was used, with red triangles representing a strain carrying pITC + pPRC-mut1.2 and purple circles pITC-OapR + pPRC-mut1.2. In each case, the gain factor was set to 110. [BA] is plotted on a symmetrical logarithmic scale. All error bars represent the mean ± standard error of the mean across four biological replicates, except for the no TF strain in panel (A) where three biological replicates were used. The solid lines were plotted using the best-fit parameters from a fit to Equation (2.4) (ligand-inducible version, parameters can be consulted in Table II.1.

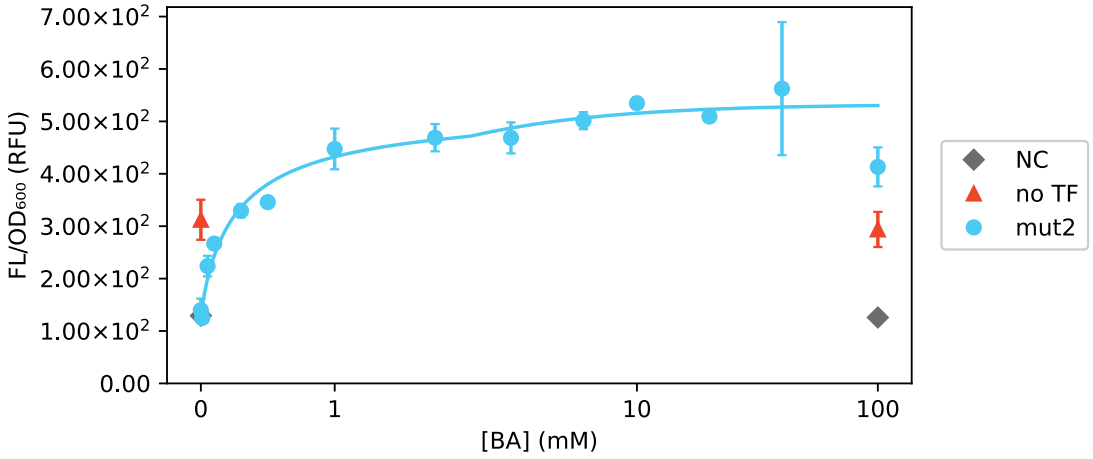

Figure 5.20: Response curve and fitted Hill function of the mut2 biosensor in the presence of 30 mg/L naringenin. The relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 35 h incubation in rich medium supplemented with different β-alanine (BA) concentrations. The gain factor was set to 110. [BA] is plotted on a symmetrical logarithmic scale. Grey diamonds represent the control strain (empty pITC + pPRC), red triangles the biosensor lacking the *oapR* gene (pITC + pPRC-mut2) and blue circles the full biosensor (pITC-OapR + pPRC-mut2). All error bars represent the mean ± standard error of the mean across three biological replicates, except for the NC and no TF strains where four biological replicates were used. The solid lines were plotted using the best-fit parameters from a fit to Equation (2.4) (ligand-inducible version, parameters can be consulted in Table II.1.

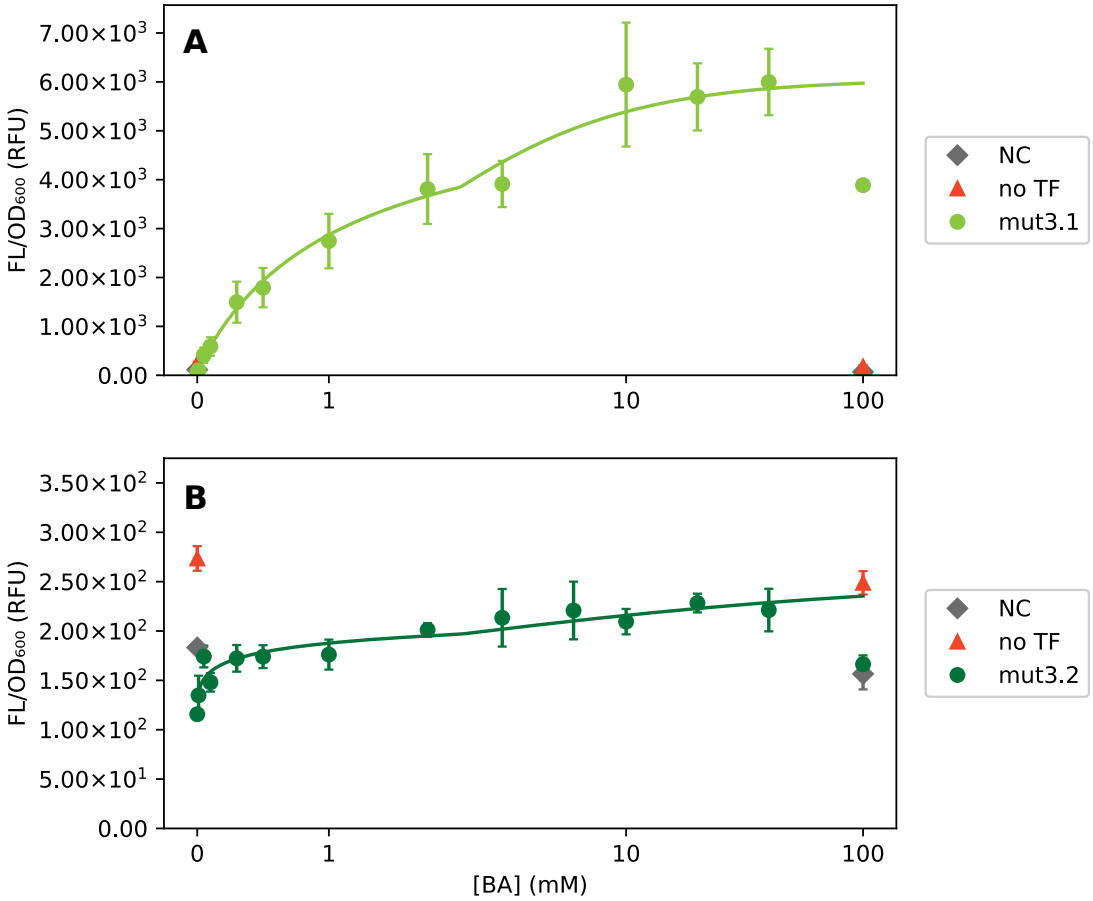

Figure 5.21: Response curve and fitted Hill functions of the mut3 biosensors in the presence of 30 mg/L naringenin. **(A)** The relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 32 h incubation in rich medium supplemented with different β-alanine (BA) concentrations. Grey diamonds represent the control strain (empty pITC + pPRC), red triangles the biosensor lacking the *oapR* gene (pITC + pPRC-mut3.1) and yellow green circles the full biosensor (pITC-OapR + pPRC-mut3.1). **(B)** FL/OD_{600} was measured after 39 h incubation in rich medium supplemented with the same [BA] gradient. Grey diamonds: pITC + pPRC, red triangles: pITC + pPRC-mut3.2, green circles: pITC-OapR + pPRC-mut3.2. **(C)** FL/OD_{600} was measured after 30 h incubation in rich medium supplemented with the same [BA] gradient. Grey diamonds: pITC + pPRC, red triangles: pITC + pPRC-mut3.2, green circles: pITC-OapR + pPRC-mut3.3. In each case, the gain factor was set to 110. [BA] is plotted on a symmetrical logarithmic scale. All error bars represent the mean ± standard error of the mean across four biological replicates, except for the mut3.3 strain where three biological replicates were used and two in the presence of 2.0 mM BA (indicated by the open circle). The solid lines were plotted using the best-fit parameters from a fit to Equation (2.4) (ligand-inducible version, parameters can be consulted in Table II.1.
(Continued on next page)

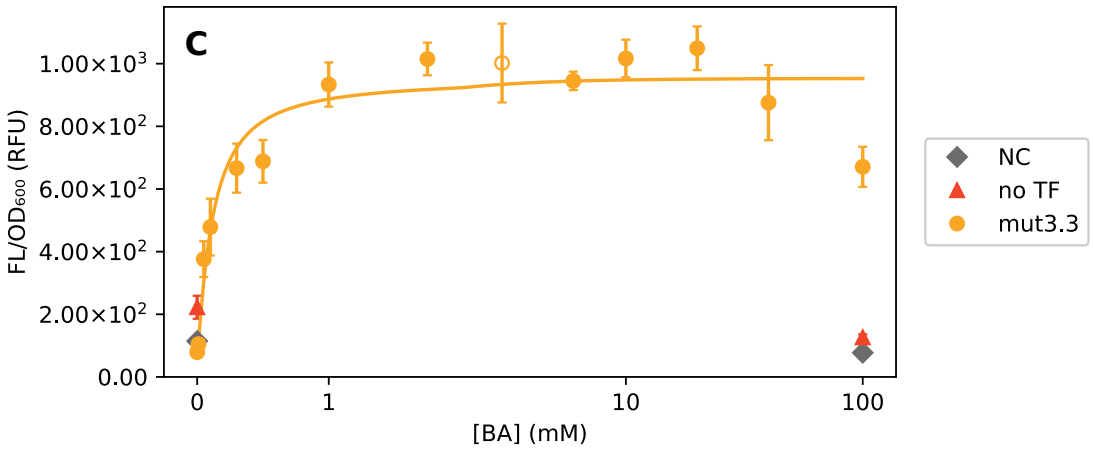

Figure 5.21: Response curve and fitted Hill functions of the mut3 biosensors in the presence of 30 mg/L naringenin. **(A)** The relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 32 h incubation in rich medium supplemented with different β-alanine (BA) concentrations. Grey diamonds represent the control strain (empty pITC + pPRC), red triangles the biosensor lacking the *oapR* gene (pITC + pPRC-mut3.1) and yellow green circles the full biosensor (pITC-OapR + pPRC-mut3.1). **(B)** FL/OD_{600} was measured after 39 h incubation in rich medium supplemented with the same [BA] gradient. Grey diamonds: pITC + pPRC, red triangles: pITC + pPRC-mut3.2, green circles: pITC-OapR + pPRC-mut3.2. **(C)** FL/OD_{600} was measured after 30 h incubation in rich medium supplemented with the same [BA] gradient. Grey diamonds: pITC + pPRC, red triangles: pITC + pPRC-mut3.2, green circles: pITC-OapR + pPRC-mut3.3. In each case, the gain factor was set to 110. [BA] is plotted on a symmetrical logarithmic scale. Error bars represent the mean ± standard error of the mean across four biological replicates, except for the mut3.3 strain where three biological replicates were used and two in the presence of 2.0 mM BA (indicated by the open circle). Note that data for mut3.1 at 6.0 mM BA is missing. The solid lines were plotted using the best-fit parameters from a fit to Equation (2.4) (ligand-inducible version, parameters can be consulted in Table II.1. (Continued)

tion abrogates the ability of OapR to activate transcription in *E. coli*. The basal output is only increased by twofold ($b = 116$ RFU (95% CI 96-136 RFU)) compared to the WT, approximating the dynamic range to 1. At 100 mM BA, the fluorescent output of the biosensor drops to the range of the NC strain. This behaviour can also be observed on other response curves (Figure 5.20A-C) and is more likely to be stress-related than caused by a direct effect on OapR. Indeed, strains grown in the presence of 100 mM BA show a delayed onset of the exponential growth phase regardless of which pPRC construct they carried. An additional control could be performed in the future, where the normalised fluorescence of the no TF strain is measured in the presence of the same BA concentration gradient. The superposition of this response curve to the one belonging to mut3.2 could clarify to which extent the observed induction is due to the action of OapR.

A next step could be to purify OapR and study its binding to the different mutated promoter regions *in vitro*. It would be especially interesting to unravel the detrimental effect of the mutation harboured by mut3.2, as it could provide further insight into the regulation by OapR.

In conclusion, different dose-response characteristics were obtained upon mutation of the predicted TFBSs. In most cases, b and a were simultaneously affected, reducing the dynamic range. The parametrisation of the obtained biosensors will allow them to be rationally selected for different applications.

5.2.3 AHOS_RS02205-based biosensors

The MRTF AHOS_RS02205 from the archaeon "*Acidianus hospitalis*" W1 was studied in *E. coli* in combination with different promoters. For this purpose, pITC-AHOS_RS02205 was used in combination with several pPRC variants (see Table 4.5). Firstly, pPRC-1S contained the promoter of the AHOS_RS02205 structural gene upstream of *mkate2*. *Escherichia coli* whole-cell biosensors containing pITC-AHOS_RS02205 and pPRC-1S plasmids conveyed a downwards trend in the fluorescence output for increasing naringenin concentrations in the absence of BA, indicating that it functions as a repressor in *E coli* (Drs. ir. Amber Bernauw, unpublished data). In the presence of BA, a concentration-dependent derepression was observed (Figure 5.22). In the presence of

10 mg/L naringenin, a relatively high basal level of expression was measured, leading to a modest dynamic range ($\approx$ 1, Figure 5.22A). When the concentration of naringenin was doubled, *b* decreased (3313 RFU (95% CI 3088-3535 RFU) to 2505 RFU (95% CI 2276-2724 RFU)) whereas *a* increased (4718 RFU (95% CI 4301-5162 RFU) to 6463 RFU (95% CI 5822-7148 RFU)), leading to a broader dynamic range ($\approx$ 2). The threshold increased simultaneously (from 0.20 mM (95% CI 0.17-0.24 mM) to 0.20 mM (95% CI 0.23-0.32 mM)). These observations support the predictions of the Mannan model for increasing a_1 values in the repressed-repressor architecture (the maximum increase in transcription factor activity, see Section 5.1.1).

On the genome of "*A. hospitalis*", *AHOS_RS02205* resides in a divergent orientation of the putative aminotransferase gene *AHOS_RS02210*. As feast/famine regulatory proteins (FFRPs) such as AHOS_RS02205 commonly regulate their divergent gene, the regulation of $P_{AHOS\_RS02210}$ by AHOS_RS02205 was assessed (Ziegler & Freddolino, 2021). Increasing the intracellular AHOS_RS02205 concentration in *E. coli* (by increasing the naringenin concentration) resulted in an increase in mKate2 fluorescence (Drs. ir. Amber Bernauw, unpublished data). This suggests that AHOS_RS02205 activates expression from $P_{AHOS\_RS02210}$ in *E. coli*. Consistent with this observation, the normalised fluorescence of the biosensor (pITC-AHOS_RS02205 + pPRC-2S) at 0 mM BA exceeded that of a strain lacking the MRTF gene (pITC + pPRC-2S) (Figure 5.23). Both strains showed overlapping output fluorescence at 10 mM BA, indicating that AHOS_RS02205 functions as a repressed-activator of the *AHOS_RS02210* promoter in *E. coli*. However, the measured signal was very weak, which can be appreciated by comparing the output fluorescence of the NC strain (harbouring empty pITC and pPRC plasmids) with the BS and no TF strains. This counterintuitive result is probably due *transcriptional readthrough* on pPRC, leading to expression of the promoterless *mkate2* in NC. Inserting the (weak) *AHOS_RS02210* promoter upstream of *mkate2* could interfere with this readthrough and reduce its effect on *mkate2* expression.

Still, this biosensor can be considered functional, as *a* significantly differed from zero for both applied naringenin concentrations (95%CIs 643-1184 RFU and 369-629 RFU, respectively). Although significant differences between both experiments were noted for *b* and *a*, these results should be interpreted with caution, due to the low output fluorescence of the BS strain and the differing output signals for the NC strains. More

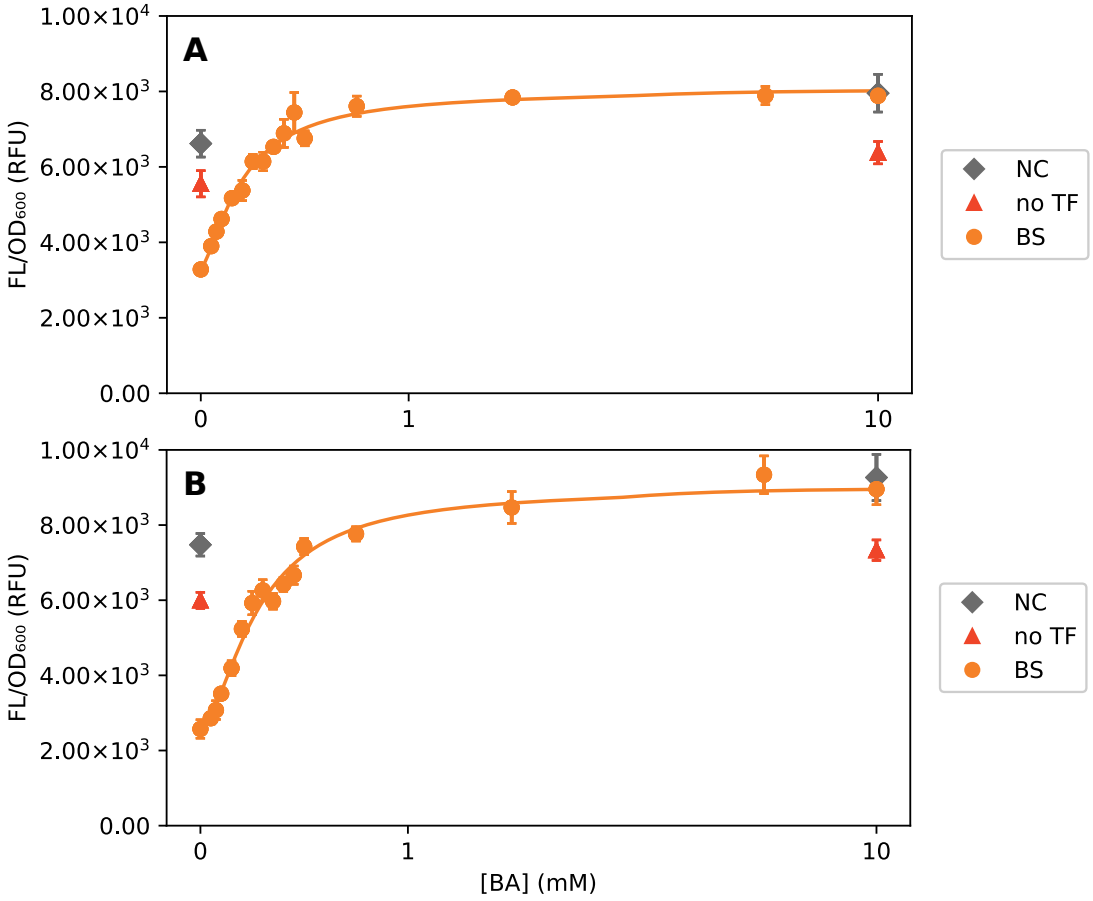

Figure 5.22: Response curves and fitted Hill functions of the AHOS_RS02205/$P_{AHOS\_RS02205}$ biosensor in the presence of 10 mg/L naringenin **(A)** and 20 mg/L naringenin **(B)**. In each case, the relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 30 h incubation in rich medium supplemented with different β-alanine (BA) concentrations. The gain factor was set to 140. [BA] is plotted on a symmetrical logarithmic scale. Grey diamonds represent the control strain (empty pITC + pPRC), red triangles the biosensor lacking the *AHOS_RS02205* gene (pITC + pPRC-1S) and orange circles the full biosensor (pITC-AHOS_RS02205 + pPRC-1S). Error bars represent the mean ± standard error of the mean across four biological replicates. The solid lines were plotted using the best-fit parameters from a fit to Equation (2.4) (ligand-inducible version, parameters can be consulted in Table II.2. Experimental data was kindly provided by Drs. ir. Amber Bernauw.

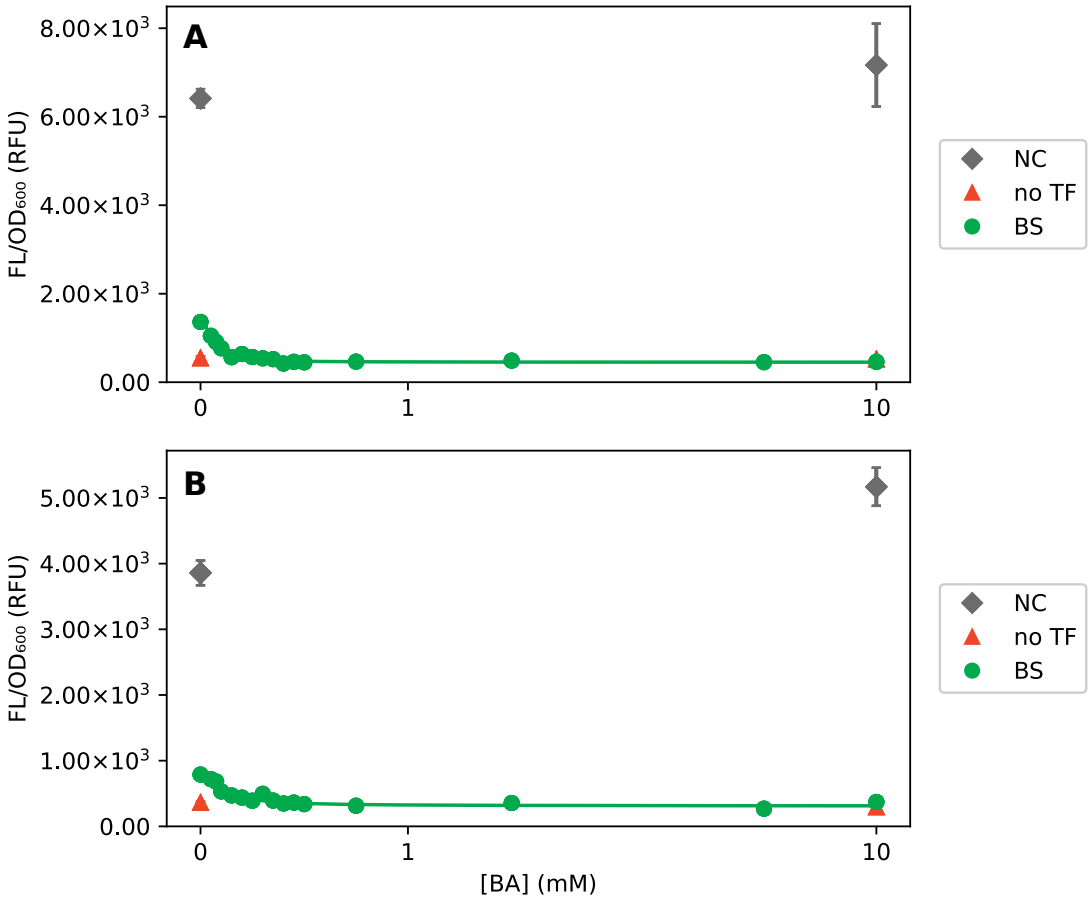

Figure 5.23: Response curves and fitted Hill functions of the AHOS_RS02205/$P_{AHOS\_RS02210}$ biosensor in the presence of 8 mg/L naringenin **(A)** and 20 mg/L naringenin **(B)**. In each case, the relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 30 h incubation in rich medium supplemented with different β-alanine (BA) concentrations. The gain factor was set to 140. [BA] is plotted on a symmetrical logarithmic scale. Grey diamonds represent the control strain (empty pITC + pPRC), red triangles the biosensor lacking the *AHOS_RS02205* gene (pITC + pPRC-2S) and green circles the full biosensor (pITC-AHOS_RS02205 + pPRC-2S). Error bars represent the mean ± standard error of the mean across four biological replicates, except for the NC strain in panel (A) where three biological replicates were used. The solid lines were plotted using the best-fit parameters from a fit to Equation (2.4) (ligand-repressible version, parameters can be consulted in Table II.2. Experimental data was kindly provided by Drs. ir. Amber Bernauw.

differences might be observed upon fortification of the *AHOS_RS02210* promoter.

Given their high level of amino acid sequence identity (66%) and common genomic arrangement (divergent with respect to a putative aminotransferase gene), AHOS_RS02205 and *Sulfolobus acidocaldarius* BarR (Sa-BarR, see also Section 2.3) can be safely assumed to be homologues. Therefore, it was investigated whether AHOS_RS02205 could regulate the expression of the *barR* and *SACI_RS10335* promoters in *E. coli*. For this purpose, plasmid vectors pPRC-54S and pPRC-55S were used, respectively.

Increasing the naringenin concentration reduced the output fluorescence of a strain harbouring pITC-AHOS_RS02205 and pPRC-54S in a concentration-dependent manner (Drs. ir. Amber Bernauw, unpublished data). This indicates that, in contrast with its native regulation in *S. acidocaldarius*, P_{barR} is repressed by AHOS_RS02205. This repression was abrogated upon BA supplementation, meaning that AHOS_RS02205 functions as a repressed-repressor of the *barR* promoter, again distinct to the BA-independent autoregulation of Sa-BarR in *S. acidocaldarius*. The dynamic range approximated to 1.2.

On the other hand, increasing naringenin concentrations revealed a gradual activation of $P_{SACI\_RS10335}$ in *E. coli* biosensors harbouring pITC-AHOS_RS02205 and pPRC-55S (Drs. ir. Amber Bernauw, unpublished data). BA acts in a reciprocal fashion, reducing the output fluorescence with $\mu \approx 5$. This larger dynamic range compared to the AHOS_RS02205/$P_{AHOS\_RS02210}$ biosensor (also functions a repressed-activator), goes at the cost of an increase basal expression level.

In conclusion, AHOS_RS02205 is functional as an MRTF in *E. coli*. Four archaeal promoters were used, resulting in four BA biosensors with differing dose-response parameters. Notably, not only repressor-based regulation was retrieved (Figures 5.22 and 5.24A), but AHOS_RS02205 could also function as an activator (Figures 5.23 and 5.24B), which is exceptional due to the evolutionary distance separating *Bacteria* and *Archaea* (see also Section 2.1.1.1). It should be noted that in each case, native archaeal promoters were used, implicating that the dose-response characteristics of these biosensors could be further improved, for instance, by exchanging the predicted core promoter sequences with endogenous *E. coli* promoter sequences.

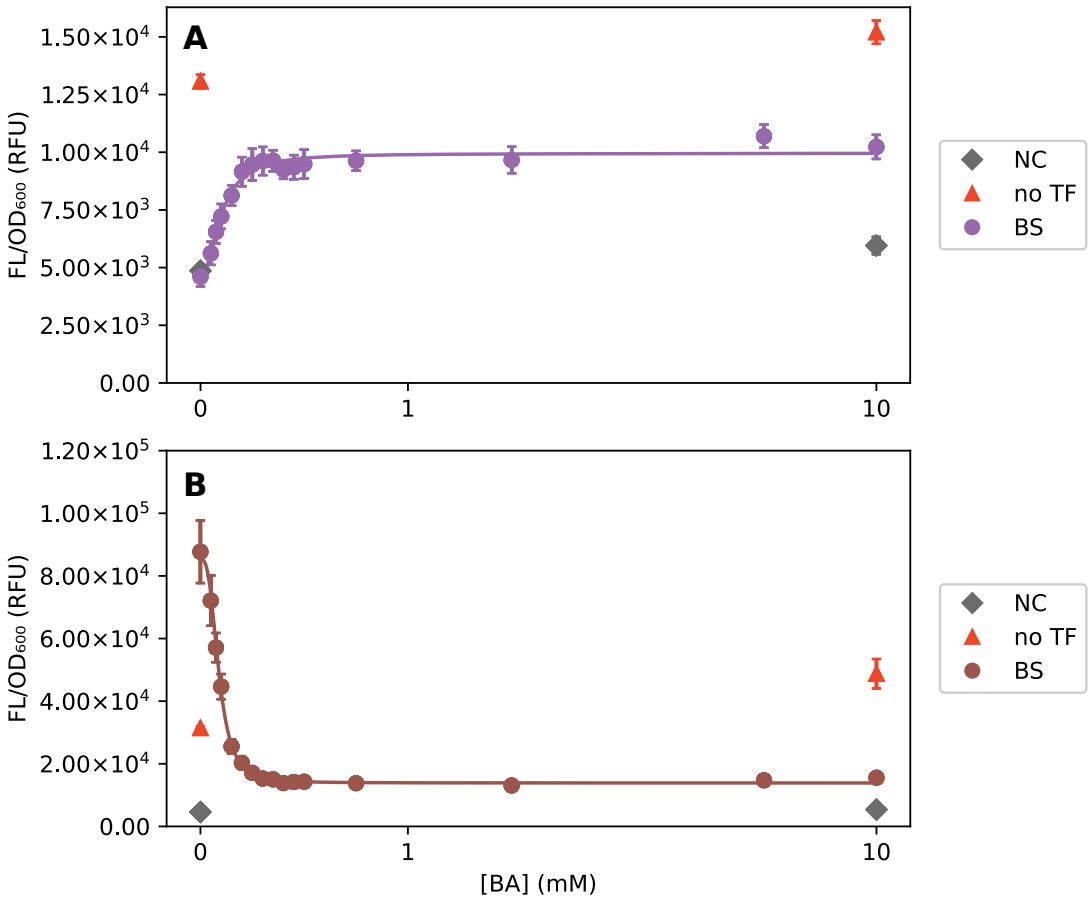

Figure 5.24: Response curve and fitted Hill function of the AHOS_RS02205/P_{barR} biosensor **(A)** and AHOS_RS02205/$P_{SACI\_RS10335}$ biosensor **(B)**. In each case, the relative fluorescence output (FL/OD_{600}, in relative fluorescence units (RFU)) was measured after 30 h incubation in rich medium supplemented with different β-alanine (BA) concentrations and 20 mg/L naringenin. The gain factor was set to 140. [BA] is plotted on a symmetrical logarithmic scale. For both panels, grey diamonds represent the control strain (empty pITC + pPRC). In panel (A), red triangles represent the biosensor lacking the *AHOS_RS02205* gene (pITC + pPRC-54S) and purple circles the full biosensor (pITC-AHOS_RS02205 + pPRC-54S). The solid line was plotted using the best-fit parameters from a fit to Equation (2.4) (ligand-inducible version, parameters can be consulted in Table II.2. The same encoding is used for panel (B), where red triangles represent a strain carrying pITC + pPRC-55S and brown circles pITC-AHOS_RS02205 + pPRC-55S. The solid line was plotted using the best-fit parameters from a fit to Equation (2.4) (ligand-repressible version, parameters can be consulted in Table II.2. Error bars represent the mean ± standard error of the mean across four biological replicates, except for the no TF strains in both panels where three biological replicates were used. Experimental data was kindly provided by Drs. ir. Amber Bernauw.

Conclusion

In the first part of this thesis, an extensive discussion of the Mannan model for biosensor response is provided, a simple phenomenological model that has received significant attention in literature but has never been re-examined by an independent group. Simulations of the repressed-repressor architecture revealed interesting behaviour in response to perturbations in tunable parameters beyond the parameters on which the original publication focused. For instance, tuning a_1 (signifying the maximum increase in transcription factor activity) affected the basal output, threshold and dynamic range of the dose-response curve at the same time, similar to the effect of tuning K_2 (representing the affinity of the transcription factor for its DNA binding site) but simpler to realise in practice by altering the transcription factor expression level. Although the observed behaviour was intuitively satisfying, the predicted effect was in part contradicted by literature, possibly implying that the a_1 parameter alone is not sufficient to model the expression level of the metabolite-responsive transcription factor (MRTF).

In addition, Mannan *et al.* (2017) proposed the promoter dynamic range (i.e. $\mu_2 = \frac{a_2}{b_2}$, where b_2 represents the basal level of promoter expression and a_2 the maximum increase in promoter expression) as a key tunable parameter allowing orthogonal control of the biosensor dynamic range. However, our results qualify this claim by defining limiting conditions, as for lower values of a_1, K_2 or n_2 (representing the sensitivity of MRTF binding to its DNA binding site), the effect on the dynamic range is moderated by knock-on effects on the basal output.

Furthermore, a parameter estimation method (PEM) serving to fit experimental biosensor characterisation data to the Mannan model was implemented in Python. Evaluation of the PEM based on simulated data sets akin to experimental data proved it to be

functional and revealed that the resulting fits could reproduce the overall behaviour of the training data for eight of the nine considered data sets. However, signs of overfitting and parameter unidentifiability suggest a requirement on high-quality experimental data, difficult to obtain in molecular biology. Potential improvements to the PEM were proposed, as its application to real-life data could provide further insight in the validity of the Mannan model.

In the second part of this study, β-alanine-responsive biosensors were studied in *Escherichia coli* due to their enormous potential in metabolic engineering applications, among which the implementation of dynamic control in the β-alanine pathway for the production of 3-hydroxypropionic acid. Biosensors were characterised *in vivo* and parameterised by fitting to a Hill function, an indispensable step that allows their rational selection for different applications.

The dose-response curve of a biosensor based on *Cupriavidus necator* OapR, an MRTF displaying both properties of a repressor and activator, was engineered by the introduction of single-nucleotide substitutions in the predicted TFBSs. This resulted in diverse dose-response characteristics that might serve different applications. For instance, the analog-like behaviour exhibited by the mut3.1 biosensor might be more suitable to differentially regulate the expression of a target gene in response to a larger concentration range, for example in dynamic pathway control. However, it should be examined whether this behaviour persists in a genotypically pure strain.

Interestingly, the 8C>G substitution in the I3 site of the mut3.2 strain decreases the dynamic range more than 100-fold. We hypothesise that the 8C residue is crucial for the action of OapR as an activator in *E. coli*. *In vitro* experiments, such as electrophoretic mobility shift assays in the presence and absence of purified RNA polymerase, are ideally suited to test this hypothesis as well as to further study the relative importance of the other DNA binding sites but exceeded the scope of this thesis.

Finally, four biosensors based on the MRTF AHOS_RS02205 from the archaeon "*Acidianus hospitalis*" were parameterised. Although their dynamic ranges are narrow, the action of AHOS_RS02205 as a repressed-activator on the *AHOS_RS02210* and *SACI_RS10335* promoters in *E. coli* is exceptional and adds two β-alanine-repressible biosensors to the synthetic biology toolbox. These biosensors are potentially useful to downregulate gene

expression in response to β-alanine, for instance, to dynamically divert the flux from endogenous β-alanine catabolism to the heterologous 3-hydroxypropionic acid production pathway in response to intracellular β-alanine accumulation. However, further response curve engineering is required to increase the utility of these biosensors.

Summary

Whole-cell biosensors are a key enabling technology in metabolic engineering. Biosensors based on metabolite-responsive transcription factors (MRTFs) have proven their use in a range of applications, of which the dynamic regulation of metabolic fluxes has received most attention by the research community. However, limited understanding of the tunability of biosensors has made their development labour-intensive and delayed their en masse industrial application. Mathematical modelling efforts have addressed this bottleneck, and one notable study has revealed theoretical constraints for biosensor design and exposed tunable parameters that allow orthogonal control of the biosensor dynamic range and sensing threshold. However, it was based on the well-characterised *lac* repressor and has never applied to new biosensors, that could possess more complex mechanisms not covered by the model.

The first part of this thesis aimed to obtain further insight in the contribution of the different tunable parameters on the biosensor response curve by investigating this phenomenological model. Simulations revealed interesting behaviour in response to perturbations in the a_1 parameter (representing the maximum increase in transcription factor activity) that did not fully correspond to the effect of tuning the expression level of repressed-repressors described in literature, potentially implying that this parameter alone cannot capture that effect. In addition, the proposed orthogonal control of the biosensor dynamic range was nuanced by identifying limiting conditions that moderate this effect. A parameter estimation method was implemented in Python to fit biosensor characterisation data to this model, which will allow to study the effect of these tunable parameters more quantitatively and provide further insight in the validity of the model.

For the second part of this thesis, biosensors responsive to the industrially important

intermediate compound β-alanine were studied in *Escherichia coli*. Six mutated biosensors were constructed by introducing mutations in the predicted DNA binding sites of *Cupriavidus necator* OapR, and their response to β-alanine was quantified *in vivo*. A range of dose-response characteristics emerged with affected dynamic ranges, thresholds and sensitivies across the mutant strains. Notably, a single-nucleotide substitution in the I3 site reduced the dynamic range 100-fold, and the resulting biosensor appeared to have lost the ability to activate the expression from the *oapTD* promoter and thus function solely as a repressor. On the other hand, four biosensors based on the archaeal MRTF AHOS_RS02205 from "*Acidianus hospitalis*" were parameterised. All biosensors were functional in *E. coli*, either as repressed-repressors or repressed-activators, which is exceptional due to the evolutionary distance separating *Bacteria* and *Archaea*.

In conclusion, this thesis project has extracted additional information from a phenomenological model that could aid biosensor response curve engineering and provided the first steps for its practical application to new biosensors. On the other hand, β-alanine-responsive biosensors were engineered, resulting in a set of novel biosensors with potential applications for the production of β-alanine-derived products, such as the promising platform molecule 3-hydroxypropionic acid.

Samenvatting

‘Whole-cell’ biosensoren zijn een sleuteltechnologie binnen de metabole engineering. Hierbij krijgen biosensoren gebaseerd op metaboliet-responsieve transcriptie factoren (MRTFs) veel aandacht nadat ze hun nut bewezen in tal van applicaties door het toepassen van dynamische regulatie van metabole fluxen. Niettemin blijft hun ontwikkeling voor gebruik in industriële toepassingen voorlopig beperkt door het arbeidsintensieve proces dat gepaard gaat met hun ontwikkeling omwille van de beperkte kennis van de afstembaarheid van biosensoren. Verschillende wiskundige modelleringspogingen hebben zich toegespitst op dit specifieke knelpunt en een opmerkelijke studie heeft theoretische beperkingen voor biosensor ontwerp onthuld en stelt hierbij afstembare parameters naar voor die orthogonale controle van het dynamisch bereik en de gevoeligheidsdrempel van de biosensor moeten toelaten. Deze studie is echter gebaseerd op de goed gekarakteriseerde *lac* repressor en werd nog niet toegepast op andere nieuwe biosensoren die volgens veel complexere mechanismen fungeren.

In het eerste deel van de thesis tracht men door het onderzoeken van dit fenomenologisch model verdere inzichten te bekomen in de bijdragen van de verschillende afstembare parameters op de respons curve van een biosensor. Simulaties onthulden interessant gedrag als reactie op verstoringen in de a_1 parameter (maximale toename in transcriptie factor activiteit) die niet volledig overeenkomt met het effect beschreven voor een gerepresseerde-repressor in de literatuur wat impliceert dat deze parameter op zichzelf dit effect niet kan neerzetten. Daarenboven, werd de vooropgestelde orthogonale controle van het dynamisch bereik van de biosensor genuanceerd door het identificeren van limiterende condities die dit effect matigen. Een parameterschattingsmethode werd geïmplementeerd in Python om biosensor karakteriseringsdata te fitten op het model op-

dat het effect van deze afstembare parameters kwantitatief bestudeerd kon worden en er verdere inzichten verworven worden voor de validatie van het model.

Voor het tweede deel van de thesis werden biosensoren, responsief aan de industrieel belangrijke verbinding β-alanine, bestudeerd in *Escherichia coli*. Zes biosensor mutanten werden gebouwd door het introduceren van mutaties in de voorspelde DNA-binding sites van OapR van *Cupriavidus necator* en hun respons op toevoeging van β-alanine werd kwantitatief opgevolgd in vivo. Een reeks van dosis-respons curves werden bekomen met gewijzigd dynamisch bereik, drempelwaarden en gevoeligheden over de verschillende gemuteerde stammen. Opmerkelijk, een enkele nucleotide substitutie in de I3 site reduceerde het dynamisch bereik met factor 100 en de resulterende biosensor lijkt het vermogen tot activatie vanaf de *oapTD* promoter verloren te hebben waardoor deze enkel nog als repressor kan fungeren. Anderzijds werden vier biosensoren gebaseerd op de archaeale MRTF AHOS_RS02205 van "*Acidianus hospitalis*" geparametriseerd. Alle biosensoren bleken functioneel in *E. coli*, als gerepresseerde-repressors of gerepresseerde-activatoren, ondanks de evolutionaire afstand tussen *Bacteria* en *Archaea*.

Tot slot leverde dit project bijkomende informatie startend van een fenomenologisch model die zal bijdragen aan toekomstige biosensor respons curve ontwikkeling en de eerste stappen werden ondernomen voor de praktische applicatie ervan op nieuwe biosensoren. Anderzijds werden verschillende β-alanine-responsieve biosensoren ontworpen resulterend in een set van nieuwe biosensoren die toegepast kunnen worden voor de productie van verbindingen afkomstig van β-alanine zoals het veelbelovende platform molecule 3-hydroxypropionzuur.

Bibliography

Ajikumar, P. K., Xiao, W.-H., Tyo, K. E. J., Wang, Y., Simeon, F., Leonard, E., Mucha, O., Phon, T. H., Pfeifer, B. and Stephanopoulos, G. (2010). Isoprenoid pathway optimization for taxol precursor overproduction in Escherichia coli. *Science* **330**, 70–74.

Alberts, A. and Vagelos, P. R. (1966). Acyl carrier protein: Studies of acyl carrier protein and coenzyme A in Escherichia coli pantothenate or β-alanine auxotrophs. *Journal of Biological Chemistry* **241**, 5201–5204.

Alcántara-Díaz, D., Breña-Valle, M. and Serment-Guerrero, J. (2004). Divergent adaptation of Escherichia coli to cyclic ultraviolet light exposures. *Mutagenesis* **19**, 349–354.

Altuvia, S., Kornitzer, D., Teff, D. and Oppenheim, A. B. (1989). Alternative mRNA structures of the cIII gene of bacteriophage λ determine the rate of its translation initiation. *Journal of Molecular Biology* **210**(2), 265–280.

Amidani, D., Tramonti, A., Canosa, A. V., Campanini, B., Maggi, S., Milano, T., di Salvo, M. L., Pascarella, S., Contestabile, R., Bettati, S. and Rivetti, C. (2017). Study of DNA binding and bending by Bacillus subtilis GabR, a PLP-dependent transcription factor. *Biochimica et Biophysica Acta - General Subjects* **1861**, 3474–3489.

Anashkin, V. A., Aksenova, V. A., Salminen, A., Lahti, R. and Baykov, A. A. (2019). Cooperativity in catalysis by canonical family II pyrophosphatases. *Biochemical and Biophysical Research Communications* **517**, 266–271.

Anderson, J. J., Quay, S. C. and Oxender, D. L. (1976). Mapping of two loci affecting the regulation of branched-chain amino acid transport in Escherichia coli K-12. *Journal of Bacteriology* **126**, 80–90.

Andrews, S. S. and Bray, D. (2004). Stochastic simulation of chemical reactions with spatial resolution and single molecule detail. *Physical Biology* **1**, 137.

Anthony, J. R., Anthony, L. C., Nowroozi, F., Kwon, G., Newman, J. D. and Keasling, J. D. (2009). Optimization of the mevalonate-based isoprenoid biosynthetic pathway

in Escherichia coli for production of the anti-malarial drug precursor amorpha-4,11-diene. *Metabolic Engineering* **11**, 13–19.

Ashyraliyev, M., Fomekong-Nanfack, Y., Kaandorp, J. A. and Blom, J. G. (2009). Systems biology: parameter estimation for biochemical models. *The FEBS Journal* **276**, 886–902.

Belitsky, B. R. and Sonenshein, A. L. (2002). GabR, a member of a novel protein family, regulates the utilization of γ-aminobutyrate in Bacillus subtilis. *Molecular Microbiology* **45**, 569–583.

Bernauw, A. J., De Kock, V. and Bervoets, I. (2022). In vivo screening method for the identification and characterization of prokaryotic, metabolite-responsive transcription factors. In: Prokaryotic Gene Regulation (E. Peeters and I. Bervoets, eds). Humana Press. New York, NY, USA. pp. 113–141.

Bervoets, I. and Charlier, D. (2019). Diversity, versatility and complexity of bacterial gene regulation mechanisms: opportunities and drawbacks for applications in synthetic biology. *FEMS Microbiology Reviews* **43**, 304–339.

Bhagwat, S. P., Rice, M. R., Matthews, R. G. and Blumenthal, R. M. (1997). Use of an inducible regulatory protein to identify members of a regulon: application to the regulon controlled by the leucine-responsive regulatory protein (Lrp) in Escherichia coli. *Journal of Bacteriology* **179**, 6254–6263.

Binder, R., Horowitz, J. A., Basilion, J. P., Koeller, D. M., Klausner, R. D. and Harford, J. B. (1994). Evidence that the pathway of transferrin receptor mRNA degradation involves an endonucleolytic cleavage within the 3′ UTR and does not involve poly(A) tail shortening. *The EMBO Journal* **13**, 1969–1980.

Binder, S., Schendzielorz, G., Stäbler, N., Krumbach, K., Hoffmann, K., Bott, M. and Eggeling, L. (2012). A high-throughput approach to identify genomic variants of bacterial metabolite producers at the single-cell level. *Genome Biology* **13**, R40.

Boada, Y., Vignoni, A., Picó, J. and Carbonell, P. (2020). Extended metabolic biosensor design for dynamic pathway regulation of cell factories. *iScience* **23**, 101305.

Borodina, I., Kildegaard, K. R., Jensen, N. B., Blicher, T. H., Maury, J., Sherstyk, S., Schneider, K., Lamosa, P., Herrgård, M. J., Rosenstand, I., Öberg, F., Forster, J. and Nielsen, J. (2015). Establishing a synthetic pathway for high-level production of 3-hydroxypropionic acid in Saccharomyces cerevisiae via β-alanine. *Metabolic Engineering* **27**, 57–64.

Breaker, R. R. (2011). Prospects for riboswitch discovery and analysis. *Molecular Cell* **43**, 867–879.

Briegel, A., Ortega, D. R., Tocheva, E. I., Wuichet, K., Li, Z., Chen, S., Müller, A., Iancu, C. V., Murphy, G. E., Dobro, M. J., Zhulin, I. B. and Jensen, G. J. (2009). Universal architecture of bacterial chemoreceptor arrays. *Proceedings of the National Academy of Sciences of the United States of America* **106**, 17181–17186.

Brown, B. D., Zipkin, I. D. and Harland, R. M. (1993). Sequence-specific endonucleolytic cleavage and protection of mRNA in Xenopus and Drosophila. *Genes & Development* **7**, 1620–1631.

Burgard, A. P., Pharkya, P. and Maranas, C. D. (2003). Optknock: a bilevel programming framework for identifying gene knockout strategies for microbial strain optimization. *Biotechnology and Bioengineering* **84**, 647–657.

Çakar, Z. P., Seker, U. O., Tamerler, C., Sonderegger, M. and Sauer, U. (2005). Evolutionary engineering of multiple-stress resistant Saccharomyces cerevisiae. *FEMS Yeast Research* **5**, 569–578.

Calvo, J. M. and Matthews, R. G. (1994). The leucine-responsive regulatory protein, a global regulator of metabolism in Escherichia coli. *Microbiological Reviews* **58**, 466–490.

Campbell, L. L. (1960). Reductive degradation of pyrimidines: enzymatic conversion of N-carbamyl-β-alanine to β-alanine, carbon dioxide, and ammonia. *Journal of Biological Chemistry* **235**, 2375–2378.

Carmany, D. O., Hollingsworth, K. and McCleary, W. R. (2003). Genetic and biochemical studies of phosphatase activity of PhoR. *Journal of Bacteriology* **185**, 1112–1115.

Carpenter, A. C., Paulsen, I. T. and Williams, T. C. (2018). Blueprints for biosensors: design, limitations, and applications. *Genes* **9**, 375.

Castaño-Cerezo, S., Fournié, M., Urban, P., Faulon, J.-L. and Truan, G. (2020). Development of a biosensor for detection of benzoic acid derivatives in Saccharomyces cerevisiae. *Frontiers in Bioengineering and Biotechnology* **7**, 372.

Ceres, P., Trausch, J. J. and Batey, R. T. (2013). Engineering modular 'ON' RNA switches using biological components. *Nucleic Acids Research* **41**(22), 10449–10461.

Charlier, D., Roovers, M., Gigot, D., Huysveld, N., Piérard, A. and Glansdorff, N. (1993). Integration Host Factor (IHF) modulates the expression of the pyrimidine-specific promoter of the carAB operons of Escherichia coli K12 and Salmonella typhimurium LT2. *Molecular and General Genetics* **237**, 273–286.

Chaurasia, A., Mohammed, N., Feki Tounsi, M. and Trabelsi, H. (2020). Microbial indicators and biosensors for bioremediation. In: Bioremediation of Pollutants (V. C. Pandey and V. Singh, eds). Elsevier. Amsterdam, The Netherlands. pp. 313–331.

Chen, S., Rosner, M. H. and Calvo, J. M. (2001). Leucine-regulated self-association of leucine-responsive regulatory protein (Lrp) from Escherichia coli. *Journal of Molecular Biology* **312**, 625–635.

Chen, X., Yang, X., Shen, Y., Hou, J. and Bao, X. (2018). Screening phosphorylation site mutations in yeast acetyl-CoA carboxylase using malonyl-CoA sensor to improve malonyl-CoA-derived product. *Frontiers in Microbiology* **9**, 47.

Cordone, A., Mauriello, E. M. F., Pickard, D. J., Dougan, G., De Felice, M. and Ricca, E. (2005). The lrp gene and its role in type I fimbriation in Citrobacter rodentium. *Journal of Bacteriology* **187**, 7009–7017.

Cordova, L. T., Lu, J., Cipolla, R. M., Sandoval, N. R., Long, C. P. and Antoniewicz, M. R. (2016). Co-utilization of glucose and xylose by evolved Thermus thermophilus LC113 strain elucidated by 13C metabolic flux analysis and whole genome sequencing. *Metabolic Engineering* **37**, 63–71.

Cox III, R. S., Surette, M. G. and Elowitz, M. B. (2007). Programming gene expression with combinatorial promoters. *Molecular Systems Biology* **3**, 145.

Crasnier, M., Dumay, V. and Danchin, A. (1994). The catalytic domain of Escherichia coli K-12 adenylate cyclase as revealed by deletion analysis of the cya gene. *Molecular and General Genetics* **243**, 409–416.

Cronan Jr, J. E. (1980). β-Alanine synthesis in Escherichia coli. *Journal of Bacteriology* **141**, 1291–1297.

Crooks, G. E., Hon, G., Chandonia, J. M. and Brenner, S. E. (2004). WebLogo: a sequence logo generator. *Genome Research* **14**, 1188–1190.

Cumming, G., Fidler, F. and Vaux, D. L. (2007). Error bars in experimental biology. *Journal of Cell Biology* **177**, 7–11.

Curran, K. A., Leavitt, J. M., Karim, A. S. and Alper, H. S. (2013). Metabolic engineering of muconic acid production in Saccharomyces cerevisiae. *Metabolic Engineering* **15**, 55–66.

Daley, D. O., Rapp, M., Granseth, E., Melén, K., Drew, D. and Von Heijne, G. (2005). Global topology analysis of the Escherichia coli inner membrane proteome. *Science* **308**, 1321–1323.

Dalwadi, M. P. and King, J. R. (2020). An asymptotic analysis of the malonyl-CoA route to 3-hydroxypropionic acid in genetically engineered microbes. *Bulletin of Mathematical Biology* **82**, 36.

David, F., Nielsen, J. and Siewers, V. (2016). Flux control at the malonyl-CoA node through hierarchical dynamic pathway regulation in Saccharomyces cerevisiae. *ACS Synthetic Biology* **5**, 224–233.

de los Rios, S. and Perona, J. J. (2007). Structure of the Escherichia coli Leucine-responsive Regulatory Protein Lrp reveals a novel octameric assembly. *Journal of Molecular Biology* **366**, 1589–1602.

De Paepe, B., Maertens, J., Vanholme, B. and De Mey, M. (2018). Modularization and response curve engineering of a naringenin-responsive transcriptional biosensor. *ACS Synthetic Biology* **7**, 1303–1314.

De Paepe, B., Peters, G., Coussement, P., Maertens, J. and De Mey, M. (2017). Tailor-made transcriptional biosensors for optimizing microbial cell factories. *Journal of Industrial Microbiology and Biotechnology* **44**, 623–645.

Delvigne, F., Zune, Q., Lara, A. R., Al-Soud, W. and Sørensen, S. J. (2014). Metabolic variability in bioprocessing: implications of microbial phenotypic heterogeneity. *Trends in Biotechnology* **32**, 608–616.

Dietrich, J. A., McKee, A. E. and Keasling, J. D. (2010). High-throughput metabolic engineering: advances in small-molecule screening and selection. *Annual Review of Biochemistry* **79**, 563–590.

Dray, K. E., Muldoon, J. J., Mangan, N. M., Bagheri, N. and Leonard, J. N. (2022). GAMES: a dynamic model development workflow for rigorous characterization of synthetic genetic systems. *ACS Synthetic Biology* **11**, 1009–1029.

Edayathumangalam, R., Wu, R., Garcia, R., Wang, Y., Wang, W., Kreinbring, C. A., Bach, A., Liao, J., Stone, T. A., Terwilliger, T. C., Hoang, Q. Q., Belitsky, B. R., Petsko, G. A., Ringe, D. and Liu, D. (2013). Crystal structure of Bacillus subtilis GabR, an autorepressor and transcriptional activator of gabT. *Proceedings of the National Academy of Sciences of the United States of America* **110**, 17820–17825.

Edelheit, O., Hanukoglu, A. and Hanukoglu, I. (2009). Simple and efficient site-directed mutagenesis using two single-primer reactions in parallel to generate mutants for protein structure-function studies. *BMC Biotechnology* **9**, 61.

Elledge, S. J. and Davis, R. W. (1989). Position and density effects on repression by stationary and mobile DNA-binding proteins. *Genes & Development* **3**, 185–197.

Elowitz, M. B., Surette, M. G., Wolf, P.-E., Stock, J. B. and Leibler, S. (1999). Protein mobility in the cytoplasm of Escherichia coli. *Journal of Bacteriology* **181**, 197–203.

Englesberg, E., Squires, C. and Meronk Jr, F. (1969). The l-arabinose operon in Escherichia coli B/r: a genetic demonstration of two functional states of the product of a regulator gene. *Proceedings of the National Academy of Sciences of the United States of America* **62**, 1100–1107.

Ernsting, B. R., Atkinson, M. R., Ninfa, A. J. and Matthews, R. G. (1992). Characterization of the regulon controlled by the leucine-responsive regulatory protein in Escherichia coli. *Journal of Bacteriology* **174**(4), 1109–1118.

Ettema, T. J., Brinkman, A. B., Tani, T. H., Rafferty, J. B. and Van Oost, J. D. (2002). A novel ligand-binding domain involved in regulation of amino acid metabolism in prokaryotes. *Journal of Biological Chemistry* **277**, 37464–37468.

Etzel, M. and Mörl, M. (2017). Synthetic riboswitches: from plug and pray toward plug and play. *Biochemistry* **56**, 1181–1198.

Falk, J., Mendler, M. and Kabisch, J. (2022). Pipette Show: an open source web application to support pipetting into microplates. *ACS Synthetic Biology* **11**, 996–999.

Falson, P., Goffeau, A., Boutry, M. and Jault, J.-M. (2004). Structural insight into the cooperativity between catalytic and noncatalytic sites of F1-ATPase. *Biochimica et Biophysica Acta - Bioenergetics* **1658**, 133–140.

Farmer, W. R. and Liao, J. C. (2000). Improving lycopene production in Escherichia coli by engineering metabolic control. *Nature Biotechnology* **18**, 533–537.

Fernandez-López, R., Ruiz, R., de la Cruz, F. and Moncalián, G. (2015). Transcription factor-based biosensors enlightened by the analyte. *Frontiers in Microbiology* **6**, 648.

Ferreira, R., Skrekas, C., Hedin, A., Sánchez, B. J., Siewers, V., Nielsen, J. and David, F. (2019). Model-assisted fine-tuning of central carbon metabolism in yeast through dCas9-based regulation. *ACS Synthetic Biology* **8**, 2457–2463.

Ferrer, L., Elsaraf, M., Mindt, M. and Wendisch, V. F. (2022). l-Serine biosensor-controlled fermentative production of l-tryptophan derivatives by Corynebacterium glutamicum. *Biology* **11**, 744.

Fic, E., Bonarek, P., Gorecki, A., Kedracka-Krok, S., Mikolajczak, J., Polit, A., Tworzydlo, M., Dziedzicka-Wasylewska, M. and Wasylewski, Z. (2009). cAMP receptor protein from Escherichia coli as a model of signal transduction in proteins – a review. *Microbial Physiology* **17**, 1–11.

Findeiß, S., Etzel, M., Will, S., Mörl, M. and Stadler, P. F. (2017). Design of artificial riboswitches as biosensors. *Sensors* **17**, 1990.

Fisher, S. L., Jiang, W., Wanner, B. L. and Walsh, C. T. (1995). Cross-talk between the histidine protein kinase VanS and the response regulator PhoB: characterization and identification of a VanS domain that inhibits activation of PhoB. *Journal of Biological Chemistry* **270**, 23143–23149.

Flachbart, L. K., Sokolowsky, S. and Marienhagen, J. (2019). Displaced by deceivers: prevention of biosensor cross-talk is pivotal for successful biosensor-based high-throughput screening campaigns. *ACS Synthetic Biology* **8**, 1847–1857.

Fong, S. S., Burgard, A. P., Herring, C. D., Knight, E. M., Blattner, F. R., Maranas, C. D. and Palsson, B. O. (2005). In silico design and adaptive evolution of Escherichia coli for production of lactic acid. *Biotechnology and Bioengineering* **91**, 643–648.

Fowler, C. C., Brown, E. D. and Li, Y. (2008). A FACS-based approach to engineering artificial riboswitches. *ChemBioChem* **9**, 1906–1911.

Frank, S. A. (2013). Input-output relations in biological systems: measurement, information and the Hill equation. *Biology Direct* **8**, 31.

Gadkar, K. G., Doyle III, F. J., Edwards, J. S. and Mahadevan, R. (2005). Estimating optimal profiles of genetic alterations using constraint-based models. *Biotechnology and Bioengineering* **89**, 243–251.

Galperin, M. Y. (2006). Structural classification of bacterial response regulators: diversity of output domains and domain combinations. *Journal of Bacteriology* **188**, 4169–4182.

Gardner, S. G., Johns, K. D., Tanner, R. and McCleary, W. R. (2014). The PhoU protein from Escherichia coli interacts with PhoR, PstB, and metals to form a phosphate-signaling complex at the membrane. *Journal of Bacteriology* **196**, 1741–1752.

Georgi, C., Buerger, J., Hillen, W. and Berens, C. (2012). Promoter strength driving TetR determines the regulatory properties of Tet-controlled expression systems. *PLOS ONE* **7**(7), e41620.

Gilbert, W. and Müller-Hill, B. (1966). Isolation of the lac repressor. *Proceedings of the National Academy of Sciences of the United States of America* **56**, 1891–1898.

Glick, B. R. (1995). Metabolic load and heterologous gene expression. *Biotechnology Advances* **13**, 247–261.

Gottesman, S. (1984). Bacterial regulation: global regulatory networks. *Annual Review of Genetics* **18**, 415–441.

Grundy, F. J. and Henkin, T. M. (1993). tRNA as a positive regulator of transcription antitermination in B. subtilis. *Cell* **74**, 475–482.

Guo, K.-H., Chen, P.-H., Lin, C., Chen, C.-F., Lee, I.-R. and Yeh, Y.-C. (2018). Determination of gold ions in human urine using genetically engineered microorganisms on a paper device. *ACS Sensors* **3**, 744–748.

Hallberg, Z. F., Su, Y., Kitto, R. Z. and Hammond, M. C. (2017). Engineering and in vivo applications of riboswitches. *Annual Review of Biochemistry* **86**, 515–539.

Halling, P. J. (1989). Do the laws of chemistry apply to living cells?. *Trends in Biochemical Sciences* **14**, 317–318.

Han, Y. and Zhang, F. (2020). Control strategies to manage trade-offs during microbial production. *Current Opinion in Biotechnology* **66**, 158–164.

Hanko, E. K., Minton, N. P. and Malys, N. (2017). Characterisation of a 3-hydroxypropionic acid-inducible system from Pseudomonas putida for orthogonal gene expression control in Escherichia coli and Cupriavidus necator. *Scientific Reports* **7**, 1724.

Hanko, E. K., Paiva, A. C., Jonczyk, M., Abbott, M., Minton, N. P. and Malys, N. (2020). A genome-wide approach for identification and characterisation of metabolite-inducible systems. *Nature Communications* **11**, 1213.

Hao, T., Li, G., Zhou, S. and Deng, Y. (2021). Engineering the reductive TCA pathway to dynamically regulate the biosynthesis of adipic acid in Escherichia coli. *ACS Synthetic Biology* **10**, 632–639.

Hartline, C. J., Schmitz, A. C., Han, Y. and Zhang, F. (2021). Dynamic control in metabolic engineering: theories, tools, and applications. *Metabolic Engineering* **63**, 126–140.

Haßlacher, M., Ivessa, A. S., Paltauf, F. and Kohlwein, S. D. (1993). Acetyl-CoA carboxylase from yeast is an essential enzyme and is regulated by factors that control phospholipid metabolism. *Journal of Biological Chemistry* **268**, 10946–10952.

Hayaishi, O., Nishizuka, Y., Tatibana, M., Takeshita, M. and Kuno, S. (1961). Enzymatic studies on the metabolism of β-alanine. *Journal of Biological Chemistry* **236**, 781–790.

Haydon, D. J. and Guest, J. R. (1991). A new family of bacterial regulatory proteins. *FEMS Microbiology Letters* **79**, 291–296.

Henry, M. F. and Cronan Jr, J. E. (1991). Escherichia coli transcription factor that both activates fatty acid synthesis and represses fatty acid degradation. *Journal of Molecular Biology* **222**, 843–849.

Herman, J. and Usher, W. (2017). SALib: an open-source Python library for sensitivity analysis. *Journal of Open Source Software* **2**(9), 97.

Hess, J. F., Oosawa, K., Kaplan, N. and Simon, M. I. (1988). Phosphorylation of three proteins in the signaling pathway of bacterial chemotaxis. *Cell* **53**, 79–87.

Hicks, M., Bachmann, T. T. and Wang, B. (2020). Synthetic biology enables programmable cell-based biosensors. *ChemPhysChem* **21**, 132–144.

Hill, A. V. (1910). The possible effects of the aggregation of the molecules of hæmoglobin on its dissociation curves. *The Journal of Physiology* **40**(suppl), 4–7.

Hong, K.-K., Vongsangnak, W., Vemuri, G. N. and Nielsen, J. (2011). Unravelling evolutionary strategies of yeast for improving galactose utilization through integrated systems level analysis. *Proceedings of the National Academy of Sciences of the United States of America* **108**, 12179–12184.

Hossain, G. S., Saini, M., Miyake, R., Ling, H. and Chang, M. W. (2020). Genetic biosensor design for natural product biosynthesis in microorganisms. *Trends in Biotechnology* **38**, 797–810.

Huber, P. J. (1981). *Robust Statistics*. John Wiley & Sons. New York, NY, USA.

Ingalls, B. (2018). *Mathematical Modelling in Systems Biology: An Introduction*. The Massachusetts Institute of Technology Press. Cambridge, MA, USA.

Jansen, M. L. A., Diderich, J. A., Mashego, M., Hassane, A., de Winde, J. H., Daran-Lapujade, P. and Pronk, J. T. (2005). Prolonged selection in aerobic, glucose-limited chemostat cultures of Saccharomyces cerevisiae causes a partial loss of glycolytic capacity. *Microbiology* **151**, 1657–1669.

Jantama, K., Haupt, M. J., Svoronos, S. A., Zhang, X., Moore, J. C., Shanmugam, K. T. and Ingram, L. O. (2008). Combining metabolic engineering and metabolic evolution to develop nonrecombinant strains of Escherichia coli C that produce succinate and malate. *Biotechnology and Bioengineering* **99**, 1140–1153.

Jiang, T., Li, C., Teng, Y., Zhang, R. and Yan, Y. (2020). Recent advances in improving metabolic robustness of microbial cell factories. *Current Opinion in Biotechnology* **66**, 69–77.

Jiang, X., Meng, X. and Xian, M. (2009). Biosynthetic pathways for 3-hydroxypropionic acid production. *Applied Microbiology and Biotechnology* **82**, 995–1003.

Jung, K., Fried, L., Behr, S. and Heermann, R. (2012). Histidine kinases and response regulators in networks. *Current Opinion in Microbiology* **15**, 118–124.

Kim, S.-K., Wilmes-Riesenberg, M. R. and Wanner, B. L. (1996). Involvement of the sensor kinase EnvZ in the in vivo activation of the response-regulator PhoB by acetyl phosphate. *Molecular Microbiology* **22**, 135–147.

Klein, J., Henrich, B. and Plapp, R. (1986). Cloning and expression of the pepD gene of Escherichia coli. *Journal of General Microbiology* **132**, 2337–2343.

Koch, M., Pandi, A., Borkowski, O., Batista, A. C. and Faulon, J.-L. (2019). Custom-made transcriptional biosensors for metabolic engineering. *Current Opinion in Biotechnology* **59**, 78–84.

Kofoid, E. C. and Parkinson, J. S. (1988). Transmitter and receiver modules in bacterial signaling proteins. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 4981–4985.

Kölling, R. and Lother, H. (1985). AsnC: an autogenously regulated activator of asparagine synthetase A transcription in Escherichia coli. *Journal of Bacteriology* **164**, 310–315.

Kreutz, C. (2019). Guidelines for benchmarking of optimization-based approaches for fitting mathematical models. *Genome Biology* **20**, 281.

Kruger, K., Grabowski, P. J., Zaug, A. J., Sands, J., Gottschling, D. E. and Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of tetrahymena. *Cell* **31**, 147–157.

Kumarevel, T., Nakano, N., Ponnuraj, K., Gopinath, S. C. B., Sakamoto, K., Shinkai, A., Kumar, P. K. R. and Yokoyama, S. (2008). Crystal structure of glutamine receptor protein from Sulfolobus tokodaii strain 7 in complex with its effector l-glutamine: implications of effector binding in molecular association and DNA binding. *Nucleic Acids Research* **36**, 4808–4820.

Kuthan, H. (2001). Self-organisation and orderly processes by individual protein complexes in the bacterial cell. *Progress in Biophysics and Molecular Biology* **75**, 1–17.

LaCroix, R. A., Sandberg, T. E., O'Brien, E. J., Utrilla, J., Ebrahim, A., Guzman, G. I., Szubin, R., Palsson, B. O. and Feist, A. M. (2015). Use of adaptive laboratory evolution to discover key mutations enabling rapid growth of Escherichia coli K-12 MG1655 on glucose minimal medium. *Applied and Environmental Microbiology* **81**, 17–30.

Lai, N., Luo, Y., Fei, P., Hu, P. and Wu, H. (2021). One stone two birds: biosynthesis of 3-hydroxypropionic acid from CO2 and syngas-derived acetic acid in Escherichia coli. *Synthetic and Systems Biotechnology* **6**, 144–152.

Lamarche, M. G., Wanner, B. L., Crépin, S. and Harel, J. (2008). The phosphate regulon and bacterial virulence: a regulatory network connecting phosphate homeostasis and pathogenesis. *FEMS Microbiology Reviews* **32**, 461–473.

Landgraf, D. (2012). *Quantifying Localizations and Dynamics in Single Bacterial Cells*. Harvard University. Cambridge, MA, USA.

Landry, B. P., Palanki, R., Dyulgyarov, N., Hartsough, L. A. and Tabor, J. J. (2018). Phosphatase activity tunes two-component system sensor detection threshold. *Nature Communications* **9**, 1433.

Lanzer, M. and Bujard, H. (1988). Promoters largely determine the efficiency of repressor action. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 8973–8977.

Latif, H., Sahin, M., Tarasova, J., Tarasova, Y., Portnoy, V. A., Nogales, J. and Zengler, K. (2015). Adaptive evolution of Thermotoga maritima reveals plasticity of the ABC transporter network. *Applied and Environmental Microbiology* **81**, 5477–5485.

Lee, D. H. and Palsson, B. O. (2010). Adaptive evolution of Escherichia coli K-12 MG1655 during growth on a nonnative carbon source, L-1,2-propanediol. *Applied and Environmental Microbiology* **76**, 4158–4168.

Lee, M. E., Aswani, A., Han, A. S., Tomlin, C. J. and Dueber, J. E. (2013). Expression-level optimization of a multi-enzyme pathway in the absence of a high-throughput assay. *Nucleic Acids Research* **41**, 10668–10678.

Levenberg, K. (1944). A method for the solution of certain non-linear problems in least squares. *Quarterly of Applied Mathematics* **2**, 164–168.

Li, S., Si, T., Wang, M. and Zhao, H. (2015). Development of a synthetic malonyl-CoA sensor in Saccharomyces cerevisiae for intracellular metabolite monitoring and genetic screening. *ACS Synthetic Biology* **4**, 1308–1315.

Li, X., Guo, D., Cheng, Y., Zhu, F., Deng, Z. and Liu, T. (2014). Overproduction of fatty acids in engineered Saccharomyces cerevisiae. *Biotechnology and Bioengineering* **111**, 1841–1852.

Liang, C., Zhang, X., Wu, J., Mu, S., Wu, Z., Jin, J. M. and Tang, S. Y. (2020). Dynamic control of toxic natural product biosynthesis by an artificial regulatory circuit. *Metabolic Engineering* **57**, 239–246.

Liberman, J. A. and Wedekind, J. E. (2012). Riboswitch structure in the ligand-free state. *Wiley Interdisciplinary Reviews: RNA* **3**, 369–384.

Lillacci, G. and Khammash, M. (2010). Parameter estimation and model selection in computational biology. *PLOS Computational Biology* **6**, e1000696.

Lin, R., D'Ari, R. and Newman, E. B. (1992). Lambda placMu insertions in genes of the leucine regulon: extension of the regulon to genes not regulated by leucine. *Journal of Bacteriology* **174**, 1948–1955.

Liu, B., Kearns, D. B. and Bechhofer, D. H. (2016). Expression of multiple Bacillus subtilis genes is controlled by decay of slrA mRNA from Rho-dependent 3′ ends. *Nucleic Acids Research* **44**, 3364–3372.

Liu, D., Evans, T. and Zhang, F. (2015). Applications and advances of metabolite biosensors for metabolic engineering. *Metabolic Engineering* **31**, 35–43.

Liu, D., Mannan, A. A., Han, Y., Oyarzún, D. A. and Zhang, F. (2018). Dynamic metabolic control: towards precision engineering of metabolism. *Journal of Industrial Microbiology and Biotechnology* **45**, 535–543.

Liu, D. and Zhang, F. (2018). Metabolic Feedback Circuits Provide Rapid Control of Metabolite Dynamics. *ACS Synthetic Biology* **7**, 347–356.

Liu, H., Orell, A., Maes, D., van Wolferen, M., Lindås, A.-C., Bernander, R., Albers, S.-V., Charlier, D. and Peeters, E. (2014). BarR, an Lrp-type transcription factor in Sulfolobus acidocaldarius, regulates an aminotransferase gene in a β-alanine responsive manner. *Molecular Microbiology* **92**, 625–639.

Liu, X., Silverman, A. D., Alam, K. K., Iverson, E., Lucks, J. B., Jewett, M. C. and Raman, S. (2020). Design of a transcriptional biosensor for the portable, on-demand detection of cyanuric acid. *ACS Synthetic Biology* **9**, 84–94.

López-Garrido, J., Puerta-Fernández, E. and Casadesús, J. (2014). A eukaryotic-like 3′ untranslated region in Salmonella enterica hilD mRNA. *Nucleic Acids Research* **42**, 5894–5906.

Lotz, T. S. and Suess, B. (2018). Small-Molecule-Binding Riboswitches. *Microbiology Spectrum* **6**(4), 26.

Lu, Z., Zhang, X., Dai, J., Wang, Y. and He, W. (2019). Engineering of leucine-responsive regulatory protein improves spiramycin and bitespiramycin biosynthesis. *Microbial Cell Factories* **18**, 1–12.

Lv, Y., Gu, Y., Xu, J., Zhou, J. and Xu, P. (2020). Coupling metabolic addiction with negative autoregulation to improve strain stability and pathway yield. *Metabolic Engineering* **61**, 79–88.

Maas, W. K. (1952). Pantothenate studies: description of the extracted pantothenate-synthesizing enzyme of Escherichia coli. *Journal of Biological Chemistry* **198**, 23–32.

Maeda, T. and Wachi, M. (2012). 3' Untranslated region-dependent degradation of the aceA mRNA, encoding the glyoxylate cycle enzyme isocitrate lyase, by RNase E/G in Corynebacterium glutamicum. *Applied and Environmental Microbiology* **78**, 8753–8761.

Magnus, W. (2021). *Genetically Encoded Biosensors for the Improvement of 3-Hydroxypropionic Acid Production in Escherichia coli*. Vrije Universiteit Brussel. Brussels, Belgium.

Mahr, R. and Frunzke, J. (2016). Transcription factor-based biosensors in biotechnology: current state and future prospects. *Applied Microbiology and Biotechnology* **100**, 79–90.

Maiwald, T., Hass, H., Steiert, B., Vanlier, J., Engesser, R., Raue, A., Kipkeew, F., Bock, H. H., Kaschek, D., Kreutz, C. and Timmer, J. (2016). Driving the model to its limit: profile likelihood based model reduction. *PLOS ONE* **11**(9), e0162366.

Makino, K., Shinagawa, H., Amemura, M., Kawamoto, T., Yamada, M. and Nakata, A. (1989). Signal transduction in the phosphate regulon of Escherichia coli involves phosphotransfer between PhoR and PhoB proteins. *Journal of Molecular Biology* **210**, 551–559.

Makino, K., Shinagawa, H., Amemura, M. and Nakata, A. (1986). Nucleotide sequence of the phoB gene, the positive regulatory gene for the phosphate regulon of Escherichia coli K-12. *Journal of Molecular Biology* **190**, 37–44.

Mannan, A. A. and Bates, D. G. (2021). Designing an irreversible metabolic switch for scalable induction of microbial chemical production. *Nature Communications* **12**, 1–11.

Mannan, A. A., Liu, D., Zhang, F. and Oyarzún, D. A. (2017). Fundamental design principles for transcription-factor-based metabolite biosensors. *ACS Synthetic Biology* **6**, 1851–1859.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* **11**(2), 431–441.

Martin, V. J. J., Pitera, D. J., Withers, S. T., Newman, J. D. and Keasling, J. D. (2003). Engineering a mevalonate pathway in Escherichia coli for production of terpenoids. *Nature Biotechnology* **21**, 796–802.

Maury, J., Kannan, S., Jensen, N. B., Öberg, F. K., Kildegaard, K. R., Forster, J., Nielsen, J., Workman, C. T. and Borodina, I. (2018). Glucose-dependent promoters for dynamic regulation of metabolic pathways. *Frontiers in Bioengineering and Biotechnology* **6**, 63.

McFarland, K. A. and Dorman, C. J. (2008). Autoregulated expression of the gene coding for the leucine-responsive protein, Lrp, a global regulator in Salmonella enterica serovar Typhimurium. *Microbiology* **154**, 2008–2016.

McKay, D. B. and Steitz, T. A. (1981). Structure of catabolite gene activator protein at 2.9 Å resolution suggests binding to left-handed B-DNA. *Nature* **290**, 744–749.

McKay, M. D., Beckman, R. J. and Conover, W. J. (1979). A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* **21**, 239–245.

Mehta, P. K., Hale, T. I. and Christen, P. (1993). Aminotransferases: demonstration of homology and division into evolutionary subgroups. *European Journal of Biochemistry* **214**, 549–561.

Michaelis, L. and Menten, M. L. (1913). Die Kinetik der Invertinwirkung. *Biochemische Zeitschrift* **49**, 333–369.

Milano, T., Angelaccio, S., Tramonti, A., Di Salvo, M. L., Contestabile, R. and Pascarella, S. (2016). A bioinformatics analysis reveals a group of MocR bacterial transcriptional regulators linked to a family of genes coding for membrane proteins. *Biochemistry Research International* **2016**, 4360285.

Mizuno, T., Chou, M. Y. and Inouye, M. (1984). A unique mechanism regulating gene expression: translational inhibition by a complementary RNA transcript (micRNA). *Proceedings of the National Academy of Sciences of the United States of America* **81**, 1966–1970.

Moré, J. J. (1978). The Levenberg-Marquardt algorithm: implementation and theory. In: Numerical Analysis (G. A. Watson, ed.). Springer. Berlin, Germany. pp. 105–116.

Moser, F., Borujeni, A. E., Ghodasara, A. N., Cameron, E., Park, Y. and Voigt, C. A. (2018). Dynamic control of endogenous metabolism with combinatorial logic circuits. *Molecular Systems Biology* **14**, e8605.

Nahvi, A., Sudarsan, N., Ebert, M. S., Zou, X., Brown, K. L. and Breaker, R. R. (2002). Genetic control by a metabolite binding mRNA. *Chemistry & Biology* **9**, 1043–1049.

Nakamura, K. and Bernheim, F. (1961). Studies of malonic semialdehyde dehydrogenase from Pseudomonas aeruginosa. *Biochimica et Biophysica Acta* **50**, 147–152.

Nardella, C., Barile, A., di Salvo, M. L., Milano, T., Pascarella, S., Tramonti, A. and Contestabile, R. (2020). Interaction of Bacillus subtilis GabR with the gabTD promoter: role of repeated sequences and effect of GABA in transcriptional activation. *The FEBS Journal* **287**, 4952–4970.

Newville, M., Stensitzki, T., Allen, D. B., Rawlik, M., Ingargiola, A. and Nelson, A. (2016). Lmfit: non-linear least-square minimization and curve-fitting for Python. **URL:** *https://doi.org/10.5281/zenodo.598352*

Nguyen-Duc, T., van Oeffelen, L., Song, N., Hassanzadeh-Ghassabeh, G., Muyldermans, S., Charlier, D. and Peeters, E. (2013). The genome-wide binding profile of the Sulfolobus solfataricus transcription factor Ss-LrpB shows binding events beyond direct transcription regulation. *BMC Genomics* **14**, 828.

Nguyen, N. H., Ainala, S. K., Zhou, S. and Park, S. (2019). A novel 3-hydroxypropionic acid-inducible promoter regulated by the LysR-type transcriptional activator protein MmsR of Pseudomonas denitrificans. *Scientific Reports* **9**, 5333.

Nguyen, N. H., Kim, J.-R. and Park, S. (2019). Development of biosensor for 3-hydroxypropionic acid. *Biotechnology and Bioprocess Engineering* **24**, 109–118.

Nguyen-Vo, T. P., Ko, S., Ryu, H., Kim, J. R., Kim, D. and Park, S. (2020). Systems evaluation reveals novel transporter YohJK renders 3-hydroxypropionate tolerance in Escherichia coli. *Scientific Reports* **10**, 19064.

Nguyen-Vo, T. P., Liang, Y., Sankaranarayanan, M., Seol, E., Chun, A. Y., Somasundar, A., Chauhan, A. S., Kim, J. R. and Park, S. (2019). Development of 3-hydroxypropionic-acid-tolerant strain of Escherichia coli W and role of minor global regulator yieP. *Metabolic Engineering* **53**, 48–58.

Nguyen-Vo, T. P., Ryu, H., Sauer, M. and Park, S. (2022). Improvement of 3-hydroxypropionic acid tolerance in Klebsiella pneumoniae by novel transporter YohJK. *Bioresource Technology* **346**, 126613.

Ni, C., Fox, K. J. and Prather, K. L. (2022). Substrate-activated expression of a biosynthetic pathway in Escherichia coli. *Biotechnology Journal* **17**, 2000433.

Ni'Bhriain, N. N., Silver, S. and Foster, T. J. (1983). Tn5 insertion mutations in the mercuric ion resistance genes derived from plasmid R100. *Journal of Bacteriology* **155**, 690–703.

Nielsen, F. C. and Christiansen, J. (1992). Endonucleolysis in the turnover of insulin-like growth factor II mRNA. *Journal of Biological Chemistry* **267**, 19404–19411.

Ninfa, A. J. and Magasanik, B. (1986). Covalent modification of the glnG product, NRI, by the glnL product, NRII, regulates the transcription of the glnALG operon in Escherichia coli. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 5909–5913.

Nishino, K., Honda, T. and Yamaguchi, A. (2005). Genome-wide analyses of Escherichia coli gene expression responsive to the BaeSR two-component regulatory system. *Journal of Bacteriology* **187**, 1763–1772.

Nixon, B. T., Ronson, C. W. and Ausubel, F. M. (1986). Two-component regulatory systems responsive to environmental stimuli share strongly conserved domains with the nitrogen assimilation regulatory genes ntrB and ntrC. *Proceedings of the National Academy of Sciences of the United States of America* **83**, 7850–7854.

Nomura, S., Horiuchi, T., Ōmura, S. and Hata, T. (1972). The action mechanism of cerulenin: effect of cerulenin on sterol and fatty acid biosyntheses in yeast. *The Journal of Biochemistry* **71**, 783–796.

Novichkov, P. S., Kazakov, A. E., Ravcheev, D. A., Leyn, S. A., Kovaleva, G. Y., Sutormin, R. A., Kazanov, M. D., Riehl, W., Arkin, A. P., Dubchak, I. and Rodionov, D. A. (2013). RegPrecise 3.0 - a resource for genome-scale exploration of transcriptional regulation in bacteria. *BMC Genomics* **14**, 745.

Nowroozi, F. F., Baidoo, E. E. K., Ermakov, S., Redding-Johanson, A. M., Batth, T. S., Petzold, C. J. and Keasling, J. D. (2014). Metabolic pathway optimization using ribosome binding site variants and combinatorial gene assembly. *Applied Microbiology and Biotechnology* **98**, 1567–1581.

Okuda, K., Kato, S., Ito, T., Shiraki, S., Kawase, Y., Goto, M., Kawashima, S., Hemmi, H., Fukada, H. and Yoshimura, T. (2015). Role of the aminotransferase domain in

Bacillus subtilis GabR, a pyridoxal 5′-phosphate-dependent transcriptional regulator. *Molecular Microbiology* **95**, 245–257.

Otero, J. M., Cimini, D., Patil, K. R., Poulsen, S. G., Olsson, L. and Nielsen, J. (2013). Industrial systems biology of Saccharomyces cerevisiae enables novel succinic acid cell factory. *PLOS ONE* **8**(1), e54144.

Otero-Muras, I., Mannan, A. A., Banga, J. R. and Oyarzún, D. A. (2019). Multiobjective optimization of gene circuits for metabolic engineering. *IFAC-PapersOnLine* **52**, 13–16.

Ozcan, S. and Johnston, M. (1995). Three different regulatory mechanisms enable yeast hexose transporter (HXT) genes to be induced by different levels of glucose. *Molecular and Cellular Biology* **15**, 1564–1572.

Palomares, L. A., Estrada-Moncada, S. and Ramírez, O. T. (2004). Production of recombinant proteins. In: Recombinant Gene Expression: Reviews and Protocols (P. Balbás and A. Lorence, eds). 2 edn. Humana Press. Totowa, NJ, USA. pp. 15–52.

Pastan, I. and Adhya, S. (1976). Cyclic adenosine 5'-monophosphate in Escherichia coli. *Bacteriological Reviews* **40**, 527–551.

Pelletier, J. and Sonenberg, N. (1985). Insertion mutagenesis to increase secondary structure within the 5′ noncoding region of a eukaryotic mRNA reduces translational efficiency. *Cell* **40**, 515–526.

Pérez-Rueda, E. and Janga, S. C. (2010). Identification and genomic analysis of transcription factors in Archaeal genomes exemplifies their functional architecture and evolutionary origin. *Molecular Biology and Evolution* **27**, 1449–1459.

Perutz, M. F. (1980). Review lecture - stereochemical mechanism of oxygen transport by haemoglobin. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **208**, 135–162.

Pfeifer, E., Gätgens, C., Polen, T. and Frunzke, J. (2017). Adaptive laboratory evolution of Corynebacterium glutamicum towards higher growth rates on glucose minimal medium. *Scientific Reports* **7**, 16780.

Pharkya, P., Burgard, A. P. and Maranas, C. D. (2004). OptStrain: a computational framework for redesign of microbial production systems. *Genome Research* **14**, 2367–2376.

Pigou, M. and Morchain, J. (2015). Investigating the interactions between physical and biological heterogeneities in bioreactors using compartment, population balance and metabolic models. *Chemical Engineering Science* **126**, 267–282.

Pittard, A. J. and Davidson, B. E. (1991). TyrR protein of Escherichia coli and its role as repressor and activator. *Molecular Microbiology* **5**, 1585–1592.

Platko, J. V. and Calvo, J. M. (1993). Mutations affecting the ability of Escherichia coli Lrp to bind DNA, activate transcription, or respond to leucine. *Journal of Bacteriology* **175**, 1110–1117.

Platko, J. V., Willins, D. A. and Calvo, J. M. (1990). The ilvIH operon of Escherichia coli is positively regulated. *Journal of Bacteriology* **172**, 4563–4570.

Pohlmann, A., Fricke, W. F., Reinecke, F., Kusian, B., Liesegang, H., Cramm, R., Eitinger, T., Ewering, C., Pötter, M., Schwartz, E., Strittmatter, A., Voß, I., Gottschalk, G., Steinbüchel, A., Friedrich, B. and Bowien, B. (2006). Genome sequence of the bioplastic-producing "Knallgas" bacterium Ralstonia eutropha H16. *Nature Biotechnology* **24**, 1257–1262.

Qin, L., Dong, S., Yu, J., Ning, X., Xu, K., Zhang, S. J., Xu, L., Li, B. Z., Li, J., Yuan, Y. J. and Li, C. (2020). Stress-driven dynamic regulation of multiple tolerance genes improves robustness and productive capacity of Saccharomyces cerevisiae in industrial lignocellulose fermentation. *Metabolic Engineering* **61**, 160–170.

Qin, L., Liu, X., Xu, K. and Li, C. (2022). Mining and design of biosensors for engineering microbial cell factory. *Current Opinion in Biotechnology* **75**, 102694.

R Core Team (2022). R: a language and environment for statistical computing. **URL:** *https://www.r-project.org/*

Rajaraman, E., Agarwal, A., Crigler, J., Seipelt-Thiemann, R., Altman, E. and Eiteman, M. A. (2016). Transcriptional analysis and adaptive evolution of Escherichia coli strains growing on acetate. *Applied Microbiology and Biotechnology* **100**, 7777–7785.

Raman, S., Rogers, J. K., Taylor, N. D. and Church, G. M. (2014). Evolution-guided optimization of biosynthetic pathways. *Proceedings of the National Academy of Sciences of the United States of America* **111**, 17803–17808.

Rathnasingh, C., Raj, S. M., Lee, Y., Catherine, C., Somasundar, A. and Park, S. (2012). Production of 3-hydroxypropionic acid via malonyl-CoA pathway using recombinant Escherichia coli strains. *Journal of Biotechnology* **157**, 633–640.

Raue, A., Becker, V., Klingmüller, U. and Timmer, J. (2010). Identifiability and observability analysis for experimental design in nonlinear dynamical models. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **20**, 045105.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U. and Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* **25**, 1923–1929.

Raue, A., Kreutz, C., Maiwald, T., Klingmüller, U. and Timmer, J. (2011). Addressing parameter identifiability by model-based experimentation. *IET Systems Biology* **5**, 120–130.

Raue, A., Schilling, M., Bachmann, J., Matteson, A., Schelker, M., Kaschek, D., Hug, S., Kreutz, C., Harms, B. D., Theis, F. J., Klingmüller, U. and Timmer, J. (2013). Lessons learned from quantitative dynamical modeling in systems biology. *PLOS ONE* **8**(9), e74335.

Ravikumar, S., Baylon, M. G., Park, S. J. and Choi, J.-i. (2017). Engineered microbial biosensors based on bacterial two-component systems as synthetic biotechnology platforms in bioremediation and biorefinery. *Microbial Cell Factories* **16**, 62.

Ravikumar, S., Yoo, I.-k., Lee, S. Y. and Hong, S. H. (2011). A study on the dynamics of the zraP gene expression profile and its application to the construction of zinc adsorption bacteria. *Bioprocess and Biosystems Engineering* **34**, 1119.

Razin, S., Bachrach, U. and Gery, I. (1958). Formation of β-alanine from spermine and spermidine by Pseudomonas aeruginosa. *Nature* **181**, 700–701.

Riehle, M. M., Bennett, A. F., Lenski, R. E. and Long, A. D. (2003). Evolutionary changes in heat-inducible gene expression in lines of Escherichia coli adapted to high temperature. *Physiological Genomics* **14**, 47–58.

Rogers, J. K. and Church, G. M. (2016*a*). Genetically encoded sensors enable real-time observation of metabolite production. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 2388–2393.

Rogers, J. K. and Church, G. M. (2016*b*). Multiplexed engineering in biology. *Trends in Biotechnology* **34**, 198–206.

Rogers, J. K., Guzman, C. D., Taylor, N. D., Raman, S., Anderson, K. and Church, G. M. (2015). Synthetic biosensors for precise gene control and real-time monitoring of metabolites. *Nucleic Acids Research* **43**, 7648–7660.

Rogers, J. K., Taylor, N. D. and Church, G. M. (2016). Biosensor-based engineering of biosynthetic pathways. *Current Opinion in Biotechnology* **42**, 84–91.

Rossbach, S., Kulpa, D. A., Rossbach, U. and de Bruijn, F. J. (1994). Molecular and genetic characterization of the rhizopine catabolism (mocABRC) genes of Rhizobium meliloti L5-30. *Molecular and General Genetics* **245**(1), 11–24.

Sandberg, T. E., Lloyd, C. J., Palsson, B. O. and Feist, A. M. (2017). Laboratory evolution to alternating substrate environments yields distinct phenotypic and genetic adaptive strategies. *Applied and Environmental Microbiology* **83**(13), e00410–17.

Sandberg, T. E., Salazar, M. J., Weng, L. L., Palsson, B. O. and Feist, A. M. (2019). The emergence of adaptive laboratory evolution as an efficient tool for biological discovery and industrial biotechnology. *Metabolic Engineering* **56**, 1–16.

Santoro, S. W., Wang, L., Herberich, B., King, D. S. and Schultz, P. G. (2002). An efficient system for the evolution of aminoacyl-tRNA synthetase specificity. *Nature Biotechnology* **20**, 1044–1048.

Savidor, A., Chalupowicz, L., Teper, D., Gartemann, K. H., Eichenlaub, R., Manulis-Sasson, S., Barash, I. and Sessa, G. (2014). Clavibacter michiganensis subsp. michiganensis Vatr1 and Vatr2 transcriptional regulators are Required for virulence in tomato. *Molecular Plant-Microbe Interactions* **27**, 1035–1047.

Schneider, F., Krämer, R. and Burkovski, A. (2004). Identification and characterization of the main β-alanine uptake system in Escherichia coli. *Applied Microbiology and Biotechnology* **65**, 576–582.

Schneider, T. D. and Stephens, R. M. (1990). Sequence logos: a new way to display consensus sequences. *Nucleic Acids Research* **18**, 6097.

Scholten, M. and Tommassen, J. (1993). Topology of the PhoR protein of Escherichia coli and functional analysis of internal deletion mutants. *Molecular Microbiology* **8**, 269–275.

Schujman, G. E., Paoletti, L., Grossman, A. D. and de Mendoza, D. (2003). FapR, a bacterial transcription factor involved in global regulation of membrane lipid biosynthesis. *Developmental Cell* **4**, 663–672.

Seok, J. Y., Han, Y. H., Yang, J.-S., Yang, J., Lim, H. G., Kim, S. G., Seo, S. W. and Jung, G. Y. (2021). Synthetic biosensor accelerates evolution by rewiring carbon metabolism toward a specific metabolite. *Cell Reports* **36**, 109589.

Seok, J. Y., Yang, J., Choi, S. J., Lim, H. G., Choi, U. J., Kim, K. J., Park, S., Yoo, T. H. and Jung, G. Y. (2018). Directed evolution of the 3-hydroxypropionic acid production pathway by engineering aldehyde dehydrogenase using a synthetic selection device. *Metabolic Engineering* **47**, 113–120.

Setny, P. and Wiśniewska, M. D. (2018). Water-mediated conformational preselection mechanism in substrate binding cooperativity to protein kinase A. *Proceedings of the National Academy of Sciences of the United States of America* **115**, 3852–3857.

Shrivastava, T., Dey, A. and Ramachandran, R. (2009). Ligand-induced structural transitions, mutational analysis, and 'open' quaternary structure of the M. tuberculosis Feast/Famine Regulatory Protein (Rv3291c). *Journal of Molecular Biology* **392**, 1007–1019.

Slotnick, I. J. and Weinfeld, H. (1957). Dihydrouracil as a growth factor for mutant strains of Escherichia coli. *Journal of Bacteriology* **74**, 122–125.

Somasundar, A., Raj, S. M., Rathnasingh, C. and Park, S. (2011). Development of recombinant Klebsiella pneumoniae ΔdhaT strain for the co-production of 3-hydroxypropionic acid and 1,3-propanediol from glycerol. *Applied Microbiology and Biotechnology* **90**, 1253–1265.

Sonderegger, M. and Sauer, U. (2003). Evolutionary engineering of Saccharomyces cerevisiae for anaerobic growth on xylose. *Applied and Environmental Microbiology* **69**, 1990–1998.

Song, C. W., Kim, J. W., Cho, I. J. and Lee, S. Y. (2016). Metabolic engineering of Escherichia coli for the production of 3-hydroxypropionic acid and malonic acid through β-alanine route. *ACS Synthetic Biology* **5**, 1256–1263.

Song, N., Nguyen Duc, T., van Oeffelen, L., Muyldermans, S., Peeters, E. and Charlier, D. (2013). Expanded target and cofactor repertoire for the transcriptional activator LysM from Sulfolobus. *Nucleic Acids Research* **41**, 2932–2949.

Stock, A. M., Robinson, V. L. and Goudreau, P. N. (2000). Two-component signal transduction. *Annual Review of Biochemistry* **69**, 183–215.

Stoebel, D. M., Hokamp, K., Last, M. S. and Dorman, C. J. (2009). Compensatory evolution of gene regulation in response to stress by Escherichia coli lacking RpoS. *PLOS Genetics* **5**(10), e1000671.

Stoeckle, M. Y. and Hanafusa, H. (1989). Processing of 9E3 mRNA and regulation of its stability in normal and Rous sarcoma virus-transformed cells. *Molecular and Cellular Biology* **9**, 4738–4745.

Sullivan, M. J., Curson, A. R., Shearer, N., Todd, J. D., Green, R. T. and Johnston, A. W. (2011). Unusual regulation of a leaderless operon involved in the catabolism of dimethylsulfoniopropionate in Rhodobacter sphaeroides. *PLOS ONE* **6**(1), e15972.

Suvorova, I. A. and Rodionov, D. A. (2016). Comparative genomics of pyridoxal 5'-phosphate-dependent transcription factor regulons in Bacteria. *Microbial genomics* **2**(1).

Sybers, D., Bernauw, A. J., El Masri, D., Maklad, H. R., Charlier, D., De Mey, M., Bervoets, I. and Peeters, E. (2022). Engineering transcriptional regulation in Escherichia coli using an archaeal TetR-family transcription factor. *Gene* **809**, 146010.

Takenaka, T., Ito, T., Miyahara, I., Hemmi, H. and Yoshimura, T. (2015). A new member of MocR/GabR-type PLP-binding regulator of d-alanyl-d-alanine ligase in Brevibacillus brevis. *The FEBS Journal* **282**, 4201–4217.

Tanna, T., Ramachanderan, R. and Platt, R. J. (2021). Engineered bacteria to report gut function: technologies and implementation. *Current Opinion in Microbiology* **59**, 24–33.

Teo, W. S. and Chang, M. W. (2015). Bacterial XylRs and synthetic promoters function as genetically encoded xylose biosensors in Saccharomyces cerevisiae. *Biotechnology Journal* **10**, 315–322.

Teo, W. S., Hee, K. S. and Chang, M. W. (2013). Bacterial FadR and synthetic promoters function as modular fatty acid sensor-regulators in Saccharomyces cerevisiae. *Engineering in Life Sciences* **13**, 456–463.

Thi Nguyen, T., Lama, S., Kumar Ainala, S., Sankaranarayanan, M., Singh Chauhan, A., Rae Kim, J. and Park, S. (2021). Development of Pseudomonas asiatica as a host for the production of 3-hydroxypropionic acid from glycerol. *Bioresource Technology* **329**, 124867.

Torres-Bacete, J., García, J. L. and Nogales, J. (2021). A portable library of phosphate-depletion based synthetic promoters for customable and automata control of gene expression in bacteria. *Microbial Biotechnology* **14**, 2643–2658.

Transtrum, M. K., Machta, B. B. and Sethna, J. P. (2010). Why are nonlinear fits to data so challenging?. *Physical Review Letters* **104**, 060201.

Transtrum, M. K., Machta, B. B. and Sethna, J. P. (2011). Geometry of nonlinear least squares with applications to sloppy models and optimization. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* **83**, 036701.

Transtrum, M. K. and Qiu, P. (2012). Optimal experiment selection for parameter estimation in biological differential equation models. *BMC Bioinformatics* **13**, 181.

Trausch, J. J. and Batey, R. T. (2015). Design of modular "plug-and-play" expression platforms derived from natural riboswitches for engineering novel genetically encodable RNA regulatory devices. In: Riboswitches as Targets and Tools (D. H. Burke-Aguero, ed.). Vol. 550. Academic Press. Cambridge, MA, USA. pp. 41–71.

Tuerk, C. and Gold, L. (1990). Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. *Science* **249**, 505–510.

Ulrich, L. E., Koonin, E. V. and Zhulin, I. B. (2005). One-component systems dominate signal transduction in prokaryotes. *Trends in Microbiology* **13**, 52–56.

Uluşeker, C., Torres-Bacete, J., García, J. L., Hanczyc, M. M., Nogales, J. and Kahramanoğulları, O. (2019). Quantifying dynamic mechanisms of auto-regulation in Escherichia coli with synthetic promoter in response to varying external phosphate levels. *Scientific Reports* **9**, 2076.

Umeyama, T., Okada, S. and Ito, T. (2013). Synthetic gene circuit-mediated monitoring of endogenous metabolites: identification of GAL11 as a novel multicopy enhancer of S-adenosylmethionine level in yeast. *ACS Synthetic Biology* **2**, 425–430.

U.S. EPA (1996). Method 8315A (SW-846): determination of carbonyl compounds by high performance liquid chromatography (HPLC).

Utrilla, J., Licona-Cassani, C., Marcellin, E., Gosset, G., Nielsen, L. K. and Martinez, A. (2012). Engineering and adaptive evolution of Escherichia coli for d-lactate fermentation reveals GatC as a xylose transporter. *Metabolic Engineering* **14**, 469–476.

van Hijum, S. A. F. T., Medema, M. H. and Kuipers, O. P. (2009). Mechanisms and evolution of control logic in prokaryotic transcriptional regulation. *Microbiology and Molecular Biology Reviews* **73**, 481–509.

Van Rossum, G. and Drake, F. L. (2009). *Python 3 Reference Manual*. CreateSpace. Scotts Valley, CA, USA.

Vandamme, P. and Coenye, T. (2004). Taxonomy of the genus Cupriavidus: a tale of lost and found. *International Journal of Systematic and Evolutionary Microbiology* **54**, 2285–2289.

Vanden Berghen, F. (2004). Levenberg-Marquardt algorithms vs trust region algorithms. **URL:** *https://www.applied-mathematics.net/LMvsTR/LMvsTR.pdf*

Vassart, A., Van Wolferen, M., Orell, A., Hong, Y., Peeters, E., Albers, S.-V. and Charlier, D. (2013). Sa-Lrp from Sulfolobus acidocaldarius is a versatile, glutamine-responsive, and architectural transcriptional regulator. *MicrobiologyOpen* **2**, 75–93.

Verhamme, D. T., Arents, J. C., Postma, P. W., Crielaard, W. and Hellingwerf, K. J. (2002). Investigation of in vivo cross-talk between key two-component systems of Escherichia coli. *Microbiology* **148**, 69–78.

Verma, B. K., Mannan, A. A., Zhang, F. and Oyarzún, D. A. (2022). Trade-offs in biosensor optimization for dynamic pathway engineering. *ACS Synthetic Biology* **11**, 228–240.

von Kamp, A. and Klamt, S. (2017). Growth-coupled overproduction is feasible for almost all metabolites in five major production organisms. *Nature Communications* **8**, 15956.

Wachsmuth, M., Domin, G., Lorenz, R., Serfling, R., Findeiß, S., Stadler, P. F. and Mörl, M. (2015). Design criteria for synthetic riboswitches acting on transcription. *RNA Biology* **12**, 221–231.

Wachsmuth, M., Findeiß, S., Weissheimer, N., Stadler, P. F. and Mörl, M. (2013). De novo design of a synthetic riboswitch that regulates transcription termination. *Nucleic Acids Research* **41**, 2541–2551.

Wang, B., Barahona, M. and Buck, M. (2015). Amplification of small molecule-inducible gene expression via tuning of intracellular receptor densities. *Nucleic Acids Research* **43**, 1955–1964.

Wang, C., Liwei, M., Park, J. B., Jeong, S. H., Wei, G., Wang, Y. and Kim, S. W. (2018). Microbial platform for terpenoid production: Escherichia coli and yeast. *Frontiers in Microbiology* **9**, 2460.

Wang, K., Sybers, D., Maklad, H. R., Lemmens, L., Lewyllie, C., Zhou, X., Schult, F., Bräsen, C., Siebers, B., Valegård, K., Lindås, A.-C. and Peeters, E. (2019). A TetR-family transcription factor regulates fatty acid metabolism in the archaeal model organism Sulfolobus acidocaldarius. *Nature Communications* **10**, 1542.

Wanner, B. L. (1992). Is cross regulation by phosphorylation of two-component response regulator proteins important in bacteria?. *Journal of Bacteriology* **174**, 2053–2058.

Wanner, B. L. (1993). Gene regulation by phosphate in enteric bacteria. *Journal of Cellular Biochemistry* **51**, 47–54.

Wanner, B. L. (1996). Phosphorus assimilation and control of the phosphate regulon. In: Escherichia coli and Salmonella: Cellular and Molecular Biology (F. C. Neidhardt, R. Curtiss III, J. L. Ingraham, E. C. C. Lin, K. B. Low, B. Magasanik, W. S. Reznikoff, M. Riley, M. Schaechter and H. E. Umbarger, eds). 2 edn. American Society for Microbiology. Washington, DC, USA. pp. 1357–1381.

Wanner, B. L., Wilmes, M. R. and Young, D. C. (1988). Control of bacterial alkaline phosphatase synthesis and variation in an Escherichia coli K-12 phoR mutant by adenyl cyclase, the cyclic AMP receptor protein, and the phoM operon. *Journal of Bacteriology* **170**, 1092–1102.

Wanner, B. L. and Wilmes-Riesenberg, M. R. (1992). Involvement of phosphotransacetylase, acetate kinase, and acetyl phosphate synthesis in control of the phosphate regulon in Escherichia coli. *Journal of Bacteriology* **174**, 2124–2130.

Waterfall, J. J., Casey, F. P., Gutenkunst, R. N., Brown, K. S., Myers, C. R., Brouwer, P. W., Elser, V. and Sethna, J. P. (2006). Sloppy-model universality class and the Vandermonde matrix. *Physical Review Letters* **97**, 150601.

Werner, F. and Grohmann, D. (2011). Evolution of multisubunit RNA polymerases in the three domains of life. *Nature Reviews Microbiology* **9**, 85–98.

Wieland, F.-G., Hauber, A. L., Rosenblatt, M., Tönsing, C. and Timmer, J. (2021). On structural and practical identifiability. *Current Opinion in Systems Biology* **25**, 60–69.

Win, M. N. and Smolke, C. D. (2007). A modular and extensible RNA-based gene-regulatory platform for engineering cellular function. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 14283–14288.

Winkler, W. C., Cohen-Chalamish, S. and Breaker, R. R. (2002). An mRNA structure that controls gene expression by binding FMN. *Proceedings of the National Academy of Sciences of the United States of America* **99**, 15908–15913.

Winkler, W. C., Nahvi, A., Roth, A., Collins, J. A. and Breaker, R. R. (2004). Control of gene expression by a natural metabolite-responsive ribozyme. *Nature* **428**, 281–286.

Woods, A., Munday, M. R., Scott, J., Yang, X., Carlson, M. and Carling, D. (1994). Yeast SNF1 is functionally related to mammalian AMP-activated protein kinase and regulates acetyl-CoA carboxylase in vivo. *Journal of Biological Chemistry* **269**, 19509–19515.

Woolston, B. M., Roth, T., Kohale, I., Liu, D. R. and Stephanopoulos, G. (2018). Development of a formaldehyde biosensor with application to synthetic methylotrophy. *Biotechnology and Bioengineering* **115**, 206–215.

Wu, Y., Chen, T., Liu, Y., Tian, R., Lv, X., Li, J., Du, G., Chen, J., Ledesma-Amaro, R. and Liu, L. (2020). Design of a programmable biosensor-CRISPRi genetic circuits for dynamic and autonomous dual-control of metabolic flux in Bacillus subtilis. *Nucleic Acids Research* **48**, 996–1009.

Wu, Y., Du, G., Chen, J. and Liu, L. (2020). Genetically encoded biosensors and their applications in the development of microbial cell factories. In: Engineering of Microbial Biosynthetic Pathways (V. Singh, A. K. Singh, P. Bhargava, M. Joshi and C. G. Joshi, eds). Springer. Singapore, Singapore. pp. 53–73.

Wuichet, K., Cantwell, B. J. and Zhulin, I. B. (2010). Evolution and phyletic distribution of two-component signal transduction systems. *Current Opinion in Microbiology* **13**(2), 219–225.

Xu, P. (2018). Production of chemicals using dynamic control of metabolic fluxes. *Current Opinion in Biotechnology* **53**, 12–19.

Xu, X., Li, X., Liu, Y., Zhu, Y., Li, J., Du, G., Chen, J., Ledesma-Amaro, R. and Liu, L. (2020). Pyruvate-responsive genetic circuits for dynamic control of central metabolism. *Nature Chemical Biology* **16**, 1261–1268.

Yamada, E. W. and Jakoby, W. B. (1960). Aldehyde oxidation: V. direct conversion of malonic semialdehyde to acetyl-coenzyme A. *Journal of Biological Chemistry* **235**, 589–594.

Yanofsky, C. (1981). Attenuation in the control of expression of bacterial operons. *Nature* **289**, 751–758.

Yu, S., Zhao, Q., Miao, X. and Shi, J. (2013). Enhancement of lipid production in low-starch mutants Chlamydomonas reinhardtii by adaptive laboratory evolution. *Bioresource Technology* **147**, 499–507.

Zhang, J., Wang, Z., Su, T., Sun, H., Zhu, Y., Qi, Q. and Wang, Q. (2020). Tuning the binding affinity of heme-responsive biosensor for precise and dynamic pathway regulation. *iScience* **23**, 101067.

Zhang, Q., Bhattacharya, S. and Andersen, M. E. (2013). Ultrasensitive response motifs: basic amplifiers in molecular signalling networks. *Open Biology* **3**, 130031.

Zhang, Y., Werling, U. and Edelmann, W. (2012). SLiCE: a novel bacterial cell extract-based DNA cloning method. *Nucleic Acids Research* **40**, e55.

Zhao, M., Wang, M., Zhang, X., Zhu, Y., Cao, J., She, Y., Cao, Z., Li, G., Wang, J. and Abd El-Aty, A. M. (2021). Recognition elements based on the molecular biological techniques for detecting pesticides in food: a review. *Critical Reviews in Food Science and Nutrition* .

Zhou, L., Grégori, G., Blackman, J. M., Robinson, J. P. and Wanner, B. L. (2005). Stochastic activation of the response regulator PhoB by noncognate histidine kinases. *Journal of Integrative Bioinformatics* **2**(1), 10–22.

Zhou, S., Ainala, S. K., Seol, E., Nguyen, T. T. and Park, S. (2015). Inducible gene expression system by 3-hydroxypropionic acid. *Biotechnology for Biofuels* **8**, 169.

Zhou, S., Catherine, C., Rathnasingh, C., Somasundar, A. and Park, S. (2013). Production of 3-hydroxypropionic acid from glycerol by recombinant Pseudomonas denitrificans. *Biotechnology and Bioengineering* **110**, 3177–3187.

Zhu, H., Mao, X.-J., Guo, X.-P. and Sun, Y.-C. (2016). The hmsT 3′ untranslated region mediates c-di-GMP metabolism and biofilm formation in Yersinia pestis. *Molecular Microbiology* **99**, 1167–1178.

Zhu, Y., Li, Y., Xu, Y., Zhang, J., Ma, L., Qi, Q. and Wang, Q. (2021). Development of bifunctional biosensors for sensing and dynamic control of glycolysis flux in metabolic engineering. *Metabolic Engineering* **68**, 142–151.

Ziegler, C. A. and Freddolino, P. L. (2021). The leucine-responsive regulatory proteins/feast-famine regulatory proteins: an ancient and complex class of transcriptional regulators in bacteria and archaea. *Critical Reviews in Biochemistry and Molecular Biology* **56**(4), 373–400.

Zobel, S., Benedetti, I., Eisenbach, L., de Lorenzo, V., Wierckx, N. and Blank, L. M. (2015). Tn7-based device for calibrated heterologous gene expression in Pseudomonas putida. *ACS Synthetic Biology* **4**, 1341–1351.

Zorraquino-Salvo, V., Quinones-Soto, S., Kim, M., Rai, N. and Tagkopoulos, I. (2014). Deciphering the genetic and transcriptional basis of cross-stress responses in Escherichia coli under complex evolutionary scenarios. *bioRxiv* p. 10595.

Zubay, G., Schwartz, D. and Beckwith, J. (1970). Mechanism of activation of catabolite-sensitive genes: a positive control system. *Proceedings of the National Academy of Sciences of the United States of America* **66**, 104–110.

Zweigenbaum, J., Heinig, K., Steinborner, S., Wachs, T. and Henion, J. (1999). High-throughput bioanalytical LC/MS/MS determination of benzodiazepines in human urine: 1000 samples per 12 hours. *Analytical Chemistry* **71**, 2294–2300.

Appendix I

Supplementary material to Section 5.1 Mathematical modelling

I.1 Additional simulations of the Mannan model

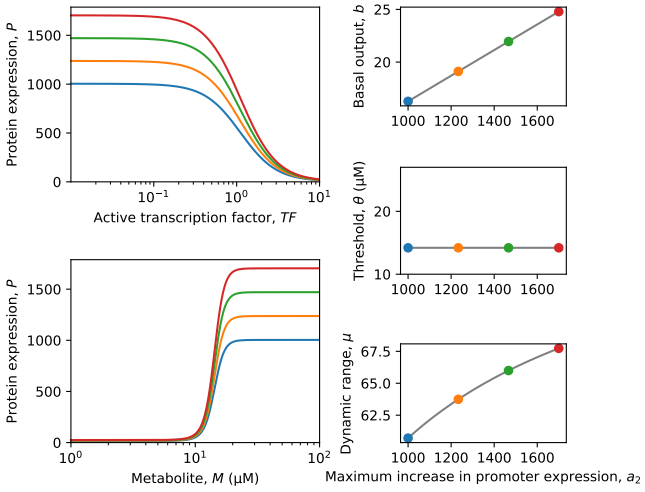

Figure I.1: Model simulations perturbing the maximum increase in promoter expression (a_2) for the repressed-repressor architecture. Data were computed from Equations (2.13), (2.14) and (4.2) to (4.4) with parameters $b_1 = 0.01$, $a_1 = 10.0$, $K_1 = 0.1$, $n_1 = 6.0$, $b_2 = 4.1$, $K_2 = 0.9$, $n_2 = 2.0$, and a_2 values span the range $a_2 = 1000$ to $a_2 = 1700$. Metabolite concentrations (M) are plotted on a logarithmic scale.

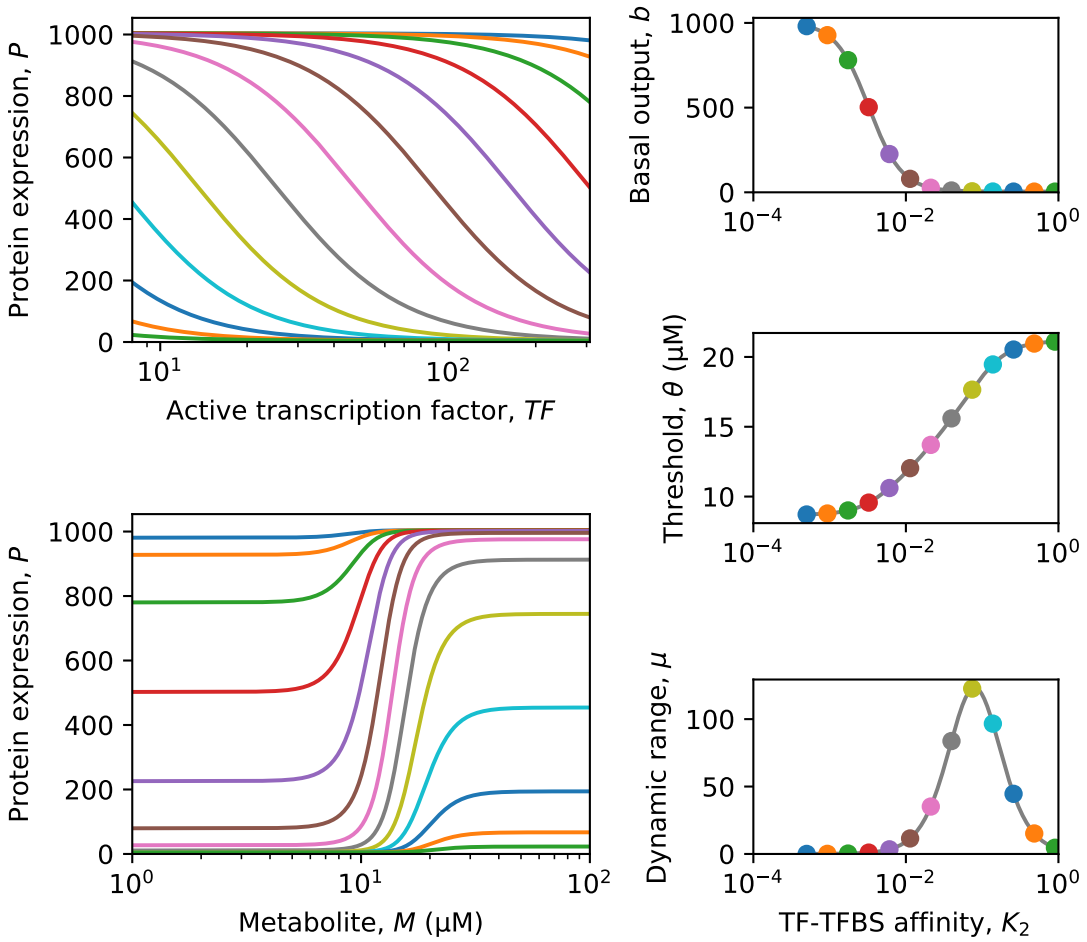

Figure I.2: Model simulations perturbing the transcription factor-TFBS affinity (K_2) for the repressed-repressor architecture. Data were computed from Equations (2.13), (2.14) and (4.2) to (4.4) with parameters $b_1 = 8$, $a_1 = 300$, $K_1 = 0.1$, $n_1 = 6.0$, $b_2 = 4.1$, $a_2 = 1000$, $n_2 = 2.0$, and K_2 values span the range $K_2 = 5 \cdot 10^{-4}$ to $K_2 = 9 \cdot 10^{-1}$. Metabolite concentrations (M) and K_2 are plotted on a logarithmic scale.

I.2 Additional parameter estimation method evaluation results

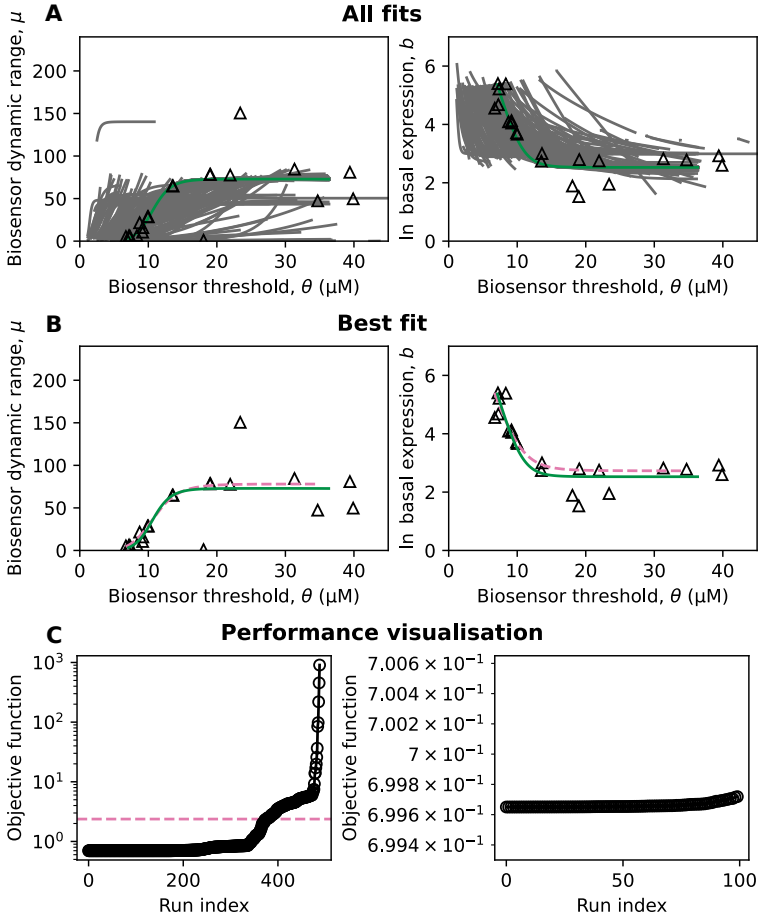

Figure I.3: Results of fitting to parameter estimation method (PEM) evaluation data set 2. **(A)** All 500 fits to the data. The best fit (displaying the lowest objective function value) is plotted as in green, the remainder of the fits is shown in grey. **(B)** Only the best fit is shown in green, along with the ideal noise-free solution (an interpolation of $(\mu_{i,True}, \theta_{i,True}, b_{i,True})$) plotted as a pink dashed line. In panels A and B, triangles represent the dose-response parameters of PEM evaluation data set 2 (the training data). **(C)** Performance visualisation for the PEM evaluation problem. Left panel: the objective function values obtained after 500 runs of least-squares optimisation (run index 0 up to 499) are sorted increasingly. The horizontal dashed line represents C_{pass}, the objective function value computed for the ideal noise-free solution. Right panel: a close-up on the 100 best runs suggests that a global optimum is reached. C_{pass} is beyond the axis limits.

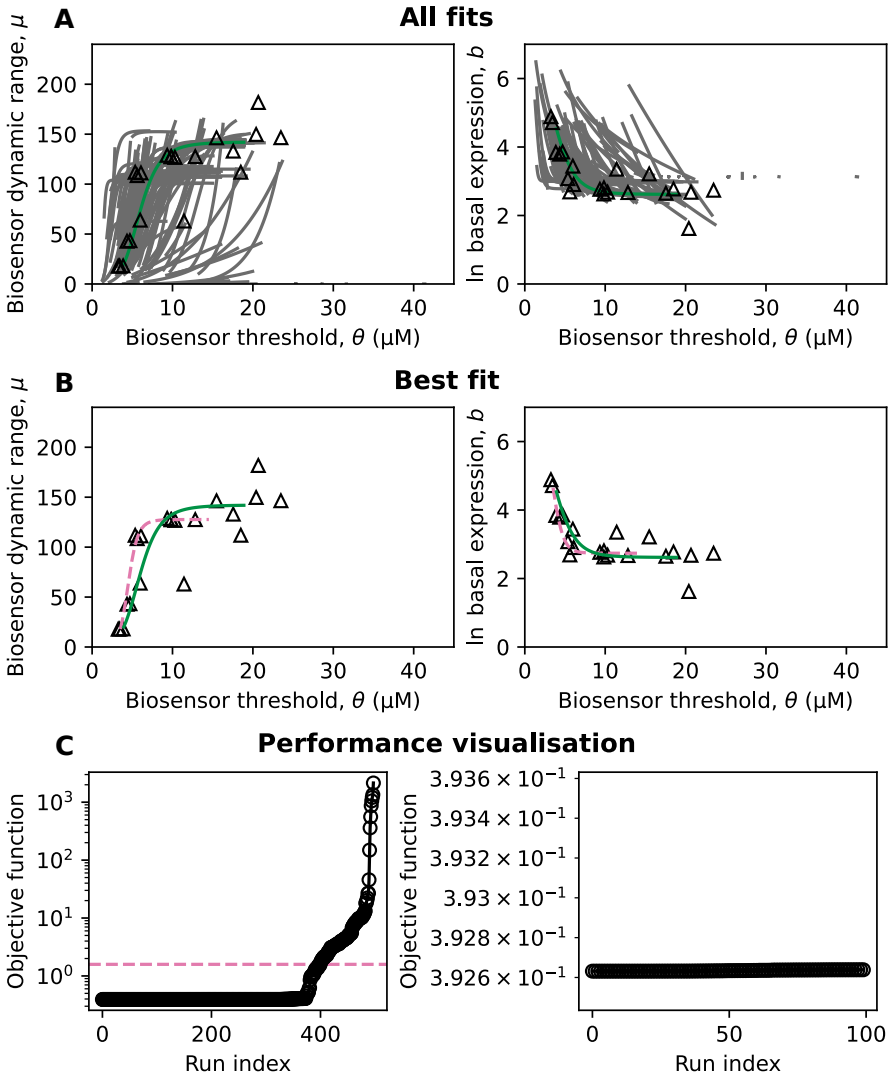

Figure I.4: Results of fitting to parameter estimation method (PEM) evaluation data set 4. **(A)** All 500 fits to the data. The best fit (displaying the lowest objective function value) is plotted as in green, the remainder of the fits is shown in grey. **(B)** Only the best fit is shown in green, along with the ideal noise-free solution (an interpolation of $(\mu_{i,True}, \theta_{i,True}, b_{i,True})$) plotted as a pink dashed line. In panels A and B, triangles represent the dose-response parameters of PEM evaluation data set 4 (the training data). **(C)** Performance visualisation for the PEM evaluation problem. Left panel: the objective function values obtained after 500 runs of least-squares optimisation (run index 0 up to 499) are sorted increasingly. The horizontal dashed line represents C_{pass}, the objective function value computed for the ideal noise-free solution. Right panel: a close-up on the 100 best runs suggests that a global optimum is reached. C_{pass} is beyond the axis limits.

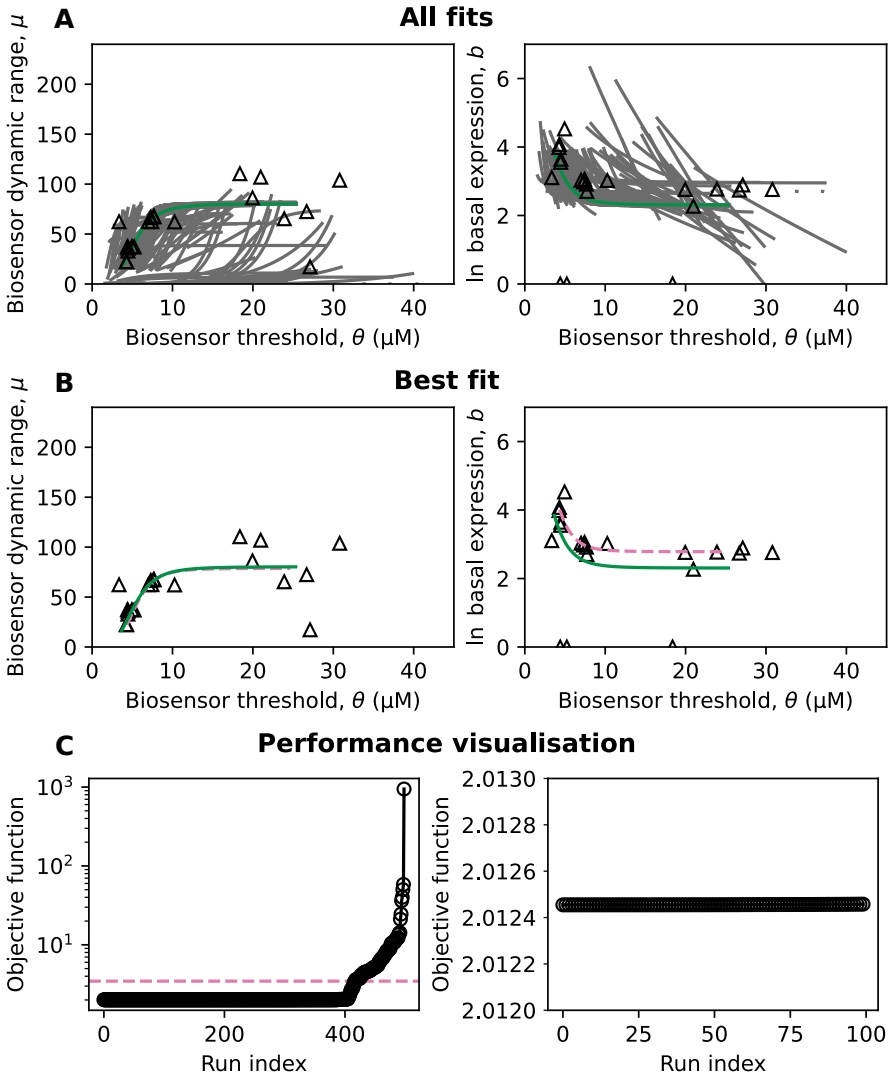

Figure I.5: Results of fitting to parameter estimation method (PEM) evaluation data set 5. **(A)** All 500 fits to the data. The best fit (displaying the lowest objective function value) is plotted as in green, the remainder of the fits is shown in grey. **(B)** Only the best fit is shown in green, along with the ideal noise-free solution (an interpolation of $(\mu_{i,True}, \theta_{i,True}, b_{i,True})$) plotted as a pink dashed line. In panels A and B, triangles represent the dose-response parameters of PEM evaluation data set 5 (the training data). **(C)** Performance visualisation for the PEM evaluation problem. Left panel: the objective function values obtained after 500 runs of least-squares optimisation (run index 0 up to 499) are sorted increasingly. The horizontal dashed line represents C_{pass}, the objective function value computed for the ideal noise-free solution. Right panel: a close-up on the 100 best runs suggests that a global optimum is reached. C_{pass} is beyond the axis limits.

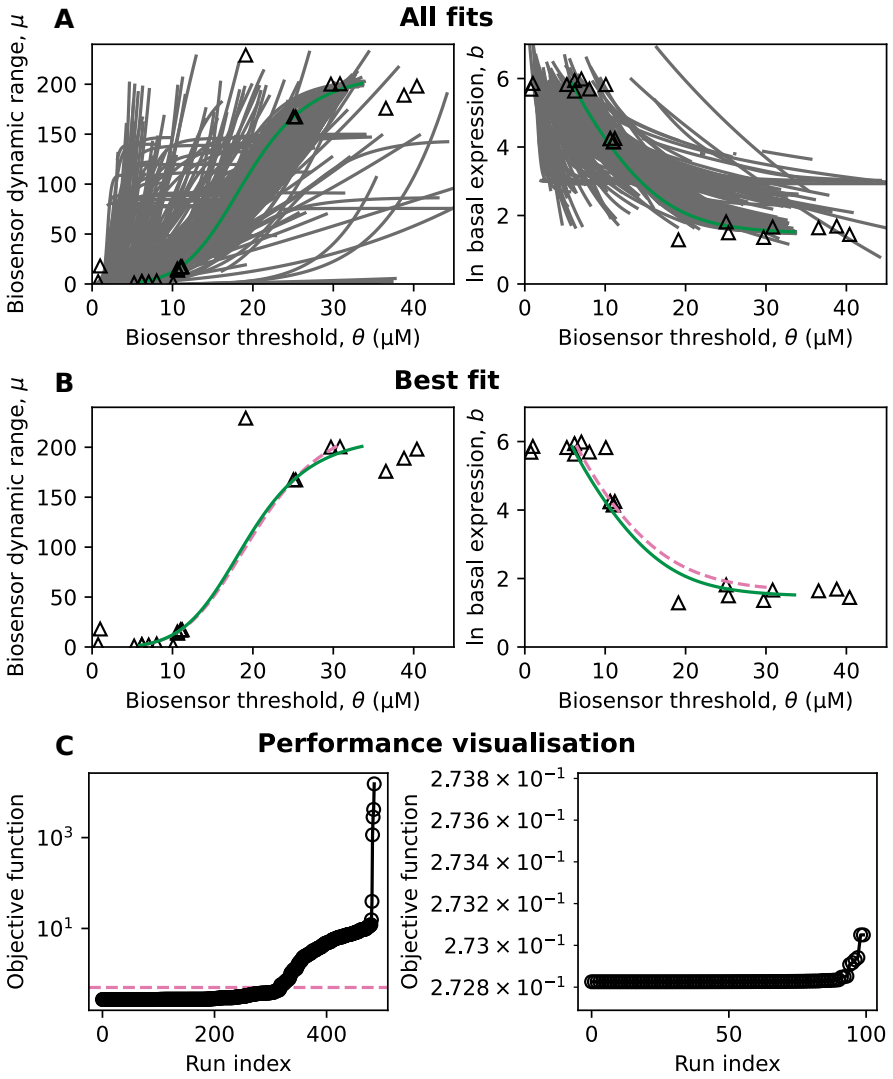

Figure I.6: Results of fitting to parameter estimation method (PEM) evaluation data set 6. **(A)** All 500 fits to the data. The best fit (displaying the lowest objective function value) is plotted as in green, the remainder of the fits is shown in grey. **(B)** Only the best fit is shown in green, along with the ideal noise-free solution (an interpolation of $(\mu_{i,True}, \theta_{i,True}, b_{i,True})$) plotted as a pink dashed line. In panels A and B, triangles represent the dose-response parameters of PEM evaluation data set 6 (the training data). **(C)** Performance visualisation for the PEM evaluation problem. Left panel: the objective function values obtained after 500 runs of least-squares optimisation (run index 0 up to 499) are sorted increasingly. The horizontal dashed line represents C_{pass}, the objective function value computed for the ideal noise-free solution. Right panel: a close-up on the 100 best runs suggests that a global optimum is reached. C_{pass} is beyond the axis limits.

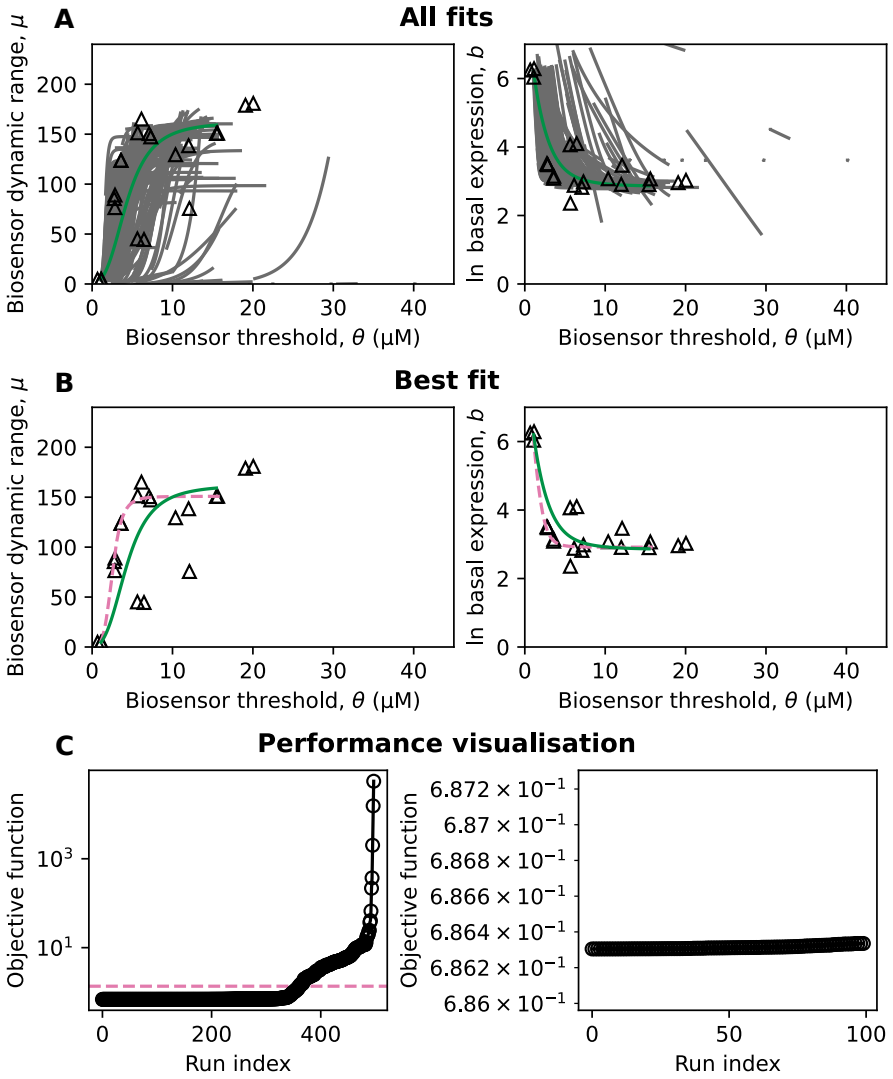

Figure I.7: Results of fitting to parameter estimation method (PEM) evaluation data set 8. **(A)** All 500 fits to the data. The best fit (displaying the lowest objective function value) is plotted as in green, the remainder of the fits is shown in grey. **(B)** Only the best fit is shown in green, along with the ideal noise-free solution (an interpolation of $(\mu_{i,True}, \theta_{i,True}, b_{i,True})$) plotted as a pink dashed line. In panels A and B, triangles represent the dose-response parameters of PEM evaluation data set 8 (the training data). **(C)** Performance visualisation for the PEM evaluation problem. Left panel: the objective function values obtained after 500 runs of least-squares optimisation (run index 0 up to 499) are sorted increasingly. The horizontal dashed line represents C_{pass}, the objective function value computed for the ideal noise-free solution. Right panel: a close-up on the 100 best runs suggests that a global optimum is reached. C_{pass} is beyond the axis limits.

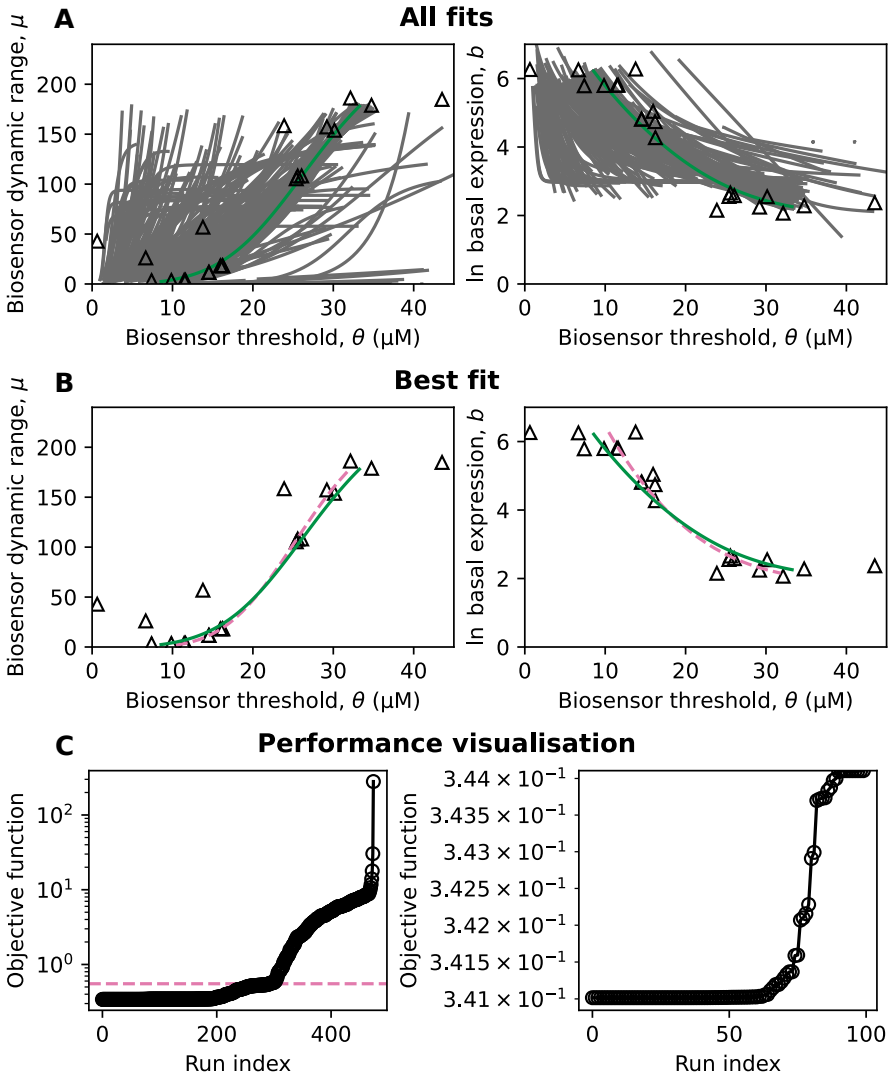

Figure I.8: Results of fitting to parameter estimation method (PEM) evaluation data set 9. **(A)** All 500 fits to the data. The best fit (displaying the lowest objective function value) is plotted as in green, the remainder of the fits is shown in grey. **(B)** Only the best fit is shown in green, along with the ideal noise-free solution (an interpolation of $(\mu_{i,True}, \theta_{i,True}, b_{i,True})$) plotted as a pink dashed line. In panels A and B, triangles represent the dose-response parameters of PEM evaluation data set 9 (the training data). **(C)** Performance visualisation for the PEM evaluation problem. Left panel: the objective function values obtained after 500 runs of least-squares optimisation (run index 0 up to 499) are sorted increasingly. The horizontal dashed line represents C_{pass}, the objective function value computed for the ideal noise-free solution. Right panel: a close-up on the 100 best runs suggests that a global optimum is reached. C_{pass} is beyond the axis limits.

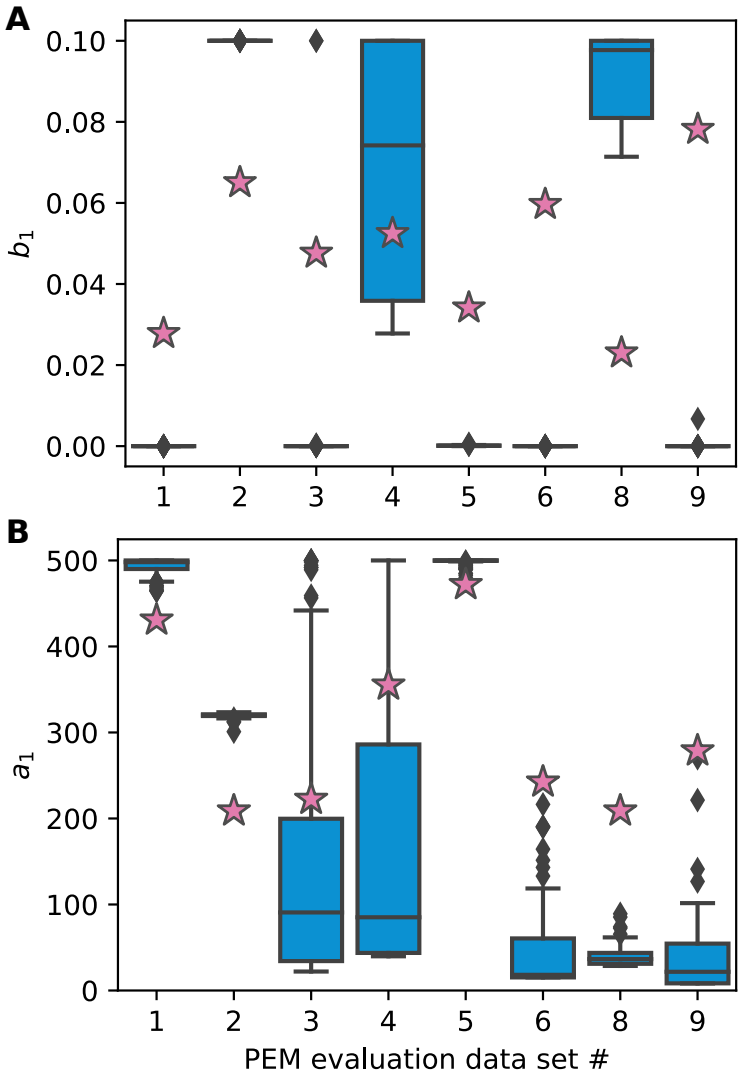

Figure I.9: Distribution of the optimised parameter values. Box plots represent the distribution of the best-fit values of b_1 (the basal level of transcription factor activity, panel **(A)**) and a_1 (the maximum increase in transcription factor activity, panel **(B)**) across the 100 best runs of the optimisation algorithm, except for parameter estimation method (PEM) evaluation data sets 2 (82 best runs), 6 (80 best runs) and 9 (59 best runs) to exclude runs that converged to a distinct objective function value. PEM evaluation data set 7 was not included due to lack-of-fit. Boxes show the quartiles of the parameter distribution while the whiskers delineate the remainder of the distribution within 1.5 interquartile ranges of the nearest box edge. Outliers (parameter values that do not reside within 1.5 interquartile ranges of the nearest quartile) are shown as grey diamonds. Pink stars represent the true parameter values of the ideal noise-free solution of each data set.

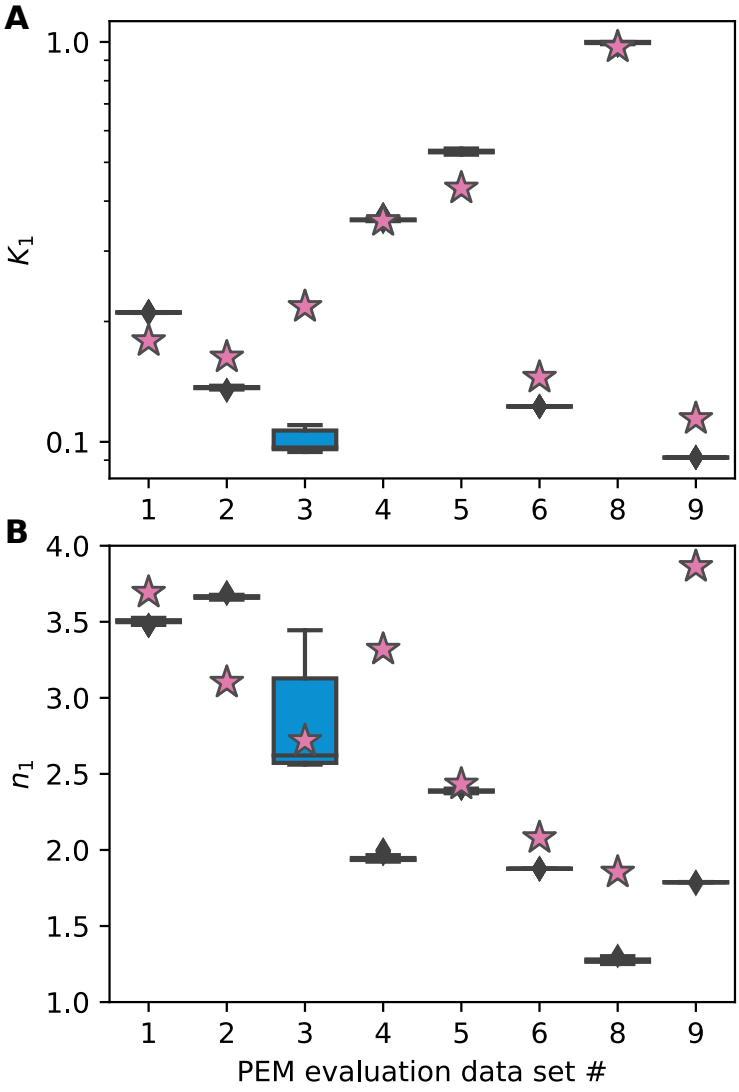

Figure I.10: Distribution of the optimised parameter values. Box plots represent the distribution of the best-fit values of K_1 (the transcription factor-metabolite affinity, panel **(A)**) and n_1 (the sensitivity of transcription factor-metabolite binding, panel **(B)**) across the 100 best runs of the optimisation algorithm, except for parameter estimation method (PEM) evaluation data sets 2 (82 best runs), 6 (80 best runs) and 9 (59 best runs) to exclude runs that converged to a distinct objective function value. PEM evaluation data set 7 was not included due to lack-of-fit. Boxes show the quartiles of the parameter distribution while the whiskers delineate the remainder of the distribution within 1.5 interquartile ranges of the nearest box edge. Outliers (parameter values that do not reside within 1.5 interquartile ranges of the nearest quartile) are shown as grey diamonds. Pink stars represent the true parameter values of the ideal noise-free solution of each data set. K_1 is shown on a logarithmic scale.

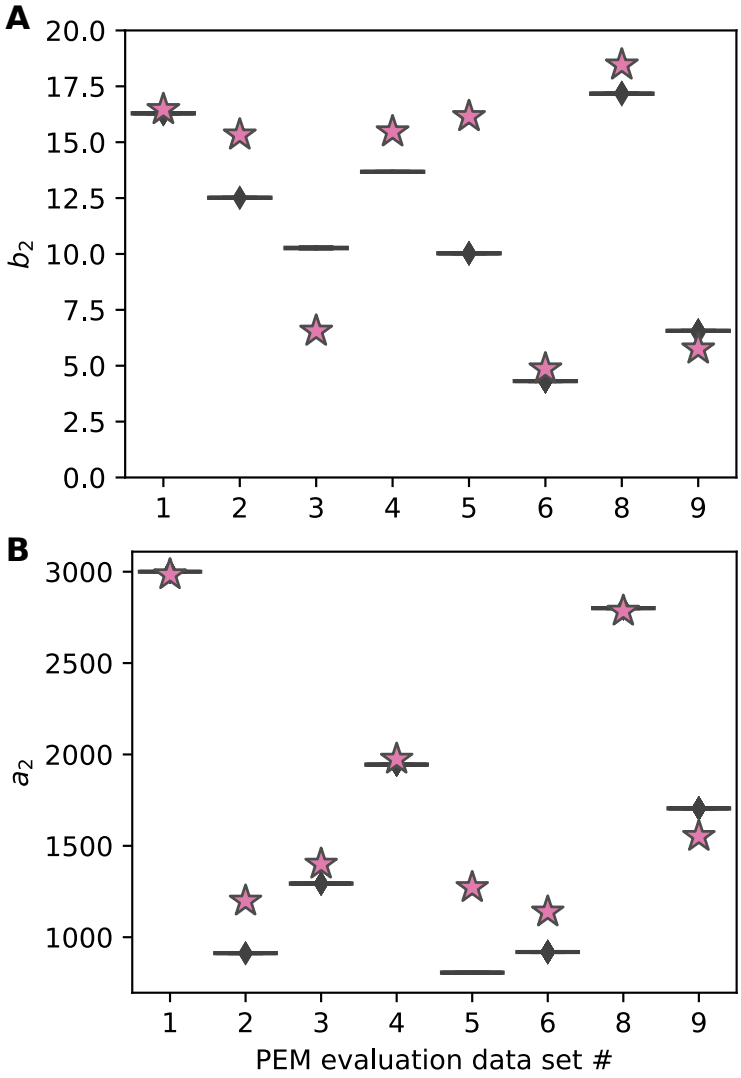

Figure I.11: Distribution of the optimised parameter values. Box plots represent the distribution of the best-fit values of b_2 (the basal level of promoter expression, panel **(A)**) and a_2 (the maximum increase in promoter expression, panel **(B)**) across the 100 best runs of the optimisation algorithm, except for parameter estimation method (PEM) evaluation data sets 2 (82 best runs), 6 (80 best runs) and 9 (59 best runs) to exclude runs that converged to a distinct objective function value. PEM evaluation data set 7 was not included due to lack-of-fit. Boxes showing the quartiles and whiskers delineating the remainder of the parameter distribution within 1.5 interquartile ranges of the nearest box edge are too small to be descried, indicating that b_2 and a_2 were reproducibly determined across all shown data sets. Outliers (parameter values that do not reside within 1.5 interquartile ranges of the nearest quartile) are shown as grey diamonds. Pink stars represent the true parameter values of the ideal noise-free solution of each data set.

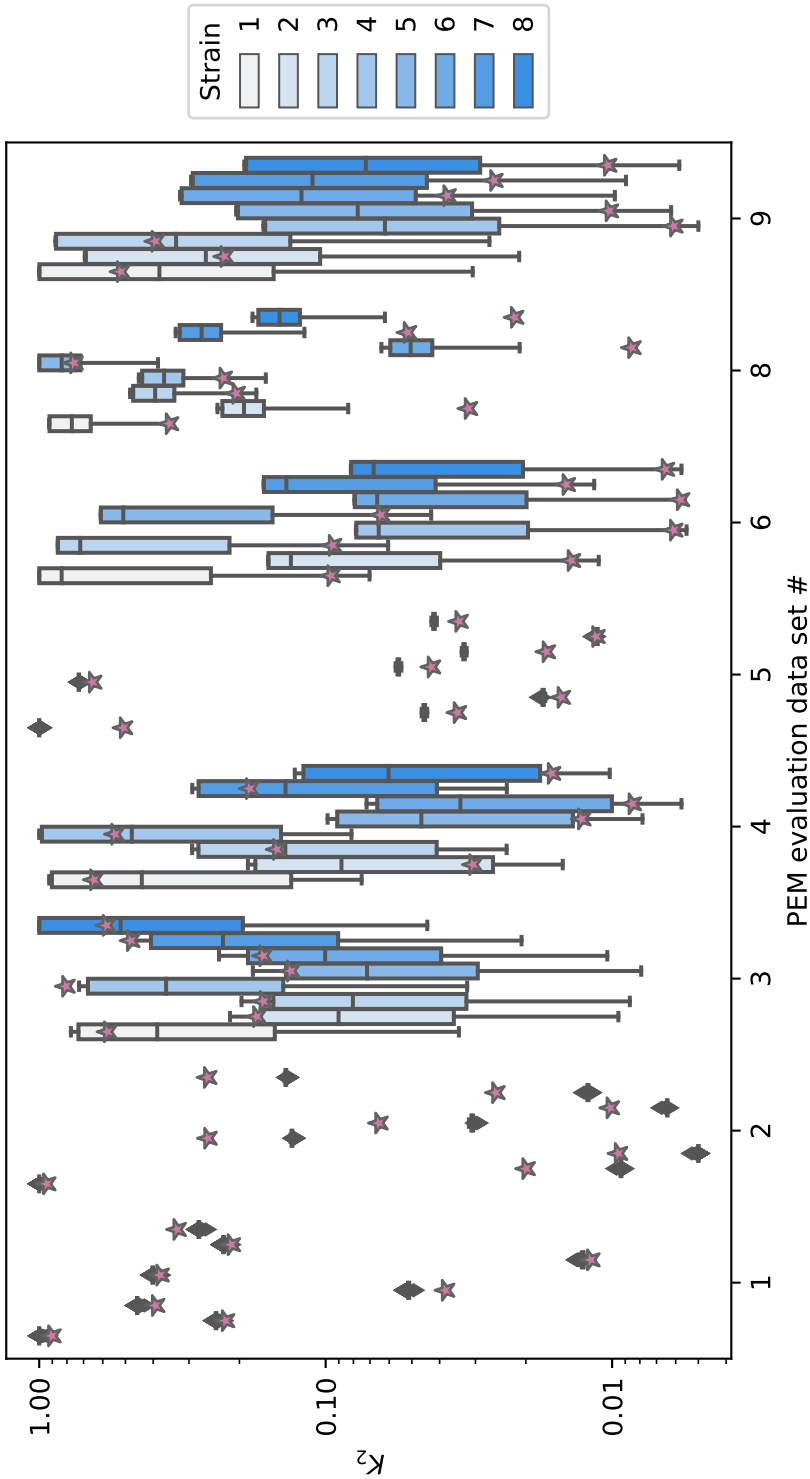

Figure I.12: Distribution of the optimised parameter values. Box plots represent the distribution of the best-fit values of K_2 (the transcription factor-TFBS affinity) for all strains across the 100 best runs of the optimisation algorithm, except for parameter estimation method (PEM) evaluation data sets 2 (82 best runs), 6 (80 best runs) and 9 (59 best runs) to exclude runs that converged to a distinct objective function value. Data set 7 was not included due to lack-of-fit. Boxes show the quartiles of the parameter distribution while the whiskers delineate the remainder of the distribution within 1.5 interquartile ranges of the nearest box edge. Note that the quartiles visibly overlap for all strains in data sets 1, 2 and 5, indicating that K_2 values could be reproducibly determined for these data sets. Outliers (parameter values that do not reside within 1.5 interquartile ranges of the nearest quartile) are shown as grey diamonds. Pink stars represent the true parameter values of the ideal noise-free solution of each PEM evaluation data set. K_2 is shown on a logarithmic scale.

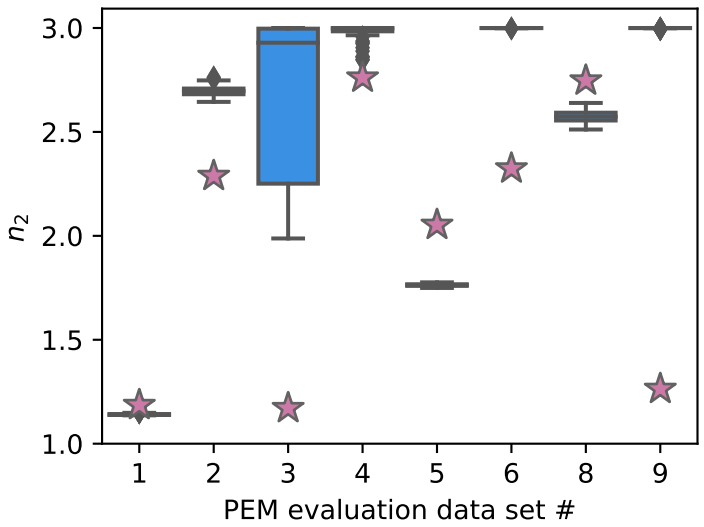

Figure I.13: Distribution of the optimised parameter values. Box plots represent the distribution of the best-fit values of n_1 (the sensitivity of transcription factor-TFBS binding) across the 100 best runs of the optimisation algorithm, except for parameter estimation method (PEM) evaluation data sets 2 (82 best runs), 6 (80 best runs) and 9 (59 best runs) to exclude runs that converged to a distinct objective function value. PEM evaluation data set 7 was not included due to lack-of-fit. Boxes show the quartiles of the parameter distribution while the whiskers delineate the remainder of the distribution within 1.5 interquartile ranges of the nearest box edge. Outliers (parameter values that do not reside within 1.5 interquartile ranges of the nearest quartile) are shown as grey diamonds. Pink stars represent the true parameter values of the ideal noise-free solution of each data set.

Appendix II

Supplementary material to Section 5.2
In vivo characterisation of biosensors

Table II.1: Dose-response parameters for all OapR-based biosensors. Reported values are best-fit values along with the 95% confidence interval between parenthesis, except for [Nar] which is an experimental parameter and μ which was computed from Equation (2.5). Abbreviations: Nar, naringenin; b, basal output; a, maximum increase in expression; μ, dynamic range; θ, threshold; n, Hill coefficient.

Figure	Biosensor	[Nar] (mg/L)	b (RFU)	a (RFU)	μ	θ (mM)	n
Figure 5.14B	*oapTD*	10	2447 (2076 - 2814)	3271 (2344 - 7548)	1.3	1.32 (0.59 - $+\infty$)	0.79 (0.32 - 1.62)
Figure 5.16	*oapTD*	30	58 (46 - 71)	7896 (7145 - 8698)	135	0.27 (0.21 - 0.36)	1.42 (1.29 - 1.58)
Figure 5.19A	mut1.1	30	124 (95 - 154)	1228 (1104 - 1360)	9.9	0.07 (0.05 - 0.10)	1.29 (0.97 - 1.78)
Figure 5.19B	mut1.2	30	109 (94 - 124)	1456 (1340 - 1583)	13	0.30 (0.23 - 0.39)	1.17 (1.01 - 1.36)
Figure 5.20	mut2	30	124 (103 - 145)	408 (348 - 487)	3.3	0.28 (0.15 - 0.56)	0.87 (0.64 - 1.21)
Figure 5.21A	mut3.1	30	51 (30 - 72)	6004 (4484 - 7991)	117	1.13 (0.57 - 2.60)	0.96 (0.80 - 1.18)
Figure 5.21B	mut3.2	30	116 (96 - 136)	139 (83 - 247)	1.2	0.84 (0.06 - $+\infty$)	0.38 (0.18 - 0.92)
Figure 5.21C	mut3.3	30	74 (56 - 92)	879 (771 - 998)	12	0.12 (0.08 - 0.20)	1.18 (0.92 - 1.62)

Table II.2: Dose-response parameters for all AHOS_RS02205-based biosensors. Reported values are best-fit values along with the 95% confidence interval between parenthesis, except for [Nar] which is an experimental parameter and μ which was computed from Equation (2.5). Abbreviations: Nar, naringenin; b, basal output; a, maximum increase in expression; μ, dynamic range; θ, threshold; n, Hill coefficient.

Figure	Promoter	[Nar] (mg/L)	b (RFU)	a (RFU)	μ	θ (mM)	n
Figure 5.22A	*AHOS_RS02205*	10	3313 (3089 - 3535)	4718 (4301 - 5162)	1.4	0.20 (0.17 - 0.24)	1.44 (1.20 - 1.74)
Figure 5.22B	*AHOS_RS02205*	20	2505 (2276 - 2724)	6463 (5822 - 7148)	2.6	0.28 (0.23 - 0.32)	1.63 (1.37 - 1.96)
Figure 5.23A	*AHOS_RS02210*	8	453 (404 - 493)	907 (643 - 1184)	2.0	0.07 (0.05 - 0.11)	1.94 (1.25 - 3.30)
Figure 5.23B	*AHOS_RS02210*	20	312 (277 - 343)	494 (369 - 629)	1.6	0.11 (0.07 - 0.16)	1.62 (1.08 - 2.52)
Figure 5.24A	*barR*	20	4595 (4050 - 5134)	5356 (4567 - 6254)	1.2	0.10 (0.08 - 0.13)	2.0 (1.27 - 3.26)
Figure 5.24B	*SACI_RS10335*	20	13873 (12733 - 14938)	71426 (57150 - 87419)	5.1	0.09 (0.07 - 0.11)	2.99 (2.32 - 4.05)