

DISS. ETH NO. 29180

Optogenetic closed-loop feedback control of the unfolded protein response optimizes protein production in yeast

A thesis submitted to attain the degree of
DOCTOR OF SCIENCES
(Dr. sc. ETH Zurich)

presented by
MORITZ HANS PETER BENISCH
MSc ETH in Chemie- und Bioingenieurwissenschaften, ETH Zurich
born on 11.09.1994

accepted on the recommendation of
Prof Mustafa Khammash, doctoral thesis supervisor
Prof José Luis Martínez Ruiz, co-examiner
Prof Jörg Stelling, co-examiner
Dr Dirk Benzinger, co-examiner

2023

To my wonderful wife . . .

Acknowledgements

First and foremost, I'm extremely grateful to my supervisor, Professor Mustafa Khammash for his support, guidance and the opportunity to do my PhD in his laboratory. Thank you for the continued unwavering support for my project, despite experimental and technical difficulties and the clear vision on how to wrap this project up into a nice paper.

I would like to express my deepest gratitude to Professor José Luis Martinez Ruiz for his support and time, despite never having met in person. Your interest in my work and the input on the biological side have been truly important and uplifting. I am looking forward to meeting you in person in Basel or Copenhagen.

I'm extremely grateful to Professor Jörg Stelling for his input on the yeast biology and modelling of my work. You kindly accepted my "short-term request" to become part of my thesis committee and have helped me since navigate my PhD. The members of your laboratory have further helped me tackle one or the other yeast specific problem.

This endeavor would not have been possible without Dr. Dirk Benzinger for his experimental expertise and all the hours of swotting together over experimental results and interpreting them. Getting me started on yeast biology and the amylase measurements really helped my project and have continuously pushed me forward.

I would like to acknowledge my chair Professor Konrad Tiefenbacher for kindly chairing my doctoral defense.

A big shout out to Dr. Sant Kumar, my only 2-day younger brother, for the time we shared in all of our diverse offices, the help with all kinds of electronics and for becoming a true friend during these 4.5 years sharing the happiness and joy of successful experiments and the frustration for failed ones.

Acknowledgements

I am sincerely grateful to Dr. Stephanie Aoki for her help in all my experiments teaching me how to do molecular cloning, never getting tired of my many questions on how to run digests, gels, PCRs... Thank you for supporting me in the writing of my paper and the work for our review.

I would like to thank Dr. Max Bahls for his openness to share his thoughts and visions of bioprocessing and teaching me how to run a real bioreactor, with all the little tweaks that are not written down in the literature.

Special thanks to all current and previous members of the CTSB group for the joy that you gave me. Coming to work was always easy and fun, knowing that we are all fighting similar biological problems. All 1135 lunches that I was able to share with you were fantastic and taught me a lot about so many diverse fields.

I would like to thank Sebastian Moxon, Esther Eggner and Alem Ahrens from the autoclaving facility for teaching me how to autoclave and responding to my countless (100+) emails about reserving the autoclave for me. Similarly, a big thank you to Jean-Pierre George Meier and Mehmet Nizam for any spare parts and expertise that you provided for getting my bioreactor and especially the gas connections up and running. Finally, a special thank you to Peter Buchmann for building a state-of-the-art light housing for my bioreactor that has always impressed all visitors of the lab.

I am grateful for all the new friends that I have met at D-BSSE with whom I could share my failures, frustration, but also joy and moments of success.

I could not have undertaken this journey without the full support, love and encouragement of my family throughout my academic life at ETH.

Lastly, but most importantly, I would express my uttermost gratitude and love to my wife Anja. You have kept me sane during these past years, showed me love, support and a positive angle in all times of frustration. Exploring the world and mountains with you is the balance that gets me excited over the weekend and I can not imagine a life without you.

Basel, 8th of February 2023

M. B.

Contents

Acknowledgements	i
Abstract (English/Deutsch)	vii
Publications and Contributions	xi
1 Introduction	1
1.1 Current practices in microbial metabolic engineering and protein production	1
1.1.1 Strain Engineering	2
1.1.2 Alternative optimization targets	4
1.2 Dynamic feedback can improve production	5
1.3 UPR and heterologous protein production	6
1.4 Introduction to optogenetics and its application in bioproduction	8
1.4.1 Introduction	8
1.4.2 Deeper Understanding of Biology with Optogenetics	11
1.4.3 Co-cultures with Optogenetics	14
1.4.4 Optogenetic Regulation in Bioproduction	15
1.4.5 Biomaterials with Optogenetics	16
1.4.6 Optogenetic Therapeutic applications	17
1.4.7 Discussion and Conclusion about Optogenetic Applications	19
1.5 Optogenetic Closed-loop Implementations	20
1.6 Objective of this work	20
	iii

Contents

2 Platform development	37
2.1 Developing a photobioreactor	39
2.2 Automatic measurements	41
2.2.1 Development of an automatic sampling and dilution module	41
2.2.2 Beckman Coulter CytFLEX S automation workflow	45
2.3 Amylase sampling and automation	46
2.3.1 Automation of amylase sampling	47
2.3.2 Automation of the amylase offline quantification	48
2.3.3 Maintaining a monoseptic fermentation	48
2.4 Matlab-bioreactor interfacing	50
2.5 Conclusion	53
2.6 Outlook	53
3 Development of a <i>S. cerevisiae</i> strain for optogenetic production of α-amylase and UPR sensing	57
3.1 Methods	59
3.1.1 Media	59
3.1.2 Plasmid construction and strains	60
3.1.3 Culture conditions	61
3.1.4 Flow cytometry	61
3.2 Final design	61
3.3 Results	63
3.4 Conclusion	67
4 Mathematical modelling	75
4.1 Detailed description of the model	77
4.1.1 Description of previous modelling work	77
4.1.2 Novel modelling performed in this PhD thesis	79

4.2	Fitting of the model to the data	83
4.3	Sensitivity analysis	85
4.4	Model inaccuracies	87
4.4.1	Jump behaviour after induction	87
4.4.2	Model inaccuracies in capturing the growth dynamics	89
4.4.3	Constant light inaccuracies	91
4.5	Prediction of control parameters	92
4.5.1	Choice of controller input	92
4.5.2	Choice of controller for initial proof of concept	93
4.5.3	Method to predict PI parameters	93
4.5.4	Comparison of different controller architectures	95
4.6	Conclusion	95
5 Closed-loop bioreactor experiments		99
5.1	Overview and mode of operation	101
5.2	Methods	103
5.2.1	Media	103
5.2.2	Culture conditions	103
5.2.3	Flow cytometry	104
5.3	First closed-loop proof of principle	104
5.3.1	Error rejection and anti-windup	104
5.4	Closed-loop control of the unfolded protein response optimizes the production of a model, hard-to-fold protein	106
5.5	Insights into the time dynamics of α -amylase production	110
5.6	Discussion	111
6 Discussion and Outlook		117
6.1	Strain development	117

Contents

6.2 UPR-yield map 120

6.3 Platform 121

6.4 Modelling and Control 123

6.5 Final Remark 125

Supplement **131**

S1 Supplementary figures 131

S2 Supplementary tables 135

Curriculum Vitae **141**

Abstract

In biotechnological protein production processes, the onset of protein unfolding at high gene expression levels leads to diminishing production yields and reduced efficiency. Here we show that *in silico* closed-loop optogenetic feedback control of the unfolded protein response (UPR) in *S. cerevisiae* clamps gene expression rates at intermediate near-optimal values, leading to significantly improved product titers. Specifically, in a fully automated custom-built 1L-photobioreactor, we used a cybergenetic control system to steer the level of UPR in yeast to a desired set-point by optogenetically modulating the expression of α -amylase, a hard-to-fold protein based on real-time feedback measurements of the UPR, resulting in 60% higher product titers.

This proof-of-concept study paves the way for advanced optimal biotechnology production strategies that diverge from and complement current strategies employing constitutive overexpression or genetically hardwired circuits.

Zusammenfassung

In biotechnologischen Proteinproduktionsprozessen führt das Einsetzen der Proteinfaltung bei hohen Genexpressionsniveaus zu sinkenden Produktionserträgen und geringerer Effizienz. Hier zeigen wir, dass wir mit Hilfe der optogenetischen, *in silico* Rückkopplungskontrolle der ungefalteten Protein Antwort (engl. Unfolded Protein Response (UPR)) in *S. cerevisiae* die Genexpressionsraten auf mittlere, nahezu optimale Werte kontrollieren können, was zu deutlich verbesserten Produkttitern führt. Konkret haben wir in einem vollautomatisierten, spezialangefertigten 1L-Photobioreaktor ein Kontrollsystem verwendet, um das Niveau der UPR in Hefe auf einen gewünschten Sollwert zu steuern. Dies konnten wir erreichen indem wir die Expression von α -Amylase, einem schwer zu faltenden Protein, basierend auf Echtzeitmessungen der UPR, optogenetisch moduliert haben, was zu 60% höheren Produkttitern führte.

Diese Machbarkeitsstudie ebnet den Weg für fortschrittliche, optimale biotechnologische Produktionsstrategien, die sich von den derzeitigen Strategien mit konstitutiver Überexpression oder genetisch verankerten Schaltkreisen unterscheiden.

Publications and Contributions

Moritz Benisch, Dirk Benzinger, Sant Kumar, Hanrong Hu, and Mustafa Khammash. Optogenetic closed-loop feedback control of the unfolded protein response optimizes protein production. *Metabolic Engineering*, 77(32-40), 2023.

Author contributions M.B., D.B. and M.K. conceptualized the study and wrote the manuscript. M.B. built the engineered *S. cerevisiae* strains, performed the characterization, closed-loop experiments and performed the data analysis. H.H. developed the initial modeling framework under the supervision of M.B.. M.B. performed the computational modeling and optimization of PI gains. M.B. and S.K. developed the automated sampling setup. D.B. performed preliminary experiments on optogenetic amylase production and co-supervised the project. M.K. secured funding and supervised the project.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (CyberGenetics; grant agreement 743269).

Moritz Benisch, Stephanie Aoki, and Mustafa Khammash. Unlocking the true potential of optogenetics in yeast applications *In preparation*.

Author contributions M.B., S.A. and M.K. conceived the review focus. M.B. and S.A. wrote the first draft of the review. M.B., S.A. and M.K. contributed to the final version.

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (CyberGenetics; grant agreement 743269).

1 Introduction

Heterologous protein production has been the basis for modern protein-based medicine the last decades. In 2010, the top three sales of biologic drugs in the US fell in the category of monoclonal antibodies (mAbs, 36.0%), hormones (21.1%) and growth factors (19.8%) with a combined sales of 40 billion USD [1]. Since then, this number has only increased and stands now at globally 400 billion USD [2]. A variety of organisms exists that share the majority of production capacity. These include *Escherichia coli*, *Saccharomyces cerevisiae*, *Pichia pastoris*, and a variety of mammalian cell lines, such as Chinese hamster ovarian (CHO) cells. Currently, 56% of biopharmaceuticals are produced in mammalian cells, while a significant portion is still produced in bacteria (24%) and yeast (13%) [3].

1.1 Current practices in microbial metabolic engineering and protein production

Saccharomyces cerevisiae is the most prevalent member of the yeast family. It is recognized for its rapid growth kinetics as well as its capacity to fold and secrete proteins [4]. Other hosts that are readily used in production of proteins and metabolites are *E. coli* [5], *P. pastoris* [6], *Streptomyces* [7], *B. subtilis* [8], chinese hamster ovarian cells (CHO) and *A. niger* [9]. Many engineering strategies have been applied to further improve the production capacity of yeast and increase protein yield. The following sections describe in greater detail some strategies to improve production.

1.1.1 Strain Engineering

The host which is used for production can be genetically altered to improve its capacity for production. This involves modifying the genetic makeup of the microbial strain to optimize its metabolic pathways for the production of the desired protein. Two main ideas exist to achieve strain engineering. Rational design involves the use of computational biology methods and a general understanding of the production pathway to predict which changes in the strain could potentially improve production of the protein. This can be done using techniques like gene knockouts, codon optimization, overexpression of key enzymes, regulatory gene engineering and synthetic genetic circuits. Alternatively, directed evolution involves the iterative selection and amplification of strain variants with desired properties from a large pool of randomly mutated strains. This approach can be used to optimize protein production without an in-depth understanding of the production pathway.

Huang et al. [10] developed a directed evolution, microfluidic droplet screening platform for *S. cerevisiae*. Mutations were introduced in the cells by illuminating the cells with ultraviolet light. Subsequently, single cells of the yeast mutant library containing 10^5 - 10^6 colonies were encapsulated with a fluorogenic α -amylase substrate and sorted. The highly active mutants were screened further in tube cultures and bioreactor fermentation. The best producing strain after two rounds of microfluidic screening resulted in 6-fold improvement in both yield and titer. The authors also whole-genome sequence the best producing strains. Building on this work, Huang et al. [11] more deeply characterized mutations in genes associated to protein secretion. The previous work had identified mutations in the genes *hda2*, *vps5*, *gos1* and *tda3* in the best producing strains. The authors combinatorially deleted these genes from the base strain and observed the effect on amylase yield. Single deletions of these genes resulted in a maximum 2-fold improvement in production, whereas combining all four deletions resulted in a nearly 5-fold increase in α -amylase production. Jakoćinas et al. [12] used CRISPR/Cas9 for selectively mutating *S. cerevisiae* genes for overproduction of mevalonate. This resulted in a strain capable of producing 41-fold more product than the wild-type base strain.

The genetic code determines how the nucleotide sequence of a gene is translated into the amino acid sequence to form a protein. Each amino acid is encoded by a set of three nucleotides called a codon. There are multiple codons that can encode the same amino acid, and different organisms have different preferences for which codons to use. Codon optimization involves analyzing the codon usage of the host organism and choosing a set of codons that are more frequently used by the host organism to replace the original codons in the gene of interest. This process can be done manually or using computer

1.1 Current practices in microbial metabolic engineering and protein production

algorithms [13]. The optimized gene can then be synthesized and inserted into the host organism for expression. Karaođlan and Erden-Karaođlan [14] optimized the pectinase production in *P. pastoris* by varying the promoter and codon optimizing the coding gene for pectinase. The gene originates from *A. niger* and was adapted to the new production host *P. pastoris* resulting in an improvement in yield of 20%. In a different example, Wiedemann and Boles [15] engineered *S. cerevisiae* for optimal biofuel production from l-arabinose. L-arabinose can only be metabolized by yeast after a set of heterologous genes is included from other hosts such as *B. subtilis* or *E. coli*. Usually, these genes are not codon optimized for expression in *S. cerevisiae*. By adapting the codons of all genes involved in the l-arabinose pathway to the highest expressing codon, production rates of ethanol could be increase 2.5-fold highlighting the effectiveness of this approach.

In the production of a metabolite, the availability of the precursor is of key essence to the final production outcome. Higher availability of precursors can improve the yield significantly. Similarly, the availability of co-factors in the final production step, such as chaperones for the folding of proteins can improve production. As an initial step, the production of precursor or co-factor can be overexpressed to improve production. Chen et al. [16] were optimizing the production of the biofuel isobutanol in *S. cerevisiae*. The simplified metabolic pathway from glucose to isobutanol contains the intermediate products glucose \rightarrow pyruvate \rightarrow L-valine \rightarrow isobutanol. The authors set about to improve the availability of the precursor valine by overexpressing the enzymes *Ilv2*, *Ilv3*, and *Ilv5* responsible for the conversion of pyruvate to valine in the mitochondria. In rich YPD media, the increased availability of L-valine resulted in an improvement of isobutanol of 71%. In a different study, Luo and Lee [17] metabolically engineered *E. coli* for the production of benzoic acid. They assessed three naturally occurring metabolic pathways for production of benzoic acid using *in silico* flux response analysis. Afterwards they applied metabolic engineering tools to improve the availability of the precursor L-phenylalanine, finally resulting in a strain capable of producing 2.37 g/L of benzoic acid.

More specifically, overexpression of cofactors for heterologous protein production can target a multitude of steps along the protein processing. Chaperone proteins assist in the proper folding of newly synthesized proteins, and can be overexpressed to improve protein folding and reduce misfolding and degradation. Signal peptides are short amino acid sequences that direct proteins to their appropriate cellular compartment. Optimizing the signal peptide sequence can improve protein secretion and reduce retention. An example demonstrating this strategy is shown by Yao et al. [18] who optimized the production and secretion of *B. stearrowthermophilus* α -amylase in *B. subtilis* by signal peptide optimization and chaperone overexpression. Initially, they screened 173 signal peptides to improve transport and secretion out of the host. The best signal peptide in-

creased activity 3.5-fold. However, the authors found elevated levels of inclusion bodies in the host, hinting towards insufficient availability of chaperones that aid folding. They thus overexpressed two sets of chaperones (GroEL–GroES and DnaK–DnaJ–GrpE) both negatively regulated by *hrcA*, by using a $\Delta hrcA$ strain. This resulted in reduced inclusion bodies and a 42% improvement in extracellular amylase activity. The stability of the mRNA encoding the desired protein can also impact protein production. Genetic engineering can be used to optimize the mRNA stability to increase protein production. Ito et al. [19] tuned the stability of mRNA in *P. pastoris* by altering the terminator sequence. In *S. cerevisiae*, terminators can affect mRNA half-life [20]. The authors thus screened a library of 72 terminator sequences with a maximum 15-fold change. They applied 10 terminator variants to the production of betaxanthin and could see a 7-fold improvement of the most stable terminator in comparison to the weakest one.

1.1.2 Alternative optimization targets

Apart from the engineering of the strain, and protein sequence, the optimization of the process around the cell is also often targeted. Examples can be found around media optimization, the design of the physical bioreactor as well as how the process is run.

The composition of the growth medium can have a significant impact on the productivity of the microbial culture. Optimization of the medium can be done by varying the concentrations of key nutrients, pH, and other growth factors. Roberts et al. [21] developed a new optimized media for production of α -amylase in *S. cerevisiae*. They split the multi-component media into five different categories (glucose, amino acids, YNB, inositol and the nitrogen source) and then used a design of experiments approach to optimize the ratio of these components. In the improved high cell density media, the amylase concentration could be improved 4-fold, although the yield per gram of glucose decreased slightly. Process conditions can also be adapted by changing temperature and pH while keeping the media composition constant. This approach is used by Lin et al. [22] who optimized the production of bioethanol by *S. cerevisiae*. Classically, the pH in fermentations of yeast is controlled to 6.0. The authors tested different conditions varying between 3.0 and 6.0 and found optimal production at pH=5.0, improving the ethanol concentration 3-fold. Interestingly, when the pH is not controlled a further 5% improvement of the concentration is achieved, although further experiments are required to confirm this change.

The design of the bioreactor can also have an impact on production efficiency. Factors like agitation rate, aeration, and temperature control can all be optimized to improve productivity. Lastly, how the production process is run can be adapted to the specific

1.2 Dynamic feedback can improve production

need of the product of interest. Classically, one differentiates between batch, fed-batch and continuous culture (or continuously stirred tank reactor, CSTR). In a batch process all media components and the production host is added at the beginning and the reaction is run until completion. A fed-batch phase can be performed following a batch phase, where high concentration media is fed to the batch culture, increasing cell density further and also potentially increasing product formation. A CSTR is operated in a single vessel with constant inflow of fresh media without cells, and constant outflow of depleted media with cells. This results in lower cell density and product concentration. At long operating times, a steady-state can be achieved in these types of setups. Generally, higher cell densities can be achieved with batch and especially fed-batch systems. If the production host is limited by substrate (substrate inhibition), the process of choice is CSTR, whereas with product inhibition, batch and fed-batch processes are favored. Also, some production strains are genetically not very stable, with loss of plasmid or mutations commonly appearing. As CSTR processes are longer, batch and fed-batch processes are favored. Bayrock and Ingledew [23] studied the production of ethanol by *S. cerevisiae* with different production methods. In a batch process, they were able to achieve ethanol concentrations of 120 g/L compared to only 60 g/L in a CSTR process. Relevant for the biofuel community is also the theoretical conversion rate of substrate into product. Here, a CSTR achieved a conversion rate of close to 90% in comparison to 80% in batch process. This shows how different production setups can address different requirements.

1.2 Dynamic feedback can improve production

The above described strategies rely mostly on gene deletions or mutations that hamper the gene expression or massive overexpression of the target. Intuitively, increasing the rate of gene expression should result in more secreted product. While this is true at low protein expression rates, maximal expression does not necessarily lead to maximal production, but can even reduce yield [24, 25]. Consequently, there exists an optimal gene expression rate that maximizes production. The diminishing rate of protein production can be attributed to an overwhelmed cell machinery, resulting in oxidative stress, product misfolding, inclusion bodies, upregulated endoplasmic reticulum-associated degradation, and stress-induced genomic instability [26–29].

Xu et al. [30] worked on improving the production of fatty acids in *E. coli* by tuning the expression rates. Production of fatty acids in bacteria is a multi-step pathway, where the conversion of substrate to product is achieved by the expression of more than 15 enzymes. The expression strength of those enzymes was combinatorially tuned. For all

enzymes, three levels of expression copy numbers were available (low, medium and high). In order to limit the number of permutations, these enzymes were divided into three subclasses (GLY: upstream glycolysis module, ACA: intermediary acetyl-CoA activation module and FAS: downstream fatty acid biosynthetic module) and expressed from the same plasmid copy number. Optimal production is achieved when ACA is expressed from low, GLY from medium and FAS from high-copy number plasmids. Applying this strategy, they were able to achieve fatty acid titers in fed-batch cultures of 8.6 g/L. This is a nice example highlighting the need of tuning of expression strength rather than deleting or maximally over-expressing targets. The mapping between optimal gene expression rate and optimal protein production is influenced by many different factors, such as growth rate, process stage, and complexity of the protein of interest [31]. Thus, fine-tuning of gene expression rates has the drawback, that a slight change in production conditions can heavily impact production.

One possible solution is to adjust protein expression intracellularly based on stress levels using burden-driven, genetic feedback circuits [32]. Dynamic pathway regulation is for example implemented by Xu et al. [33] for the production of fatty acids in *E. coli*. The main bottleneck in production was found to be malonyl-CoA. The authors genetically rewired the metabolism, so that both source production and consumption of malonyl-CoA was regulated. This allowed the cells to compensate for critical enzymes and redirect metabolic flux towards fatty acid synthesis. The resulting engineered strain improved fatty acid yield 2.1-fold in comparison to uncontrolled fatty acid production. A similar approach was used by Dahl et al. [34] in *E. coli* for the production of amorpha-4,14-diene. A negative feedback was genetically engineered to control farnesyl pyrophosphate production, which led to a twofold improvement. However, for this strategy to work effectively, the synthetic circuits must be carefully calibrated for a specific product and production environment.

To circumvent this, direct *in silico* control on the internal states of cells, similar to how the industry regulates process parameters like pH and dissolved oxygen, can be applied. The feedback strength is then implemented by a computer, rather than genetically engineering the cells. This enables the adjustment of feedback intensity to meet the design specifications, possibly resulting in enhanced protein production.

1.3 UPR and heterologous protein production

For such an *in silico* feedback control loop a measurement of a relevant cellular state is required. One option is to develop a biosensor that is specific for the product of interest as

1.3 UPR and heterologous protein production

shown for small molecules such as muconic acid [35], xylose [36] or branched chain amino acid metabolism [37] as well as for some proteins such as glucose dehydrogenase [38], β -lactamase [39], anti-apoptosis protein BCL-2, IgG1 Fc domain and HER2 receptor [40]. The development of a biosensor with good sensitivity, specificity and a high dynamic range properties is challenging [41] and could limit the applicability of this approach to a large selection of proteins of interest. Thus, we focus our attention on the more general target that is the unfolded protein response (UPR). The UPR coordinates the cellular response in yeast to elevated levels of unfolded proteins in the ER and can trigger the expression of chaperones, cofactors and the ER-associated degradation (ERAD) pathway [42]. It is thus central to the proper folding of proteins.

The first step in protein expression is the synthesis of mRNA from DNA through the process of transcription. After transcription, the mRNA molecule undergoes several modifications, including the addition of a 5' cap and a 3' poly-A tail, as well as the removal of introns. These modifications are necessary for proper mRNA stability, export, and translation. Once the mRNA is processed, it is exported from the nucleus to the cytosol through nuclear pore complexes. In the cytosol, the mRNA is bound by ribosomes, and the process of translation begins, where the sequence of nucleotides in the mRNA is translated into a sequence of amino acids to form a polypeptide chain. As the polypeptide chain is synthesized by the ribosome, it is translocated across the membrane of the rough endoplasmic reticulum (ER) through a channel formed by the translocon complex. This process is known as co-translational translocation and ensures that proteins are translocated into the ER lumen as they are being synthesized [43]. Once the protein enters the ER lumen, it undergoes folding and post-translational modification, which includes the formation of disulfide bonds, glycosylation, and proteolytic cleavage. These modifications are critical for the proper folding, stability, and function of the protein. The unfolded protein response targets specifically this step in the endoplasmic reticulum.

The UPR is controlled as follows. In the inactive state, Hac1 mRNA is constitutively transcribed. Unspliced Hac1 mRNA contains an intron forming a stem loop to the 5' UTR, preventing full translation and additionally reducing translation rate. Ire1 (Inositol requiring enzyme 1) is a sensor in the endoplasmic reticulum. The cytosolic side of Ire1 consists of a kinase and an endoribonuclease. Upon presence of unfolded proteins, those unfolded proteins bind to Ire1 and make them cluster on the ER-lumen-cytosol surface. Due to the proximity of the kinases of Ire1, the kinases transautophosphorylate each other and thus result in an activation of the endoribonucleases, which excise the Hac1mRNA intron [42]. Ribosomes are able to now efficiently translate the spliced Hac1mRNA to functional Hac1p transcription factor. Downstream targets of Hac1p are

chaperones (e.g. Kar2), oxidoreductases, glycosylating enzymes, and ER degradation components.

The UPR has been genetically modified to enhance protein production. Valkonen et al. [45] studied the effect of overexpression of the UPR master regulator *Hac1* in *S. cerevisiae* on production of α -amylase. Initial deletion of *Hac1* resulted in a 3-fold decrease of amylase production. In a next step, the authors expressed *Hac1* under the *PGK1* promoter, resulting in a 70% increase in amylase secretion. A similar approach was applied in the production host *A. niger*. Overexpression of *hacA* resulted in a 7-fold improvement of *Trametes versicolor* laccase production. Huang et al. [46] overexpressed *Hac1* from the *AOX1* promoter increasing amylase activity 621%. As the *AOX1* promoter was also used for the expression of amylase transcript, the promoter for *HAC1* overexpression was replaced by a *GAP* promoter further increasing activity by 68.2%. The UPR also integrates over disturbances stemming from growth kinetics, nutrient availability and protein folding state. We thus believe that implementing the sensing channel through the UPR allows this approach to be widely generalizable beyond the expression of α -amylase. As the basis for sensing, we use the promoter topology from Merksamer et al. [47]. The reporter consists of a minimal *CYC1* promoter and four unfolded protein response elements (UPRE) in series. It is known that the master UPR regulator *Hac1* binds to these elements, so that this is a good readout for UPR activity [48].

1.4 Introduction to optogenetics and its application in bio-production

This section stems from the review on "Unlocking the potential of optogenetics in microbial applications" by Moritz Benisch, Stephanie Aoki and Mustafa Khammash that is currently in preparation.

1.4.1 Introduction

Optogenetics is an interdisciplinary and rapidly growing field that combines optics and genetics to precisely control cellular processes with light [49, 50]. By expressing light-sensitive proteins in cells, researchers can manipulate their activity with high precision, offering a range of advantages, such as temporal reversibility and spatial control, while being fast, often biologically orthogonal, and cheap. While often highlighting these advantages, however, many researchers use optogenetics solely as a replacement for chemical inducers without taking advantage of its desirable features.

1.4 Introduction to optogenetics and its application in bioproduction

Synthetic biology optogenetics is a powerful tool that allows researchers to control the behavior of cells and organisms using light. Some of the first microbial synthetic optogenetic constructs were published approximately twenty years ago in yeast [51] and bacteria [52]. By engineering cells to express light-sensitive proteins, researchers can precisely control a wide array of cell behavior including gene expression, signaling pathways, metabolism, and even motility [53]. These capabilities make optogenetics an attractive tool for identifying and characterizing biological models. For example, researchers can use optogenetics to activate or inhibit specific signaling pathways to study their roles in different cellular processes.

In order to clearly delineate use-cases from each other, we classify applications in a matrix consisting of induction, and measurement scheme and temporal and spatial usage of optogenetics (figure 1.1). We believe that this classification will allow researchers to more precisely define their induction system's requirements and better use optogenetics in their studies.

The induction scheme describes how light is used to regulate the activity of the biological system. In many publications, optogenetic tools are used similarly to chemical inducers, as it is just used in bulk cultures with or without light. This approach results in no significant differences between the two methods, as varying the light intensity has a similar effect on induction strength as varying the concentration of a chemical inducer [54]. However, reversible induction experiments using chemical inducers require a thorough washing step to eliminate any residual traces of the inducer. This can limit their application in terms of speed and throughput. In contrast, optogenetics offers several advantages, such as easy modulation of light intensity and the ability to generate complex time-varying induction profiles on a fast time scale, including single light pulses, pulse-width modulation (PWM), and arbitrary light inputs.

PWM is a special case of time-varying induction that requires further study. In an experiment conducted by Bennett et al. [55], *S. cerevisiae* strains were subjected to periodic changes in the carbon source, and the metabolic response of the cells was tracked. It was found that cells acted as a low-pass filter with a maximum response frequency of $0.88 h^{-1}$. Given an approximate growth rate of $0.2h^{-1}$, any input changes that occur faster than approximately 1/5 of the growth rate of a microbiological organism are filtered out and behave like constant inputs. Therefore, the maximum modulation frequency that is distinguishable from constant input is limited and should be taken into consideration.

In the spatial domain, optogenetics can be applied with ease to specific population

Chapter 1. Introduction

regions, single-cells or subcellular compartments of the cells rather than the whole bulk. This spatial specificity can be achieved using a variety of methods, including photomasks, which are physical masks that can be placed over the sample to block or direct light; and digital projection systems, which use computer-controlled light sources, such as digital micromirror devices (DMDs), to project patterns of light onto the sample [56]. In all cases, it is also possible to combine spatial control with temporal control, thereby increasing the range of induction schemes.

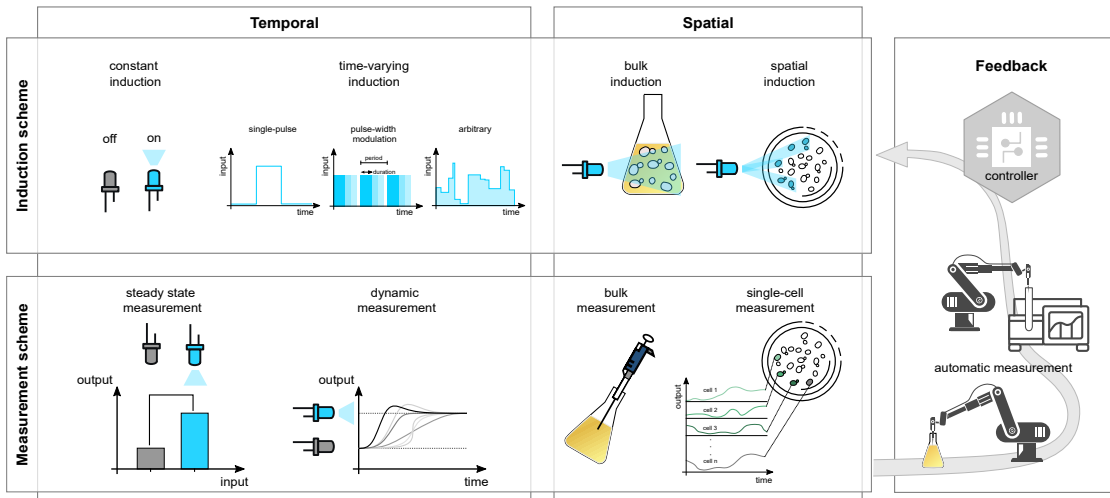


Figure 1.1: Classification matrix for applications. **(Top row)** Induction of cells can be achieved temporally either with constant or time-varying light input. Spatially, we discern between induction of the entire bulk or by spatial (i.e. population, single-cell or subcellular) induction. **(Bottom row)** Measurement of cellular or global parameters in experiments can be performed using either steady-state measurements or dynamics measurements (temporally) and in bulk or at the single cell level. **(Right column)** Feedback can be performed with optogenetics through automatic measurements and time-varying induction of the cells.

The measurement scheme is typically divided into two categories: steady-state and dynamic measurements. While both types of measurements are useful, dynamic measurements can be particularly useful to discriminate between model topologies and for parameter inference. On the other hand, steady-state measurements can be sufficient for certain applications, particularly when combined with time-varying or spatial induction. Nonetheless, the higher information content and accuracy of dynamic measurements make them a valuable tool, and they should be considered as an important component of any experimental design.

Another advantage optogenetics offers is its use in real-time closed-loop feedback control, enabling precise manipulation and study of cellular processes. This approach relies on automated acquisition of dynamic output measurements of cellular state, followed by a time-varying light input computed by a controller algorithm based on these mea-

1.4 Introduction to optogenetics and its application in bioproduction

surements. Optogenetics is particularly well-suited for feedback applications, as light intensity can be instantaneously adjusted both positively and negatively, allowing for a high degree of control.

Optogenetics is a rapidly expanding field with numerous comprehensive reviews of optogenetic tools already available (e.g. Figueroa et al. [57]; Ohlendorf and Möglich [50]). In this short review, we highlight the unique and desirable abilities of microbial optogenetics for spatiotemporal control, and present a select number of exciting and innovative applications that arise from this, including its potential for understanding biology, co-cultures, protein and metabolic production, biomaterials, and therapeutics. By presenting some of the latest research and emerging trends in this area, we aim to complement existing literature and inspire the field to fully explore the potential of microbial optogenetics in a diverse range of applications.

1.4.2 Deeper Understanding of Biology with Optogenetics

In dynamic experimental settings, optogenetics can give a faster response compared to chemical inducers. The import of inducers such as galactose or β -estradiol into cells depends on diffusion and can take longer than optogenetic activation. This rapid response time is a key reason for using optogenetics to study fast-acting signalling pathways and cellular processes. For example, Sumner et al. [69] investigated the brownian motion dynamics of chromosomal loci towards the nuclear pore complex and confirmed initial computational simulations of repositioning dynamics to the nuclear periphery. Similarly, light-inducible nuclear exporter (LINX) and light-activated nuclear shuttle (LANS) were used to control the localization of proteins inside the nucleus, enabling researchers to study the temporal dynamics of histone ubiquitination and methylation [70, 71].

Subcellular compartmentalization is a powerful tool to locally increase the concentration of enzymes inside the cell [58, 72, 73]. Bracha et al. [58] studied phase separation in yeast and mammalian cells using local oligomerization induced by photo-activatable iLID domains (figure 1.2a) [74]. Full-cell oligomerization in yeast and subcellular induction of oligomerization in mammalian epithelia cells allows insights into diffusion dynamics and allow predictions of liquid-liquid phase separations *in vivo*.

Harrigan et al. [59] studied internal feedback structures in yeast, by deleting feedback components in the pheromone response pathway and re-introducing this feedback with closed loop optogenetic control (figure 1.2a). With this platform, they were able to identify not only which nodes are involved in biological signaling pathways, but additionally what dynamic requirements are computed in each biological node. This in-depth analy-

sis of biological signalling dynamics can only be achieved with the unique features that optogenetics offer.

Single-cell feedback requires additional algorithm to allow single-cell tracking and measurement. Once implemented, this unique approach can reveal dynamics that may be hidden in bulk, as e.g. periodic oscillations in single-cell parameters are not lost to population dynamics. Initial work of such a platform for *S. cerevisiae* was implemented by Rullan et al. [75] and has since been used to study the emergence of checkerboard patterns in cell signaling systems [76]. Furthermore, Kumar et al. [77] designed and prototyped biomolecular controllers, testing non-ideal circuit behaviors and qualitatively demonstrating improvements in controller function with certain network modifications.

By precisely controlling the timing and location of cellular activity through the use of light-sensitive proteins, optogenetics provides a means to perturb and measure complex biological networks with high spatiotemporal resolution. This has enabled researchers to develop new experimental techniques for parameter inference, network topology identification, and quantification of information flows within and between cells.

Davidovic et al. [78] used a platform previously developed in [60]. The red/green CcaS/CcaR optogenetic gene expression system [79] in *E. coli* is used in combination with a mother machine to spatially deliver different dynamic light patterns/perturbations in parallel to individual cells and observe their output response with frequent measurements over time (figure 1.2b). The application of multiple perturbations allows for the generation of high-throughput dynamic data for inference of parameters that may not otherwise be identifiable from a single perturbation experiment. Another example of parameter identification leverages trains of reversible light pulses to measure the maturation times of a set of fluorescent proteins, overcoming the limitations of translational inhibitors [80].

Lee et al. [81] measured the response of an epigenome regulator to AM and PWM optogenetic inputs to interrogate the maximum information content transferred through a single promoter and unveils that chromatin regulators tune the maximum information. Benzinger et al. [61] leveraged the difference in dark-state reversion kinetics of two different mutants of EL222 [82] to construct complex biological circuits. The fast-reverting mutant was mapped to a repressor, the slow-reverting mutant to an activator. After constant light induction where repressor and activator cancel each other out, a falling edge triggers the unbinding of the fast-reverting repressor, while the slow-reverting activator stayed bound to allow the generation of a falling-edge detector. This work demonstrate the unique possibility to leverage electrical engineering concepts in synthetic biology.

1.4 Introduction to optogenetics and its application in bioproduction

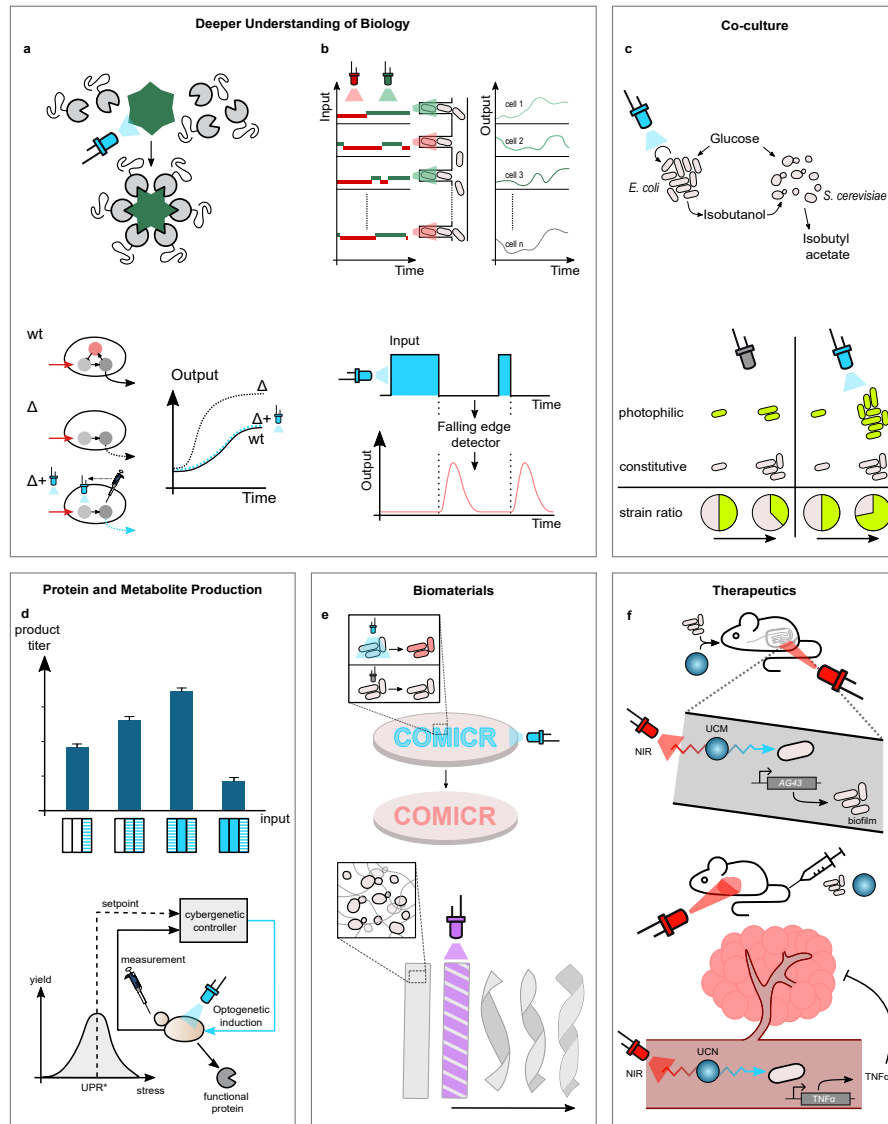


Figure 1.2: Optogenetic applications in *S. cerevisiae* and *E. coli*. (**Deeper Understanding of Biology**) **a.** Bracha et al. [58]: Spatial induction of oligomerization using iLID. Harrigan et al. [59]: Understanding dynamic requirements of feedback in *S. cerevisiae*. **b.** Chait et al. [60]: Depicted is the usage of a red/green optogenetic system to generate high-throughput data in a mothermachine. Ben-zinger et al. [61]: Optogenetic usage in complex information processing applications. (**Co-cultures**) **c.** Lalwani et al. [62]: Co-culture of *E. coli* and *S. cerevisiae* for the production of isobutyl acetate. Gutiérrez Mena et al. [63]: Long-term stabilization of a two-strain bacteria culture using dynamic cybergenetic feedback. (**Protein and Metabolite Production**) **d.** Zhao et al. [64]: Sequential optimization of process conditions in the production of metabolites. Benisch et al. [65]: Closed-loop control of the unfolded protein response maximizes protein production. (**Biomaterials**) **e.** Optogenetic patterning for bacteriography applications. Rivera-Tarazona et al. [66]: 3D patterning of a hydrogel with embedded yeasts. (**Therapeutics**) **f.** Cui et al. [67]: Conversion of near infrared light (NIR) to blue light by upconversion microgels (UCM) optogenetically activates the expression of AG43 and subsequent biofilm formation of *E. coli* in the colon for colitis treatment. Pan et al. [68]: Conversion of near infrared light (NIR) to blue light by upconversion nanoparticles (UCN) optogenetically activates the expression of TNF α for cancer treatment.

1.4.3 Co-cultures with Optogenetics

Co-cultures have become increasingly important in synthetic biology and bioproduction. One key advantage of co-cultures is the ability to perform complex tasks that are not possible with single-cell systems, by distributing the task among the members of the co-culture. For example, they can be engineered to perform multi-step biosynthetic pathways, where each co-culture member is responsible for a specific reaction in the pathway. Additionally, they can be used to study naturally occurring co-culture systems, mimicking the complex microbial interactions that occur in the gut. This can provide insights into the metabolic pathways and signaling networks that govern the gut microbiota and its interactions with the host. Stabilizing synthetic co-cultures of multiple microorganisms is difficult as a slight growth advantage of one member will lead to outcompetition of all other organisms. This section details the applications of optogenetics to stabilize and use synthetic co-culture systems.

Lalwani et al. [62] achieve a short-lived, unstable co-culture system of *E. coli* and *S. cerevisiae* by finely tuning the inoculation ratios of both strains. The optogenetic pDusk/pDawn system [83] controls the growth rate of *E. coli*, which extends the inoculation ratios that lead to co-cultures. It should be noted, that optogenetics here is used with high-frequency PWM acting chemical inducer-like. They applied this complex co-culture in the synthesis of isobutyl acetate. Production of isobutanol from glucose is performed in *E. coli* followed by the conversion of isobutanol and acetyl-CoA to the final product in yeast (figure 1.2c).

In comparison to the short-lived, unstable co-culture system, Aditya et al. [84] successfully maintained a two-component yeast system for up to five days. They engineered a single yeast strain, to optogenetically express the Cre recombinase and recombine its genome from mCerulean to mNeonGreen production. As reversible differentiation is not possible with this system, they resorted to re-introducing the original uninduced mCerulean producer by continuously feeding it with fresh media. The Achilles' heel of this approach is the need to replenish one type of cell continuously to maintain the co-culture.

Gutiérrez Mena et al. [63] move another step further and achieve and maintain arbitrary ratios of two strains of *E. coli* using an automated platform and *in silico* feedback control. Opto-T7RNAP [85] is used to regulate expression of the antibiotic resistance gene chloramphenicol acetyltransferase in response to blue light (also done in [86]). The higher the expression, the faster the strain grows in the presence of chloramphenicol (figure 1.2c). By regulating only one strain, it was possible to dynamically tune and

stabilize the co-culture ratio for 40 hours.

1.4.4 Optogenetic Regulation in Bioproduction

For bioproduction applications, optogenetics allows one to control gene expression and metabolic pathways in a non-invasive and reversible manner. This allows for precise control over the production of desired products, such as pharmaceuticals, biofuels, and industrial enzymes. Additionally, optogenetics can improve the efficiency and scalability of bioprocessing by allowing for real-time monitoring and control of cellular processes. Overall, the use of optogenetics in bioprocessing and bioproduction has the potential to revolutionize the field by enabling more precise and efficient control over biological processes.

Zhao et al. [87] are pioneers in the field of optogenetic bioproduction demonstrating the first optogenetically-driven bioproduction process in lab-scale bioreactors. The authors engineered *S. cerevisiae* to optogenetically control the expression of *PDC1* required for ethanol biosynthesis and growth. Their final 2-stage process consisted of a full-light induced biomass forming phase followed by intermittent-light (PWM) production phase taking advantage of the tunability and reversibility of optogenetics resulting in six-fold higher isobutanol formation. A similar approach was also implemented in *E. coli* to control between light induced growth and synthesis of muconic acid [88] and polyhydroxybutyrate [89]. A variety of research groups have started to use optogenetics in bioproduction scenarios in larger volumes [90, 91], but are treating the optogenetics like a chemical inducer without the benefits of reversibility, or speed. Care also needs to be taken, that the product of interest is not photosensitive as in the case of Duplus-Bottin et al. [92], where the use of optogenetics should be avoided.

Optogenetics can be used to increase the degrees of freedom in a production process. By combinatorially testing the induction timing and type (no, PWM, constant light), product formation can be maximized (figure 1.2d, [64]). Zhao et al. [64] applied this principle to the production of lactic acid, isobutanol and naringenin using amplification circuits. Cheng et al. [93] used OptoLAC [62] in *E. coli* to dynamically switch between the production of photoenzyme fatty acid photodecarboxylase from *Chlorella variabilis* and photocatalysis in a one-pot system for asymmetric decarboxylation and increased yield of L-phosphinothricin. To apply more targeted optimization of process parameters e.g. design of experiments [21] or the use of deep learning [94], higher throughput techniques for optogenetic application and product measurement will be necessary [56].

The idea of using closed-loop optogenetic feedback control for optimal bioproduction

has been extensively discussed in the community [95, 96]. Benisch et al. [65] and a recent preprint [97] demonstrate closed-loop control of the unfolded protein response (UPR) as a mean to maximize production of proteins (figure 1.2d). Automation of the measurement of UPR stress resulting from recombinant expression of a protein of interest allows for closed-loop feedback control. Adjusting the rate of transcription optogenetically resulted in 60% higher production of the secreted protein of interest.

Another application is presented in [98], improving the catalytic rate of the TEV protease, otherwise inaccessible by conventional tools. Yeast cells expressing a library of TEV proteases under Cry2-Cib1 [99] control were exposed to short light pulses, so that only the catalytically fast TEV versions would produced elevated signals. They were subsequently enriched and sorted by FACS improving the TEV catalytic rate 5.4-fold. They eventually used this optimized protease in FLARE [100] and SPARK [101] tools reducing the processing time for the protease from 10-15 minutes to 30 seconds, potentially having a significant impact on the biotech industry.

1.4.5 Biomaterials with Optogenetics

Optogenetics can also be used in the design and production of biomaterials [102]. By using optogenetic tools, scientists can precisely manipulate the metabolic pathways of microorganisms, leading to the production of specific molecules that can be used to build biomaterials with desired properties. For example, microbes can be programmed to synthesize proteins that can self-assemble into complex structures, or to produce bioplastics that are biodegradable and environmentally friendly. Additionally, optogenetics can be used to control the growth and differentiation of microbial communities, enabling the construction of biofilms and other complex structures with defined shapes and functions. Light can be spatially targeted at specific areas, single cells or subcellular regions of the new material with the help of DMDs or photomasks.

The initial transfer of an optogenetic tool (Cph8) into *E. coli* Levskaya et al. [52] also included spatial patterning of an agar plate (figure 1.2e). Since this initial demonstration of optogenetically induced bacteriography, more complex image generation has been demonstrated by a variety of groups [54, 84, 103, 104]. While a nice demonstration of the spatial property, no new information is unravelled from this spatial patterning. A more advanced technique for bacteriography is shown by Frangipane et al. [105] who use proteorhodopsin [106] in *E. coli* to control cell density in response to light by regulating proton-motive force-dependent motility. The authors were able to both spatially and temporally control complex patterning within a time scale of minutes in contrast to the hours required for adhesion-based biofilm lithography. Additionally, they observed

1.4 Introduction to optogenetics and its application in bioproduction

and addressed memory effects in light response by using feedback control to improve the spatial resolution. Combining the speed and accuracy of this system with biofilm lithography and the production of polymers such as bacterial cellulose could lead to the development of more complex biomaterials.

Huang et al. [11] and Walker et al. [107] move in the direction of real-world applications, by expressing genes in *E. coli* optogenetically, which allows cells to attach to surfaces. In addition, Moser et al. [108] demonstrate that surfaces such as fabric and plastic are suitable for *E. coli* attachment, expanding the range of material types that can be developed. Jin and Riedel-Kruse [109] used pDawn [83] to regulate expression of Ag43 and biofilm formation in *E. coli*. With the ability to spatially control and tune biofilm formation, the authors were able to quantitatively characterize their system and propose a biophysical model that helps to understand the role that Ag43 plays in biofilms and may be useful in designing strategies to further improve spatial resolution.

Bacterial lithography has also been taken one step further. Zhao et al. [110] used pDawn [83] to pattern naturally electroactive *Shewanella oneidensis* onto electrode surfaces and tune the conductance and electrochemical activity of the biofilm. They used this material to study the intrinsic conductivity of a living biofilm. Recently, a preprint from Walker et al. [107] describe the use of Opto-T7RNAP [85] to spatially control synthesis of eumelanin in cellulose-producing *Komagataeibacter rhaeticus* for the purpose of producing self-dyeing patterned textiles. The resulting product had high background accumulation of eumelanin, which made it difficult to see the desired patterns. Even so, this work shows potential for growing patterned textiles.

A very impressive demonstration of the spatial properties of light is shown in Rivera-Tarazona et al. [66]. The authors embedded yeast in a hydrogel and controlled proliferation signals with optogenetics resulting in the controllable increase of the material of 400%. Especially worth mentioning is the supplementary movie S2 showing the controlled UV-light patterning and subsequent folding of aforementioned material into a helix (figure 1.2e). One can easily imagine the rise of optogenetic origami based off this work. Collaborations between topologists and bio-engineers could enable the creation of more complex structures.

1.4.6 Optogenetic Therapeutic applications

Live therapeutics comprise microorganisms engineered to have a therapeutic effect on the human body. Mechanisms of action can range from interacting with the host's immune system or metabolic processes to direct targeting and killing of tumor cells, thereby

promoting health and preventing or treating diseases [111]. Ideally, the therapeutic cells should specifically target the affected area, apply the therapeutic effector, and then clear from the body after treatment. Often, the therapeutic strain is engineered to constitutively produce the effector of interest. However, prolonged and uncontrolled delivery of some therapeutics can have adverse effects [112, 113], highlighting a need for regulation. In response, researchers are starting to design strains that produce the therapeutic only where and when needed but this requires the availability of appropriate biosensors [113, 114]. The use of optogenetics has the potential to confer spatiotemporal control capabilities to cells without the need for advanced biosensors.

Efforts are being made to build a variety of bacterial strains engineered to produce therapeutics upon light stimulation. For example, Sankaran et al. [115], showed that engineered *E. coli* embedded in a hydrogel matrix could produce and secrete the drug deoxyviolacein when stimulated by light. More recently, the efficacy of optogenetic live therapeutics has been demonstrated in mouse disease models. Light-regulated effectors range from IL-10 [67] and TGF- β 1 [67, 116] for the treatment of ulcerative colitis to TNF- α [68] and IFN- γ [116] for the treatment of cancer in mice. In addition, Pan et al. [117] also engineered *L. lactis* to produce gamma-aminobutyric acid, granulocyte-colony stimulating factor, and glucagon-like peptide-1 in mice to treat anxiety and Parkinson's disease, and to influence the central nervous system by stimulating the vagus nerve, respectively. Delivering light to optogenetic cells in animals is a challenge, as implanting a light source is invasive. While tissue-penetrating near-infrared light (NIR) can be externally applied, optogenetic systems that respond to NIR light are currently limited. One solution to this problem is the use of upconversion materials, which can convert NIR light to blue light [118]. This approach enables the use of more commonly used blue light optogenetic systems. In these examples, the optogenetic systems were constructed in either *E. coli* or *L. lactis* and introduced into mice with upconversion material either orally [67, 67, 116] or via tail vein injection [68] (figure 1.2f). Post-inoculation, NIR light was transiently and externally applied. For the oral inoculations, light was applied to the abdomens of the mice. For the tail vein injections, light was applied directly at the tumor site. In this way, it was possible to spatiotemporally control the application of therapy without using specific biosensors.

Many of these live therapeutic applications rely on efficient delivery of the therapeutic strain to the host gut. Diseases such as ulcerative colitis only affect the colon and therefore strains that can specifically colonize the colon for longer periods of time are preferable to therapies that more generally affect the entire gastrointestinal tract. Cui et al. [67] address this issue by engineering *E. coli* Nissle 1917 to produce the adhesin Ag43 and immunosuppressive cytokine TGF- β 1 upon blue light stimulation using the

1.4 Introduction to optogenetics and its application in bioproduction

pDawn [83] system. Bacteria and an NIR to blue light upconversion microgel, were orally-delivered to a dextran sulfate sodium-colitis mouse model. NIR light was externally and periodically applied to the abdomens of the mice to stimulate cell attachment to the colon and TGF- β 1 secretion. With this system, the authors were able to both specifically increase colon colonization as well as reduce gut inflammation.

1.4.7 Discussion and Conclusion about Optogenetic Applications

Microbial optogenetics holds immense potential for a wide range of applications, including model identification, co-cultures, protein production, biomaterials, and therapeutics by imparting the ability to control gene expression, metabolic and signaling pathways, with high precision, speed, and spatiotemporal control. Despite the progress made in optogenetic tool development, these tools are still not often used to their full potential. In this review, we aimed to present a variety of example applications, in which the unique properties of optogenetics are used to realize goals not previously possible via other methods.

In its most basic and common use, optogenetics is directly used as a replacement for chemical induction with added speed benefits, since light does not depend on small molecule diffusion for delivery. Light can be reversibly applied, allowing one to incorporate more complex dynamic induction schemes for better temporal control [63, 98]. Finally, light can be used to spatially control populations of cells all the way down to single cell and subcellular resolution [58, 66, 77, 117].

The studies presented here, are still in the early proof-of-concept stage. The next step is moving towards the ultimate goal of real-world use. If experiments are to be run in temporal and dynamic ways, new experimental setups will be required for higher throughput inputs and measurements [56]. For more industrial applications such as biomaterial synthesis and bioproduction, this will require scale up and likely the need for new light-delivery hardware and technologies. In terms of biomedical applications, challenges that will need to be addressed include testing of the optogenetic systems for cytotoxicity in humans and the development of NIR tools for better application of light, preferably without reliance on upconverting nanoparticles.

Overall, the future of microbial optogenetics looks bright, and we can expect to see many exciting developments and applications in the years to come. With further research and development, microbial optogenetics has the potential to revolutionize the way we study and manipulate microbial systems, and to open up new avenues for tackling a wide range of challenges.

1.5 Optogenetic Closed-loop Implementations

Bulk feedback of yeast liquid cultures was first demonstrated by Milias-Argeitis et al. [119]. The authors used the light-responsive Phy/PIF red/far-red optogenetic system to achieve robust regulation of gene expression fold change. The feedback was implemented by manual sampling of the cultures. Melendez et al. [120] improved this setup by removing the need for manual sampling, with automatic fluorescence imaging attached to a microfluidic device. With this setup, tight control over gene expression was achieved for 2 days. Milias-Argeitis et al. [121] used the green/red CcaS/CcaR system in *E. coli* to track the GFP level at different set points leveraging model predictive control to achieve the target levels. They additionally, controlled *MetE* a key metabolic enzyme for growth in *E. coli* and were able to control the growth rate with this setup.

Steel et al. [122] more recently developed a cheap custom-built 25mL optogenetic device and demonstrated closed-loop feedback control of GFP in *E. coli*. These papers have in common that they use reversible light inputs, and are focused on development of optogenetic regulators and regulation strategies in general. However, these proof of principles are operated at small volumes (<20mL) and neglect the challenges related to bioprocessing such as aeration, high cell densities and upscaling. Lastly, none of the above described papers focus on the optimization of the production of a protein or metabolite of interest. The idea of using closed-loop optogenetic feedback control for optimal bioproduction has been extensively discussed in the community [95, 96, 123, 124]. In this thesis, we close this gap.

1.6 Objective of this work

The goal of this thesis is to demonstrate the applicability and advantage of closed-loop, *in silico*, optogenetic feedback control for the production of proteins. Specifically, we are targeting the production of the hard-to-fold protein α -amylase by optogenetic control through the EL222 system [125].

The thesis is structured in four main results chapters, all describing a part of this interdisciplinary work. Each chapter includes a short introduction of the relevant literature, necessary to understand the field in question. Chapter 2 describes the development of a light illumination device centered around a commercial bioreactor that allows us to perform optogenetic experiments. We additionally built an automatic sampling and dilution system, in order to get real-time insights of the bioprocess and allow for automation. Chapter 3 describes the development of a *S. cerevisiae* strain that responds

to blue light and produces amylase. We genetically encode two biomolecular sensors, one for transcriptional output, one to monitor the unfolded protein response. We go on to characterize this strain and improve the growth media, to maintain unstressed cells for a longer duration. We develop a mathematical model in chapter 4 that describes the progression of the two fluorescent reporters and the cell density with a set of ordinary differential equations. The final chapter 5 shows closed-loop control of the unfolded protein response in *S. cerevisiae* and how intermediate closed-loop setpoints of the UPR maximize production of amylase.

Bibliography

- [1] Saurabh Aggarwal. What's fueling the biotech engine-2010 to 2011. *Nature Biotechnology*, 29(12):1083–1089, 2011. ISSN 10870156. doi: 10.1038/nbt.2060.
- [2] Sullivan and Frost. Size of the global chemical drugs and biologics pharmaceutical market from 2014 to 2023, 2023.
- [3] Nabih a Baeshen, Mohammed Baeshen, Abdullah Sheikh, Roop S Bora, Mohamed Morsi M Ahmed, Hassan A I Ramadan, Kulvinder Singh Saini, and Elrashdy M Redwan. Cell factories for insulin production. *Microbial cell factories*, 13(4):141, 2014. ISSN 1475-2859. doi: 10.1186/s12934-014-0141-0. URL <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4203937&tool=pmcentrez&rendertype=abstract>.
- [4] Mingtao Huang, Jichen Bao, and Jens Nielsen. Biopharmaceutical protein production by *Saccharomyces cerevisiae* : current state and future prospects. *Pharmaceutical Bioprocessing*, 2(2):167–182, 2014. ISSN 2048-9145. doi: 10.4155/pbp.14.8.
- [5] Dongsoo Yang, Seon Young Park, Yae Seul Park, Hyunmin Eun, and Sang Yup Lee. Metabolic Engineering of *Escherichia coli* for Natural Product Biosynthesis. *Trends in Biotechnology*, 38(7):745–765, 2020. ISSN 18793096. doi: 10.1016/j.tibtech.2019.11.007. URL <https://doi.org/10.1016/j.tibtech.2019.11.007>.
- [6] Mohsen Karbalaeei, Seyed A. Rezaee, and Hadi Farsiani. *Pichia pastoris*: A highly successful expression system for optimal synthesis of heterologous proteins. *Journal of Cellular Physiology*, 235(9):5867–5881, 2020. ISSN 10974652. doi: 10.1002/jcp.29583.
- [7] Rudi Emerson de Lima Procópio, Ingrid Reis da Silva, Mayra Kassawara Martins, João Lúcio de Azevedo, and Janete Magali de Araújo. Antibiotics produced by *Streptomyces*. *Brazilian Journal of Infectious Diseases*, 16(5):466–471, 2012. ISSN 14138670. doi: 10.1016/j.bjid.2012.08.014. URL <http://dx.doi.org/10.1016/j.bjid.2012.08.014>.
- [8] Wolfgang Schumann. Production of Recombinant Proteins in *Bacillus subtilis*. *Advances in Applied Microbiology*, 62(07):137–189, 2007. ISSN 00652164. doi: 10.1016/S0065-2164(07)62006-1.
- [9] Cen Li, Jingwen Zhou, Guocheng Du, Jian Chen, Shunji Takahashi, and Song Liu. Developing *Aspergillus niger* as a cell factory for food enzyme production. *Biotechnology Advances*, 44(March):107630, 2020. ISSN 07349750. doi: 10.1016/j.biotechadv.2020.107630. URL <https://doi.org/10.1016/j.biotechadv.2020.107630>.

-
- [10] Mingtao Huang, Yunpeng Bai, Staffan L. Sjoström, Björn M. Hallström, Zihe Liu, Dina Petranovic, Mathias Uhlén, Haakan N. Joensson, Helene Andersson-Svahn, and Jens Nielsen. Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 112(34):E4689–E4696, 2015. ISSN 10916490. doi: 10.1073/pnas.1506460112.
- [11] Mingtao Huang, Guokun Wang, Jiufu Qin, Dina Petranovic, and Jens Nielsen. Engineering the protein secretory pathway of *Saccharomyces cerevisiae* enables improved protein production. *PNAS*, 115(47):E11025–E11032, 2018. doi: 10.1073/pnas.1809921115.
- [12] Tadas Jakočinas, Ida Bonde, Markus Herrgård, Scott J. Harrison, Mette Kristensen, Lasse E. Pedersen, Michael K. Jensen, and Jay D. Keasling. Multiplex metabolic pathway engineering using CRISPR/Cas9 in *Saccharomyces cerevisiae*. *Metabolic Engineering*, 28:213–222, 2015. ISSN 10967184. doi: 10.1016/j.ymben.2015.01.008.
- [13] Hongguang Fu, Yanbing Liang, Xiuqin Zhong, Zhi Ling Pan, Lei Huang, Hai Lin Zhang, Yang Xu, Wei Zhou, and Zhong Liu. Codon optimization with deep learning to enhance protein expression. *Scientific Reports*, 10(1):1–9, 2020. ISSN 20452322. doi: 10.1038/s41598-020-74091-z. URL <https://doi.org/10.1038/s41598-020-74091-z>.
- [14] Mert Karaođlan and Fidan Erden-Karaođlan. Effect of codon optimization and promoter choice on recombinant endo-polygalacturonase production in *Pichia pastoris*. *Enzyme and Microbial Technology*, 139(March), 2020. ISSN 18790909. doi: 10.1016/j.enzmictec.2020.109589.
- [15] Beate Wiedemann and Eckhard Boles. Codon-optimized bacterial genes improve L-arabinose fermentation in recombinant *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 74(7):2043–2050, 2008. ISSN 00992240. doi: 10.1128/AEM.02395-07.
- [16] Xiao Chen, Kristian F. Nielsen, Irina Borodina, Morten C. Kielland-Brandt, and Kaisa Karhumaa. Increased isobutanol production in *Saccharomyces cerevisiae* by overexpression of genes in valine metabolism. *Biotechnology for Biofuels*, 4:1–12, 2011. ISSN 17546834. doi: 10.1186/1754-6834-4-21.
- [17] Zi Wei Luo and Sang Yup Lee. Metabolic engineering of *Escherichia coli* for the production of benzoic acid from glucose. *Metabolic Engineering*, 62(October):

- 298–311, 2020. ISSN 10967184. doi: 10.1016/j.ymben.2020.10.002. URL <https://doi.org/10.1016/j.ymben.2020.10.002>.
- [18] Dongbang Yao, Lingqia Su, Na Li, and Jing Wu. Enhanced extracellular expression of *Bacillus stearothermophilus* α -amylase in *Bacillus subtilis* through signal peptide optimization, chaperone overexpression and α -amylase mutant selection. *Microbial Cell Factories*, 18(1):1–12, 2019. ISSN 14752859. doi: 10.1186/s12934-019-1119-8.
- [19] Yoichiro Ito, Goro Terai, Misa Ishigami, Noriko Hashiba, Yasuyuki Nakamura, Takahiro Bamba, Ryota Kumokita, Tomohisa Hasunuma, Kiyoshi Asai, Jun Ishii, and Akihiko Kondo. Exchange of endogenous and heterogeneous yeast terminators in *Pichia pastoris* to tune mRNA stability and gene expression. *Nucleic Acids Research*, 48(22):13000–13012, 2020. ISSN 13624962. doi: 10.1093/nar/gkaa1066.
- [20] Kathleen A. Curran, Ashty S. Karim, Akash Gupta, and Hal S. Alper. Use of expression-enhancing terminators in *Saccharomyces cerevisiae* to increase mRNA half-life and improve gene expression control for metabolic engineering applications. *Metabolic Engineering*, 19:88–97, 2013. ISSN 10967176. doi: 10.1016/j.ymben.2013.07.001. URL <http://dx.doi.org/10.1016/j.ymben.2013.07.001>.
- [21] Tania Michelle Roberts, Hans-Michael Kaltenbach, and Fabian Rudolf. Development and Optimisation of a Defined High Cell Density Yeast Medium. *Yeast*, 2019.
- [22] Yan Lin, Wei Zhang, Chunjie Li, Kei Sakakibara, Shuzo Tanaka, and Hainan Kong. Factors affecting ethanol fermentation using *Saccharomyces cerevisiae* BY4742. *Biomass and Bioenergy*, 47:395–401, 2012. ISSN 18732909. doi: 10.1016/j.biombioe.2012.09.019. URL <http://dx.doi.org/10.1016/j.biombioe.2012.09.019>.
- [23] Dennis P. Bayrock and W. M. Ingledew. Ethanol production in multistage continuous, single stage continuous, *Lactobacillus*-contaminated continuous, and batch fermentations. *World Journal of Microbiology and Biotechnology*, 21(1):83–88, 2005. ISSN 09593993. doi: 10.1007/s11274-004-2781-4.
- [24] A. Mellitzer, R. Weis, A. Glieder, and K. Flicker. Expression of lignocellulolytic enzymes in *Pichia pastoris*. *Microbial Cell Factories*, 11, 2012. ISSN 01681656. doi: 10.1186/1475-2859-11-61.
- [25] Elena Camara, Joan Albiol, and Pau Ferrer. Droplet Digital PCR-Aided Screening and Characterization of *Pichia pastoris* Multiple Gene Copy Strains. *Biotechnology and Bioengineering*, 113(7), 2016.

- [26] Carissa L. Young and Anne S. Robinson. Protein folding and secretion: Mechanistic insights advancing recombinant protein production in *S. cerevisiae*. *Current Opinion in Biotechnology*, 30:168–177, 2014. ISSN 18790429. doi: 10.1016/j.copbio.2014.06.018. URL <http://dx.doi.org/10.1016/j.copbio.2014.06.018>.
- [27] Claudia Rubio, David Pincus, Alexei Korennykh, Sebastian Schuck, Hana El-Samad, and Peter Walter. Homeostatic adaptation to endoplasmic reticulum stress depends on Ire1 kinase activity. *Journal of Cell Biology*, 193(1):171–184, 2011. ISSN 00219525. doi: 10.1083/jcb.201007077.
- [28] Sijie Yu, Liangtian Miao, He Huang, Yin Li, and Taicheng Zhu. High-level production of glucose oxidase in *Pichia pastoris*: Effects of Hac1p overexpression on cell physiology and enzyme expression. *Enzyme and Microbial Technology*, 141(1):109671, 2020. ISSN 18790909. doi: 10.1016/j.enzmictec.2020.109671. URL <https://doi.org/10.1016/j.enzmictec.2020.109671>.
- [29] Carine Beaupere and Vyacheslav M. Labunskyy. (Un)folding mechanisms of adaptation to ER stress: lessons from aneuploidy. *Current Genetics*, 65(2):467–471, 2019. ISSN 14320983. doi: 10.1007/s00294-018-0914-9. URL <http://dx.doi.org/10.1007/s00294-018-0914-9>.
- [30] Peng Xu, Qin Gu, Wenya Wang, Lynn Wong, Adam G.W. Bower, Cynthia H. Collins, and Mattheos A.G. Koffas. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nature Communications*, 4:1408–1409, 2013. ISSN 20411723. doi: 10.1038/ncomms2425. URL <http://dx.doi.org/10.1038/ncomms2425>.
- [31] José L. Martínez, Eugenio Meza, Dina Petranovic, and Jens Nielsen. The impact of respiration and oxidative stress response on recombinant α -amylase production by *Saccharomyces cerevisiae*. *Metabolic Engineering Communications*, 3:205–210, 2016. ISSN 22140301. doi: 10.1016/j.meteno.2016.06.003. URL <http://dx.doi.org/10.1016/j.meteno.2016.06.003>.
- [32] Francesca Ceroni, Alice Boo, Simone Furini, Thomas E. Goroehowski, Olivier Borkowski, Yaseen N. Ladak, Ali R. Awan, Charlie Gilbert, Guy Bart Stan, and Tom Ellis. Burden-driven feedback control of gene expression. *Nature Methods*, 15(5):387–393, 2018. ISSN 15487105. doi: 10.1038/nmeth.4635.
- [33] P. Xu, L. Li, F. Zhang, G. Stephanopoulos, and M. Koffas. Improving fatty acids production by engineering dynamic pathway regulation and metabolic control. *Proceedings of the National Academy of Sciences*, 111(31):11299–11304, 2014. ISSN

- 0027-8424. doi: 10.1073/pnas.1406401111. URL <http://www.pnas.org/cgi/doi/10.1073/pnas.1406401111>.
- [34] Robert H. Dahl, Fuzhong Zhang, Jorge Alonso-Gutierrez, Edward Baidoo, Tanveer S. Batth, Alyssa M. Redding-Johanson, Christopher J. Petzold, Aindrila Mukhopadhyay, Taek Soon Lee, Paul D. Adams, and Jay D. Keasling. Engineering dynamic pathway regulation using stress-response promoters. *Nature Biotechnology*, 31(11):1039–1046, 2013. ISSN 10870156. doi: 10.1038/nbt.2689.
- [35] John M. Leavitt, James M. Wagner, Cuong C. Tu, Alice Tong, Yanyi Liu, and Hal S. Alper. Biosensor-Enabled Directed Evolution to Improve Muconic Acid Production in *Saccharomyces cerevisiae*. *Biotechnology Journal*, 12, 2017.
- [36] Wei Suong Teo and Matthew Wook Chang. Bacterial XylRs and synthetic promoters function as genetically encoded xylose biosensors in *Saccharomyces cerevisiae*. *Biotechnology Journal*, 10(2):315–322, 2015. ISSN 18607314. doi: 10.1002/biot.201400159.
- [37] Yanfei Zhang, Jeremy D. Cortez, Sarah K. Hammer, César Carrasco-López, Sergio García Echaury, Jessica B. Wiggins, Wei Wang, and José L. Avalos. Biosensor for branched-chain amino acid metabolism in yeast and applications in isobutanol and isopentanol production. *Nature Communications*, 13(1):1–14, 2022. ISSN 20411723. doi: 10.1038/s41467-021-27852-x.
- [38] Zhong Guo, Wayne A. Johnston, Jason Whitfield, Patricia Walden, Zhenling Cui, Elvira Wijker, Selvakumar Edwardraja, Ignacio Retamal Lantadilla, Fernanda Ely, Claudia Vickers, Jacobus P.J. Ungerer, and Kirill Alexandrov. Generalizable Protein Biosensors Based on Synthetic Switch Modules. *Journal of the American Chemical Society*, 141(20):8128–8135, 2019. ISSN 15205126. doi: 10.1021/jacs.8b12298.
- [39] Saurabh R. Nirantar, Kun Song Yeo, Sharon Chee, David P. Lane, and Farid J. Ghadessy. A generic scaffold for conversion of peptide ligands into homogenous biosensors. *Biosensors and Bioelectronics*, 47:421–428, 2013. ISSN 09565663. doi: 10.1016/j.bios.2013.03.049. URL <http://dx.doi.org/10.1016/j.bios.2013.03.049>.
- [40] Alfredo Quijano-Rubio, Hsien Wei Yeh, Jooyoung Park, Hansol Lee, Robert A. Langan, Scott E. Boyken, Marc J. Lajoie, Longxing Cao, Cameron M. Chow, Marcos C. Miranda, Jimin Wi, Hyo Jeong Hong, Lance Stewart, Byung Ha Oh, and David Baker. De novo design of modular and tunable protein biosensors. *Nature*, 591(7850):482–487, 2021. ISSN 14764687. doi: 10.1038/s41586-021-03258-z. URL <http://dx.doi.org/10.1038/s41586-021-03258-z>.

- [41] Colin Jackson, Alisha Anderson, and Kirill Alexandrov. The present and the future of protein biosensor engineering. *Current Opinion in Structural Biology*, 75(July):102424, 2022. ISSN 1879033X. doi: 10.1016/j.sbi.2022.102424. URL <https://doi.org/10.1016/j.sbi.2022.102424>.
- [42] Sebastián Bernales, Feroz R. Papa, and Peter Walter. Intracellular signaling by the unfolded protein response. *Annual Review of Cell and Developmental Biology*, 22:487–508, 2006. ISSN 10810706. doi: 10.1146/annurev.cellbio.21.122303.120200.
- [43] Yvonne Nyathi, Barrie M. Wilkinson, and Martin R. Pool. Co-translational targeting and translocation of proteins to the endoplasmic reticulum. *Biochimica et Biophysica Acta - Molecular Cell Research*, 1833(11):2392–2402, 2013. ISSN 01674889. doi: 10.1016/j.bbamcr.2013.02.021. URL <http://dx.doi.org/10.1016/j.bbamcr.2013.02.021>.
- [44] Mari Valkonen, Merja Penttilä, and Markku Saloheimo. Effects of inactivation and constitutive expression of the unfolded-protein response pathway on protein production in the yeast *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 69(4):2065–2072, 2003. ISSN 00992240. doi: 10.1128/AEM.69.4.2065-2072.2003.
- [45] Mari Valkonen, Michael Ward, Huaming Wang, Merja Penttilä, and Markku Saloheimo. Improvement of Foreign-Protein Production in *Aspergillus niger* var. *awamori* by Constitutive Induction of the Unfolded-Protein Response. *Applied and Environmental Microbiology*, 69(12):6979–6986, 2003. ISSN 00992240. doi: 10.1128/AEM.69.12.6979-6986.2003.
- [46] Mengmeng Huang, Yanyun Gao, Xiangshan Zhou, Yuanxing Zhang, and Menghao Cai. Regulating unfolded protein response activator HAC1p for production of thermostable raw-starch hydrolyzing alpha-amylase in *Pichia pastoris*. *Bioprocess and Biosystems Engineering*, 40(3):341–350, 2017. ISSN 16157605. doi: 10.1007/s00449-016-1701-y.
- [47] Philip I. Merksamer, Ala Trusina, and Feroz R. Papa. Real-Time Redox Measurements during Endoplasmic Reticulum Stress Reveal Interlinked Protein Folding Functions. *Cell*, 135(5):933–947, 2008. ISSN 00928674. doi: 10.1016/j.cell.2008.10.011. URL <http://dx.doi.org/10.1016/j.cell.2008.10.011>.
- [48] Kazutoshi Mori, Naoki Ogawa, Tetsushi Kawahara, Hideki Yanagi, and Takashi Yura. Palindrome with spacer of one nucleotide is characteristic of the cis-acting unfolded protein response element in *Saccharomyces cerevisiae*. *Jour-*

Chapter 1. Introduction

- nal of Biological Chemistry*, 273(16):9912–9920, 1998. ISSN 00219258. doi: 10.1074/jbc.273.16.9912. URL <http://dx.doi.org/10.1074/jbc.273.16.9912>.
- [49] Barbara Di Ventura and Wilfried Weber. The rise of molecular optogenetics, 2021.
- [50] Robert Ohlendorf and Andreas Möglich. Light-regulated gene expression in Bacteria: Fundamentals, advances, and perspectives. *Frontiers in Bioengineering and Biotechnology*, 10(October):1–39, 2022. ISSN 22964185. doi: 10.3389/fbioe.2022.1029403.
- [51] Sae Shimizu-Sato, Enamul Huq, James M. Tepperman, and Peter H. Quail. A light-switchable gene promoter system. *Nature Biotechnology*, 20(10):1041–1044, 2002. ISSN 10870156. doi: 10.1038/nbt734.
- [52] Anselm Levskaya, Aaron A Chevalier, Jeffrey J. Tabor, Zachary Booth Simpson, Laura A Lavery, Matthew Levy, Eric A. Davidson, Alexander Scouras, Andrew D Ellington, Edward M Marcotte, and Christopher A. Voigt. Engineering *Escherichia coli* to see light. *Nature*, 2005.
- [53] Jingjing Wei and Fan Jin. Illuminating bacterial behaviors with optogenetics. *Current Opinion in Solid State and Materials Science*, 26(6):101023, 2022.
- [54] Edoardo Romano, Armin Baumschlager, Emir Bora Akmeriç, Navaneethan Palanisamy, Moustafa Houmani, Gregor Schmidt, Mehmet Ali Öztürk, Leonard Ernst, Mustafa Khammash, and Barbara Di Ventura. Engineering AraC to make it responsive to light instead of arabinose. *Nature Chemical Biology*, 17(7):817–827, 2021. ISSN 15524469. doi: 10.1038/s41589-021-00787-6. URL <http://dx.doi.org/10.1038/s41589-021-00787-6>.
- [55] Matthew R Bennett, Wyming Lee Pang, Natalie A Ostroff, L Bridget, Sujata Nayak, Lev S Tsimring, and Jeff Hasty. Metabolic gene regulation in a dynamically changing environment. *Nature*, 454(7208):1119–1122, 2008. doi: 10.1038/nature07211.Metabolic.
- [56] Sant Kumar and Mustafa Khammash. Platforms for Optogenetic Stimulation and Feedback Control. *Frontiers in Bioengineering and Biotechnology*, 10(June):1–15, 2022. doi: 10.3389/fbioe.2022.918917.
- [57] David Figueroa, Vicente Rojas, Andres Romero, Luis F. Larrondo, and Francisco Salinas. The rise and shine of yeast optogenetics. *Yeast*, 38(2):131–146, 2021. ISSN 10970061. doi: 10.1002/yea.3529.

-
- [58] Dan Bracha, Mackenzie T. Walls, Ming Tzo Wei, Lian Zhu, Martin Kurian, José L. Avalos, Jared E. Toettcher, and Clifford P. Brangwynne. Mapping Local and Global Liquid Phase Behavior in Living Cells Using Photo-Oligomerizable Seeds. *Cell*, 175(6):1467–1480.e13, 2018. ISSN 10974172. doi: 10.1016/j.cell.2018.10.048.
- [59] Patrick Harrigan, Hiten D. Madhani, and Hana El-Samad. Real-Time Genetic Compensation Defines the Dynamic Demands of Feedback Control. *Cell*, 175(3):877–886.e10, 2018. ISSN 10974172. doi: 10.1016/j.cell.2018.09.044. URL <https://doi.org/10.1016/j.cell.2018.09.044>.
- [60] Remy Chait, Jakob Ruess, Tobias Bergmiller, Gašper Tkačik, and Călin C. Guet. Shaping bacterial population behavior through computer-interfaced control of individual cells. *Nature Communications*, 8(1), 2017. ISSN 20411723. doi: 10.1038/s41467-017-01683-1.
- [61] Dirk Benzinger, Serguei Ovinnikov, and Mustafa Khammash. Synthetic gene networks recapitulate dynamic signal decoding and differential gene expression. *Cell Systems*, 13(5):353–364.e6, 2022. ISSN 24054720. doi: 10.1016/j.cels.2022.02.004. URL <https://doi.org/10.1016/j.cels.2022.02.004>.
- [62] Makoto A. Lalwani, Hinako Kawabe, Rebecca L. Mays, Shannon M. Hoffman, and José L. Avalos. Optogenetic Control of Microbial Consortia Populations for Chemical Production. *ACS Synthetic Biology*, 10(8):2015–2029, 2021. ISSN 21615063. doi: 10.1021/acssynbio.1c00182.
- [63] Joaquín Gutiérrez Mena, Sant Kumar, and Mustafa Khammash. Dynamic cybergenetic control of bacterial co-culture composition via optogenetic feedback. *Nature Communications*, 13(1):1–16, 2022. ISSN 20411723. doi: 10.1038/s41467-022-32392-z.
- [64] Evan M. Zhao, Makoto A. Lalwani, Jhong Min Chen, Paulina Orillac, Jared E. Toettcher, and José L. Avalos. Optogenetic Amplification Circuits for Light-Induced Metabolic Control. *ACS Synthetic Biology*, 10(5):1143–1154, 2021. ISSN 21615063. doi: 10.1021/acssynbio.0c00642.
- [65] Moritz Benisch, Dirk Benzinger, Sant Kumar, Hanrong Hu, and Mustafa Khammash. Optogenetic closed-loop feedback control of the unfolded protein response optimizes protein production. *Metabolic Engineering*, 77(March):32–40, 2023. ISSN 1096-7176. doi: 10.1016/j.ymben.2023.03.001. URL <https://doi.org/10.1016/j.ymben.2023.03.001>.

- [66] L. K. Rivera-Tarazona, V. D. Bhat, H. Kim, Z. T. Campbell, and T. H. Ware. Shape-morphing living composites. *Science Advances*, 6(3):1–10, 2020. ISSN 23752548. doi: 10.1126/sciadv.aax8582.
- [67] Meihui Cui, Gaoju Pang, Tao Zhang, Tao Sun, Lili Zhang, Ruru Kang, Xin Xue, Huizhuo Pan, Chun Yang, Xinyu Zhang, Jin Chang, Jing Liu, Shufang Zhang, and Hanjie Wang. Optotheranostic Nanosystem with Phone Visual Diagnosis and Optogenetic Microbial Therapy for Ulcerative Colitis At-Home Care. *ACS Nano*, 15(4):7040–7052, 2021. ISSN 1936086X. doi: 10.1021/acsnano.1c00135.
- [68] Huizhuo Pan, Lianyue Li, Gaoju Pang, Chunli Han, Baona Liu, Yingying Zhang, Yue Shen, Tao Sun, Jing Liu, Jin Chang, and Others. Engineered NIR light-responsive bacteria as anti-tumor agent for targeted and precise cancer therapy. *Chemical Engineering Journal*, 426:130842, 2021.
- [69] Michael Chas Sumner, Steven B. Torrisi, Donna G. Brickner, and Jason H. Brickner. Random sub-diffusion and capture of genes by the nuclear pore reduces dynamics and coordinates inter- chromosomal movement. *eLife*, 10:1–25, 2021. ISSN 2050084X. doi: 10.7554/eLife.66238.
- [70] Hashem A. Meriesh, Andrew M. Lerner, Mahesh B. Chandrasekharan, and Brian D. Strahl. The histone H4 basic patch regulates SAGA-mediated H2B deubiquitination and histone acetylation. *Journal of Biological Chemistry*, 295(19):6561–6569, 2020. ISSN 1083351X. doi: 10.1074/jbc.RA120.013196. URL <http://dx.doi.org/10.1074/jbc.RA120.013196>.
- [71] Andrew M. Lerner, Austin J. Hepperla, Gregory R. Keele, Hashem A. Meriesh, Hayretin Yumerefendi, David Restrepo, Seth Zimmerman, James E. Bear, Brian Kuhlman, Ian J. Davis, and Brian D. Strahl. An optogenetic switch for the Set2 methyltransferase provides evidence for transcription-dependent and -independent dynamics of H3K36 methylation. *Genome Research*, 30(11):1605–1617, 2020. ISSN 15495469. doi: 10.1101/gr.264283.120.
- [72] Ellen H. Reed, Benjamin S. Schuster, Matthew C. Good, and Daniel A. Hammer. SPLIT: Stable Protein Coacervation Using a Light Induced Transition. *ACS Synthetic Biology*, 9(3):500–507, 2020. ISSN 21615063. doi: 10.1021/acssynbio.9b00503.
- [73] Mikael V. Garabedian, Wentao Wang, Jorge B. Dabdoub, Michelle Tong, Reese M. Caldwell, William Benman, Benjamin S. Schuster, Alexander Deiters, and Matthew C. Good. Designer membraneless organelles sequester native factors for control of cell behavior. *Nature Chemical Biology*, 17(9):998–1007, 2021. ISSN

15524469. doi: 10.1038/s41589-021-00840-4. URL <http://dx.doi.org/10.1038/s41589-021-00840-4>.
- [74] Gurkan Guntas, Ryan A Hallett, Seth P Zimmerman, Tishan Williams, Hayretin Yumerefendi, James E Bear, and Brian Kuhlman. Engineering an improved light-induced dimer (iLID) for controlling the localization and activity of signaling proteins. *Proceedings of the National Academy of Sciences*, 112(1):112–117, 2015.
- [75] Marc Rullan, Dirk Benzinger, Gregor W. Schmidt, Andreas Miliias-Argeitis, and Mustafa Khammash. An Optogenetic Platform for Real-Time, Single-Cell Interrogation of Stochastic Transcriptional Regulation. *Molecular Cell*, 70(4):745–756.e6, 2018. ISSN 10974164. doi: 10.1016/j.molcel.2018.04.012. URL <https://doi.org/10.1016/j.molcel.2018.04.012>.
- [76] Melinda Liu Perkins, Dirk Benzinger, Murat Arcak, and Mustafa Khammash. Cell-in-the-loop pattern formation with optogenetically emulated cell-to-cell signaling. *Nature Communications*, 11(1):1–10, 2020. ISSN 20411723. doi: 10.1038/s41467-020-15166-3. URL <http://dx.doi.org/10.1038/s41467-020-15166-3>.
- [77] Sant Kumar, Marc Rullan, and Mustafa Khammash. Rapid prototyping and design of cybergenetic single-cell controllers. *Nature Communications*, 12(1), 2021. ISSN 20411723. doi: 10.1038/s41467-021-25754-6. URL <http://dx.doi.org/10.1038/s41467-021-25754-6>.
- [78] Andela Davidovic, Remy Chait, Gregory Batt, and Jakob Ruess. Parameter inference for stochastic biochemical models from perturbation experiments parallelised at the single cell level. *PLoS Computational Biology*, 18(3):1–22, 2022. ISSN 15537358. doi: 10.1371/journal.pcbi.1009950.
- [79] Sebastian R Schmidl, Ravi U Sheth, Andrew Wu, and Jeffrey J Tabor. Refactoring and optimization of light-switchable Escherichia coli two-component systems. *ACS synthetic biology*, 3(11):820–831, 2014.
- [80] Paolo Guerra, Luc Alban Vuilleminot, Brady Rae, Valeriia Ladyhina, and Andreas Miliias-Argeitis. Systematic in Vivo Characterization of Fluorescent Protein Maturation in Budding Yeast. *ACS Synthetic Biology*, 11(3):1129–1141, 2022. ISSN 21615063. doi: 10.1021/acssynbio.1c00387.
- [81] Jessica B. Lee, Leandra M. Caywood, Jennifer Y. Lo, Nicholas Levering, and Albert J. Keung. Mapping the dynamic transfer functions of eukaryotic gene regulation. *Cell Systems*, 12(11):1079–1093.e6, 2021. ISSN 24054720. doi: 10.1016/j.cels.2021.08.003. URL <https://doi.org/10.1016/j.cels.2021.08.003>.

- [82] Brian D Zoltowski, Laura B Motta-Mena, and Kevin H Gardner. Blue light-induced dimerization of a bacterial LOV–HTH DNA-binding protein. *Biochemistry*, 52(38):6653–6661, 2013.
- [83] Robert Ohlendorf, Roe R Vidavski, Avigdor Eldar, Keith Moffat, and Andreas Möglich. From dusk till dawn: one-plasmid systems for light-regulated gene expression. *Journal of molecular biology*, 416(4):534–542, 2012.
- [84] Chetan Aditya, François Bertaux, Gregory Batt, and Jakob Ruess. A light tunable differentiation system for the creation and control of consortia in yeast. *Nature Communications*, 12(1), 2021. ISSN 20411723. doi: 10.1038/s41467-021-26129-7. URL <http://dx.doi.org/10.1038/s41467-021-26129-7>.
- [85] Armin Baumschlager, Stephanie K. Aoki, and Mustafa Khammash. Dynamic Blue Light-Inducible T7 RNA Polymerases (Opto-T7RNAPs) for Precise Spatiotemporal Gene Expression Control. *ACS Synthetic Biology*, 6(11):2157–2167, 2017. ISSN 21615063. doi: 10.1021/acssynbio.7b00169.
- [86] Michael B. Sheets, Nathan Tague, and Mary J. Dunlop. An optogenetic toolkit for light-inducible antibiotic resistance. *Nature communications*, 14(1):1034, 2023. ISSN 20411723. doi: 10.1038/s41467-023-36670-2.
- [87] Evan M. Zhao, Yanfei Zhang, Justin Mehl, Helen Park, Makoto A. Lalwani, Jared E. Toettcher, and José L. Avalos. Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature*, 555(7698):683–687, 2018. ISSN 14764687. doi: 10.1038/nature26141.
- [88] Peiling Wu, Yufen Chen, Mingyu Liu, Gezhi Xiao, and Jifeng Yuan. Engineering an Optogenetic CRISPRi Platform for Improved Chemical Production. *ACS Synthetic Biology*, 10(1):125–131, 2021. ISSN 21615063. doi: 10.1021/acssynbio.0c00488.
- [89] Sumeng Wang, Yue Luo, Wei Jiang, Xiaomeng Li, Qingsheng Qi, and Quanfeng Liang. Development of optogenetic dual-switch system for rewiring metabolic flux for polyhydroxybutyrate production. *Molecules*, 27(3):617, 2022.
- [90] Simona Gramazio, Jonathan Trauth, Filipp Bezold, Lars Oliver Essen, Christof Taxis, and Roberta Spadaccini. Light-induced fermenter production of derivatives of the sweet protein monellin is maximized in prestationary *Saccharomyces cerevisiae* cultures, 2022. ISSN 18607314.
- [91] Sebastian Tommi Tandar, Sachie Senoo, Yoshihiro Toya, and Hiroshi Shimizu. Optogenetic switch for controlling the central metabolic flux of *Escherichia coli*.

- Metabolic Engineering*, 55(June):68–75, 2019. ISSN 10967184. doi: 10.1016/j.ymben.2019.06.002. URL <https://doi.org/10.1016/j.ymben.2019.06.002>.
- [92] H el ene Duplus-Bottin, Martin Spichty, G erard Triqueneaux, Christophe Place, Philippe Emmanuel Mangeot, Th eophile Ohlmann, Franck Vittoz, and Ga el Yvert. A single-chain and fast-responding light-inducible cre recombinase as a novel optogenetic switch. *eLife*, 10:1–52, 2021. ISSN 2050084X. doi: 10.7554/eLife.61268.
- [93] Feng Cheng, Dong-Yang Wu, Xi-Hang Liang, Cheng-Jiao Wang, Jia-Qi Weng, Shu-Ping Zou, Jian-Miao Xu, Ya-Ping Xue, and Yu-Guo Zheng. A light-controlled biocatalytic system for precise regulation of enzymatic decarboxylation. *Catalysis Science & Technology*, 12(11):3421–3425, 2022.
- [94] Robert J. Lovelett, Jos e L. Avalos, and Ioannis G. Kevrekidis. Partial Observations and Conservation Laws: Gray-Box Modeling in Biotechnology and Optogenetics. *Industrial and Engineering Chemistry Research*, 59(6):2611–2620, 2020. ISSN 15205045. doi: 10.1021/acs.iecr.9b04507.
- [95] Makoto A. Lalwani, Evan M. Zhao, and Jos e L. Avalos. Current and future modalities of dynamic control in metabolic engineering. *Current Opinion in Biotechnology*, 52:56–65, 2018. ISSN 18790429. doi: 10.1016/j.copbio.2018.02.007.
- [96] Sylvain Pouzet, Alvaro Banderas, Matthias Le Bec, Thomas Lautier, Gilles Truan, and Pascal Hersen. The Promise of Optogenetics for Bioproduction : Dynamic Control Strategies and Scale-Up Instruments. *Bioengineering*, 2020.
- [97] Sebasti an Sosa Carrillo, Henri Galez, Sara Napolitano, Fran ois Bertaux, and Gregory Batt. Maximizing protein production by keeping cells at optimal secretory stress levels using real - time control approaches. *bioRxiv*, 2022.
- [98] Mateo I. Sanchez and Alice Y. Ting. Directed evolution improves the catalytic efficiency of TEV protease. *Nature Methods*, 17(2):167–174, 2020. ISSN 15487105. doi: 10.1038/s41592-019-0665-7. URL <http://dx.doi.org/10.1038/s41592-019-0665-7>.
- [99] Matthew J. Kennedy, Robert M. Hughes, Leslie A. Peteya, Joel W. Schwartz, Michael D. Ehlers, and Chandra L. Tucker. Rapid blue-light-mediated induction of protein interactions in living cells. *Nature Methods*, 7(12):973–975, 2010. ISSN 15487091. doi: 10.1038/nmeth.1524.
- [100] Wenjing Wang, Craig P Wildes, Tanyaporn Pattarabanjird, Mateo I Sanchez, Gordon F Guber, Gillian A Matthews, Kay M Tye, and Alice Y Ting. A light- and calcium-gated transcription factor for imaging and manipulating activated neurons. *Nature biotechnology*, 35(9):864–871, 2017.

- [101] Min Woo Kim, Wenjing Wang, Mateo I Sanchez, Robert Coukos, Mark Von Zastrow, and Alice Y Ting. Time-gated detection of protein-protein interactions with transcriptional readout. *Elife*, 6:e30233, 2017.
- [102] O. Burgos-Morales, M. Gueye, L. Lacombe, C. Nowak, R. Schmachtenberg, M. Hörner, C. Jerez-Longres, H. Mohsenin, H. J. Wagner, and W. Weber. Synthetic biology as driver for the biologization of materials sciences. *Materials Today Bio*, 11(May), 2021. ISSN 25900064. doi: 10.1016/j.mtbio.2021.100115.
- [103] Charlie Gilbert, Tzu Chieh Tang, Wolfgang Ott, Brandon A. Dorr, William M. Shaw, George L. Sun, Timothy K. Lu, and Tom Ellis. Living materials with programmable functionalities grown from engineered microbial co-cultures. *Nature Materials*, 20(5):691–700, 2021. ISSN 14764660. doi: 10.1038/s41563-020-00857-5. URL <http://dx.doi.org/10.1038/s41563-020-00857-5>.
- [104] Zhiqian Wang, Yunjun Yan, and Houjin Zhang. Design and Characterization of an Optogenetic System in *Pichia pastoris*. *ACS Synthetic Biology*, 11(1):297–307, 2022. ISSN 21615063. doi: 10.1021/acssynbio.1c00422.
- [105] Giacomo Frangipane, Dario Dell’Arciprete, Serena Petracchini, Claudio Maggi, Filippo Saglimbeni, Silvio Bianchi, Gaszton Vizsnyiczai, Maria Lina Bernardini, and Roberto Di Leonardo. Dynamic density shaping of photokinetic *E. coli*. *Elife*, 7:e36608, 2018.
- [106] Jessica M Walter, Derek Greenfield, Carlos Bustamante, and Jan Liphardt. Light-powering *Escherichia coli* with proteorhodopsin. *Proceedings of the National Academy of Sciences*, 104(7):2408–2412, 2007.
- [107] Kenneth T Walker, Jennifer Keane, Vivianne J Goosens, Wenzhe Song, Koon-Yang Lee, and Tom Ellis. Self-dyeing bacterial cellulose biomaterials 1 Self-dyeing textiles grown from cellulose-producing bacteria with engineered tyrosinase expression. *bioRxiv*, 2023. URL <https://doi.org/10.1101/2023.02.28.530172>.
- [108] Felix Moser, Eléonore Tham, Lina M González, Timothy K. Lu, and Christopher A. Voigt. Light-Controlled, High-Resolution Patterning of Living Engineered Bacteria Onto Textiles, Ceramics, and Plastic. *Advanced Functional Materials*, 2019.
- [109] Xiaofan Jin and Ingmar H Riedel-Kruse. Biofilm Lithography enables high-resolution cell patterning via optogenetic adhesin expression. *Proceedings of the National Academy of Sciences*, 115(14):3698–3703, 2018.

- [110] Fengjie Zhao, Marko S Chavez, Kyle L Naughton, Christina M Niman, Joshua T Atkinson, Jeffrey A Gralnick, Mohamed Y El-Naggar, and James Q Boedicker. Light-induced patterning of electroactive bacterial biofilms. *ACS Synthetic Biology*, 11(7):2327–2338, 2022.
- [111] Junheng Ma, Yuhong Lyu, Xin Liu, Xu Jia, Fangyun Cui, Xiaoheng Wu, Shanshan Deng, and Changwu Yue. Engineered probiotics. *Microbial cell factories*, 21(1): 1–9, 2022.
- [112] Yan Huang, Xiaojun Lin, Siyang Yu, Ruiyue Chen, and Weizhao Chen. Intestinal Engineered Probiotics as Living Therapeutics: Chassis Selection, Colonization Enhancement, Gene Circuit Design, and Biocontainment. *ACS Synthetic Biology*, 11(10):3134–3153, 2022.
- [113] Benjamin M. Scott, Cristina Gutiérrez-Vázquez, Liliana M. Sanmarco, Jessica A. da Silva Pereira, Zhaorong Li, Agustín Plasencia, Patrick Hewson, Laura M. Cox, Madelynn O’Brien, Steven K. Chen, Pedro M. Moraes-Vieira, Belinda S.W. Chang, Sergio G. Peisajovich, and Francisco J. Quintana. Self-tunable engineered yeast probiotics for the treatment of inflammatory bowel disease. *Nature Medicine*, 27(7):1212–1222, 2021. ISSN 1546170X. doi: 10.1038/s41591-021-01390-x. URL <http://dx.doi.org/10.1038/s41591-021-01390-x>.
- [114] Kristin J. Adolfsen, Isolde Callihan, Catherine E. Monahan, Per Jr Greisen, James Spoonamore, Munira Momin, Lauren E. Fitch, Mary Joan Castillo, Lindong Weng, Lauren Renaud, Carl J. Weile, Jay H. Konieczka, Teodelinda Mirabella, Andres Abin-Fuentes, Adam G. Lawrence, and Vincent M. Isabella. Improvement of a synthetic live bacterial therapeutic for phenylketonuria with biosensor-enabled enzyme engineering. *Nature Communications*, 12(1):1–13, 2021. ISSN 20411723. doi: 10.1038/s41467-021-26524-0.
- [115] Shrikrishnan Sankaran, Judith Becker, Christoph Wittmann, and Aránzazu Del Campo. Optoregulated Drug Release from an Engineered Living Material: Self-Replenishing Drug Depots for Long-Term, Light-Regulated Delivery. *Small*, 2019.
- [116] Chun Yang, Meihui Cui, Yingying Zhang, Huizhuo Pan, Jing Liu, Shixing Wang, Ning Ma, Jin Chang, Tao Sun, and Hanjie Wang. Upconversion optogenetic micro-nanosystem optically controls the secretion of light-responsive bacteria for systemic immunity regulation. *Communications biology*, 3(1):561, 2020.
- [117] Huizhuo Pan, Tao Sun, Meihui Cui, Ning Ma, Chun Yang, Jing Liu, Gaoju Pang, Baona Liu, Lianyue Li, Xinyu Zhang, and Others. Light-sensitive Lactococcus

- lactis for microbe–gut–brain Axis regulating via upconversion optogenetic micro-nano system. *ACS nano*, 16(4):6049–6063, 2022.
- [118] Yinyan Lin, Yuanfa Yao, Wanmei Zhang, Qiuyu Fang, Luhao Zhang, Yan Zhang, and Yingke Xu. Applications of upconversion nanoparticles in cellular optogenetics. *Acta Biomaterialia*, 135:1–12, 2021. ISSN 18787568. doi: 10.1016/j.actbio.2021.08.035.
- [119] Andreas Miliadis-Argeitis, Sean Summers, Jacob Stewart-Ornstein, Ignacio Zuleta, David Pincus, Hana El-Samad, Mustafa Khammash, and John Lygeros. In silico feedback for in vivo regulation of a gene expression circuit. *Nature Biotechnology*, 29(12):1114–1116, 2011. ISSN 10870156. doi: 10.1038/nbt.2018.
- [120] Justin Melendez, Michael Patel, Benjamin L. Oakes, Ping Xu, Patrick Morton, and Megan N. McClean. Real-time optogenetic control of intracellular protein concentration in microbial cell cultures. *Integrative Biology (United Kingdom)*, 6(3):366–372, 2014. ISSN 17579708. doi: 10.1039/c3ib40102b.
- [121] Andreas Miliadis-Argeitis, Marc Rullan, Stephanie K. Aoki, Peter Buchmann, and Mustafa Khammash. Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth. *Nature Communications*, 7(May):1–11, 2016. ISSN 20411723. doi: 10.1038/ncomms12546. URL <http://dx.doi.org/10.1038/ncomms12546>.
- [122] Harrison Steel, Robert Habgood, Ciarán Kelly, and Antonis Papachristodoulou. In situ characterisation and manipulation of biological systems with Chi.Bio. *PLoS Biology*, 18(7):1–12, 2020. ISSN 15457885. doi: 10.1371/journal.pbio.3000794.
- [123] Evan M. Zhao, Makoto A. Lalwani, Robert J. Lovelett, Sergio A. García-Echauri, Shannon M. Hoffman, Christopher L. Gonzalez, Jared E. Toettcher, Ioannis G. Kevrekidis, and José L. Avalos. Design and Characterization of Rapid Optogenetic Circuits for Dynamic Control in Yeast Metabolic Engineering. *ACS Synthetic Biology*, 9(12):3254–3266, 2020. ISSN 21615063. doi: 10.1021/acssynbio.0c00305.
- [124] César Carrasco-López, Sergio A. García-Echauri, Therese Kichuk, and José L. Avalos. Optogenetics and biosensors set the stage for metabolic cybergenetics. *Current Opinion in Biotechnology*, 65:296–309, 2020. ISSN 18790429. doi: 10.1016/j.copbio.2020.07.012.
- [125] Dirk Benzinger and Mustafa Khammash. Pulsatile inputs achieve tunable attenuation of gene expression variability and graded multi-gene regulation. *Nature Communications*, 9(2018):1–38, 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05882-2. URL <http://dx.doi.org/10.1038/s41467-018-05882-2>.

2 Platform development

The production of recombinant proteins in yeast is performed in tightly controlled stirred tank reactors (STRs) at optimal oxygen, nutrient and pH conditions [1, 2]. Commercial bioreactors exist for a wide range of volumes and can be made of glass, steel or plastic (single-use).

The ultimate goal of this PhD thesis is to implement closed-loop optogenetic feedback control on a bioreactor for microbiology applications. We want to leverage optogenetics for actuation of the cells. In order to achieve this, we developed an illumination device that is able to supply light to the bioreactor contents. Furthermore, real-time data of the cell state is required to feed into the control algorithms. We thus extended the platform to additionally include automatic sampling and dilution.

Usage of optogenetics for protein production requires proper illumination devices. Most experiments in the microbiology setting are performed at smaller volume scales [3–7]. Steel et al. [8] developed a 20mL platform termed “Chi.Bio” which allows for the automatic measurement of a variety of fluorophores and optical density. Additionally, the authors implemented continuous culture capabilities in their platform and an LED for the usage as an optogenetic input. Tandar et al. [9] used the red-light CcaS/CcaR system in *E. coli* to regulate the metabolic allocation between the Embden-Meyerhof-Parnas (EMP) and the oxidative pentose phosphate (oxPP) pathways at the 20mL scale. Here, the authors placed single LEDs below shaking flasks or tubes. The test tube scale is good to quickly test the functionality of optogenetics, but lacks control over pH and aeration that bioreactors offer and which are required in full-scale bioprocesses.

The field of photobioreactors in the algae world is much more mature [10]. The goal here is to maximize the surface area so that liquid culture algae can transform CO₂ and ambient light into biomass for later energy usage. While these platforms are more

Chapter 2. Platform development

mature, the focus here is on delivering the maximum amount of light to the cells. Thus, control over the light wavelength or keeping the system dark from external influences is not important. Additionally, the operating volumes of algae reactors are much larger than necessary for laboratory experiments. For our applications however, it is crucial to apply a defined wavelength to the cells, as well as to be able to turn off all light activation.

Chen et al. [11] and Chang et al. [12] have demonstrated the usage of optogenetic tools in a bioreactor setting. In both cases however the setup can not efficiently be shielded from ambient light. A hallmark publication stems from Zhao et al. [13], where they demonstrated the usage of optogenetics at the 5L scale to produce isobutanol and 2-methyl-1-butanol optogenetically and achieve considerably higher titers. They achieved this by wrapping the glass-walled bioreactor with flexible LED panels. The same approach was used in more recent publications from the Avalos Research Group [14, 15]. What is missing for further development is a robust platform that allows the easy application of light for optogenetic induction at the bioreactor scale without the need of manual attachment of LEDs to the bioreactor after sterilization.

All taken together, there is no commercial standard for a photobioreactor platform for microbiology cultivations available giving rise to the need to develop a platform ourselves. We used a commercially available bioreactor and developed an light delivery system around it in order to run optogenetic experiments under industrial conditions. In the first part of this chapter (section 2.1), we discuss the development of the photobioreactor platform, which allows the illumination of a 1.2L bioreactor with red, green and blue LEDs.

In recombinant protein production, the organism of choice acts as a catalyst converting substrate to a product. Consequently, a prerequisite for high productivity is a sufficiently large cell density [16]. Conditions allowing for this are achieved in batch and fed-batch high cell density cultivations (HCDC). Most bioreactor platforms are equipped with sensors for temperature, dissolved oxygen and pH. Additionally, platforms can include conductivity, pressure and turbidity measurements, off-gas analysis systems and the *in situ* measurement of a variety of metabolites (e.g. glucose, acetate). For our specific purpose, we would like to measure the abundance of fluorescent proteins inside of our cells. Quantification of fluorescent proteins is usually performed by a fluorometer or flow cytometer. These devices have a narrow operating range, meaning that the cell densities encountered in HCDCs pose a significant challenge for measurements. We have thus implemented an automatic sampling and dilution system to meet the requirements posed by the measurement devices.

2.1 Developing a photobioreactor

Automatic sampling devices have been around for a long time [17, 18] and some systems are commercially available [19] or can be assembled for a relatively low price [8, 20]. Sampling systems have also been developed in our research group [5, 21] and provide some inspiration on how to develop automatic sampling. These devices were all used in experiments where the cell density was controlled to a low, constant optical density, allowing direct measurement of the sample without prior dilution. With the rise of pipetting robots, automatic dilution of samples is also possible (e.g. Opentrons). An affordable combination of monoseptic sampling and dilution is however not affordable and often not open for modifications.

Due to the lack of available and affordable combined monoseptic sampling and dilutions systems, we developed an automatic sampling and dilution setup (section 2.2), which allows for the robust sampling of bioreactor culture and dilutions of up to 400-fold. It gives us an easy way to modify the system, is inexpensive and integrateable with a flow cytometer, resulting in an automatic way to quantify the fluorescence of cells in exponentially growing cultures.

2.1 Developing a photobioreactor

The bioreactor we used as the basis for the photobioreactor platform is an Eppendorf BioFlo®120 with a total volume of 1.2L and an operating volume of 1L (figure 2.1). It is equipped with sensors for measuring pH, dissolved oxygen (DO) and temperature, and has a control unit for regulating these parameters. To provide heating and cooling, we used a water-jacketed version rather than a heat-blanketed bioreactor, so that the bioreactor liquid could be continuously illuminated from the side. The bioreactor is placed on a heat plate to warm the the water-jacket from below, and cold water (5°C) can be pumped into the water jacket for cooling.



Figure 2.1: Picture of the bioreactor BioFlo®120 [22]. Shown is the heat-blanketed 2L version.

For illuminating the bioreactor, we built a hexagonal prism with sides of 18.5cm length and a height of 28cm. Into the sides of the prism, we placed six Adafruit NeoPixel NeoMatrix LED light panels with 8x8 individual LEDs (see figure 2.2). To prevent overheating of the light panels, we placed custom-made 7.5x10cm heat sinks on the outside of each panel. The prism can be sealed at the top with covers to completely

Chapter 2. Platform development

block ambient light, resulting in a measured irradiance of $0.02 \frac{\mu W}{cm^2}$ when all light panels are turned off.

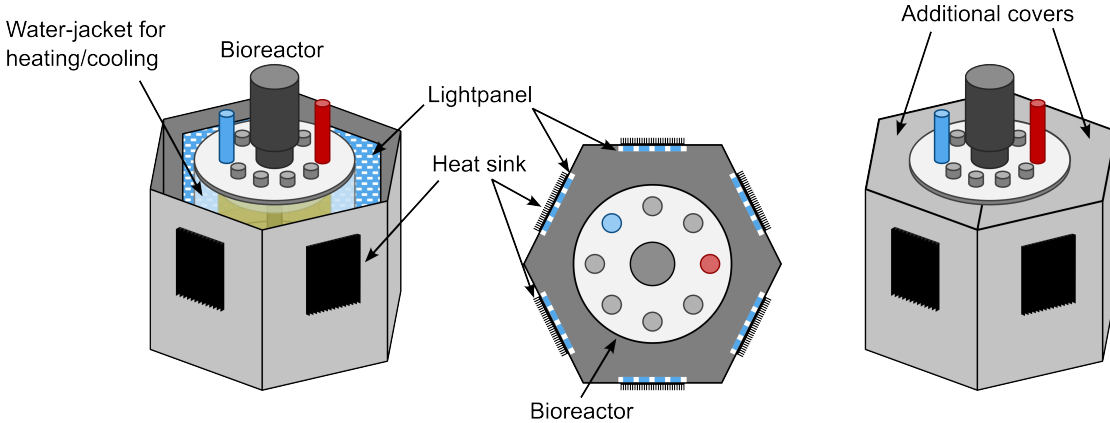


Figure 2.2: Schematic view of the photobioreactor from the side (**left**) and top (**center**). The bioreactor is placed inside a hexagon container, which shields the bioreactor from ambient light. Six light panels are placed on the inside of the hexagon with six heat sinks on the outside, to avoid overheating of the light panels. **right** Additional covers can be placed at the top of the hexagon to shield all sides of the bioreactor from light.

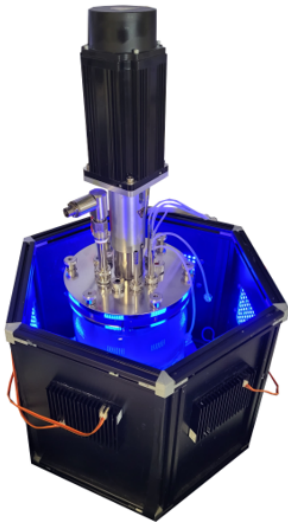


Figure 2.3: Picture of the photobioreactor platform inside the hexagonal prism without the additional covers.

All of the light panels are controlled by an Arduino Mega 2560, with two panels connected in series and each pair controlled by one digital pin of the Arduino. The electrical current for the light panels is supplied by a custom-built power driver (5V) with a maximum current of 2A per light panel (12A total). A picture of the bioreactor with all LEDs set to 10% brightness can be seen in figure 2.3.

To understand the range of brightness levels at which the photobioreactor can be operated, we measured the irradiance inside the bioreactor at various input brightness levels of the Arduino. The measurements were performed with the Thorlabs PM100USB Power and Energy Meter at the excitation wavelength of EL222 ($\lambda = 450nm$) [23] with averaging over 1'000 datapoints. The sensor was placed inside the bioreactor adjacent to the glass wall facing one individual light panel and only this light panel was turned on for measurements. Figure 2.4 shows the irradiance plotted against the digital brightness input level on a log-log plot with a clear linear correlation (Adj. R-squared = 0.99997). A threshold input level ≥ 0.008 needs to be supplied to be in the linear range of the light panels. The maximum

2.2 Automatic measurements

irradiance with one light panel is $267.5 \frac{\mu W}{cm^2}$. When all light panels are turned on, this number increases to $625 \frac{\mu W}{cm^2}$. The actual experienced irradiance of cells at the bioreactor wall is a function of the cell density, as light panels opposite of the measurement point will contribute less irradiance at higher cell densities than when the cell density is close to zero. The two maximum values can be used to assess whether cells could be exposed to phototoxic effects when illuminated with maximum brightness in the bioreactor.

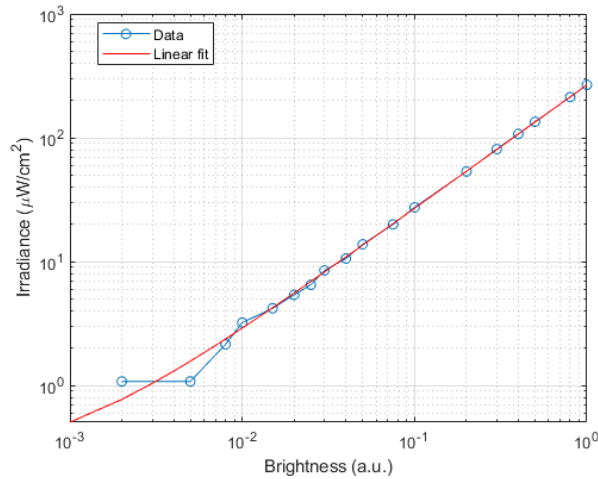


Figure 2.4: Irradiance given in $\frac{\mu W}{cm^2}$ plotted against the brightness in a.u. that is set with the Arduino Mega.

2.2 Automatic measurements

2.2.1 Development of an automatic sampling and dilution module

The bioreactor is operated in batch mode to closely mimic the operating conditions in industry. This is because higher cell densities result in more “biocatalyst” that is producing the product leading to higher productivities. We inoculate our bioreactor at $OD_{600} \approx 0.02$ and cell densities reach $OD_{600} \approx 20$ at the end resulting in a 1’000 fold range where reliable measurements need to be obtained. Typical lab measurement devices can produce reliable measurement output in a much smaller cell density range. For example, the NanoDropTM has a 20-fold measurement range, a Beckman Coulter CytFLEX S flow cytometer a 30-fold range. This results in the need for dilution of the sample before measurement, to bring the cell density into the measurement range.

We built an automatic sampling and dilution platform which is able to automatically sample from the bioreactor, dilute the sample and transfer it to a measurement device. Initially in the batch run, the sample can be fed into the sampling device undiluted. At

Chapter 2. Platform development

the end of the bioreactor run a dilution of 50-fold is required to be in the measurement regime of our flow cytometer. We split this dilution task into two dilution stages so that each dilution step doesn't exceed a 10-fold dilution (lab standard for dilution series). Both dilution stages are shown in figure 2.14.

The sampling relies on the usage of a set of pumps taking inspiration from Miliadis-Argentis et al. [5]. Four pumps are dedicated to each dilution stage, one for air, sample, diluent and waste, as well as a glass vial for sample storage, dilution and mixing. For practical implementation a pump board from Chi.Bio [8] containing four peristaltic pump heads is used for each dilution stage. Finally, one custom developed pump is used to transfer the sample from the second dilution vial to the measurement device.

Figure 2.5 depicts the steps performed at each dilution stage. Simply put, sample is drawn by the air pump (P1) into tubing connecting the bioreactor to the glass vial. A defined pulse of sample is drawn by the sample pump (P2) and pushed to the sample vial. The sample residing in the tubing is pushed back to the bioreactor and the sample is diluted with diluent (P3). The sample is mixed with the diluent and transferred to the next dilution stage and finally to the measurement device. Lastly, the vial is rinsed with diluent and all liquids are transferred to a waste container (P4).

The amount of dilution depends on the measurement device. For our application, we used the CytFLEX S. If the cell density in the sample is too large, then cells will enter the flow cell too close to each other and the instrument can not discern between the cells. In order to avoid wrong measurements, the instrument labels those events as aborted. The abortion rate is now all cells aborted divided by all events measured by the instrument. Operating at low cell densities also ensures that the flow cell doesn't clog. We aim for an abortion rate $< 2\%$ which for yeast cells corresponds to event counts of $5000 \frac{\text{events}}{\mu\text{L}}$ (see figure 2.6). In order to safely work in this regime we set a target cell density of $2500 \frac{\text{events}}{\mu\text{L}}$ to be fed into the flow cytometer. The dilution rate for each new measurement is calculated from the previously measured cell density and previous dilution rate. With a fixed sample pump rate and duration, the dilution is tuned by varying the time that the diluent pump is turned on.

A picture of the automatic sampling and dilution can be seen in figure 2.7. Additionally, figure 2.8 shows a sample 150-fold dilution achieved with the automatic sampler

2.2 Automatic measurements

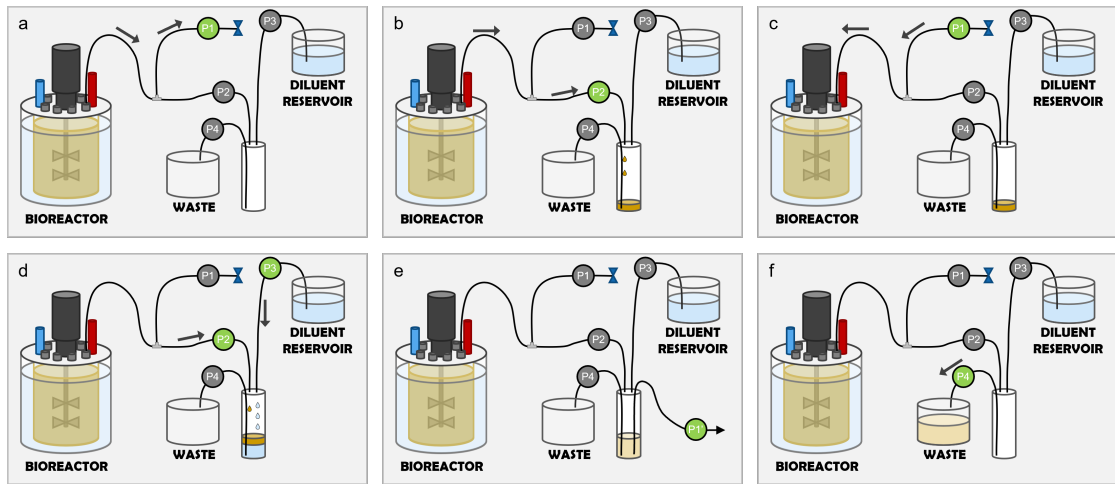


Figure 2.5: Automatic sampling setup in detail. (a) Initially, pump P1 is turned on and the pump direction of P1 is reversed three times to ensure proper mixing of sample in the sample tubing. Then, the sample is pumped from the bioreactor through a T-piece into tubing of pump P1. This allows a more precise amount of sample to be transferred through pump P2 towards the sample vial, as the sample is primed until the T-piece. (b) The sample is pumped by pump P2 into the sample vial. (c) The pump direction of P1 is reversed and the sample remaining in the tube is pushed back into the bioreactor. A sterile filter behind P1 ensures continuous sterility during sampling. (d) The remaining sample between T-piece and vial is pushed into the sample vial, while the sample is diluted with pump P3 which pumps diluent from the diluent reservoir into the vial. (e) The sample is mixed with pump P1' of the next dilution stage and finally transferred on to the next dilution step or measurement device, similar as in step a from bioreactor to sample vial 1. (f) After transfer, the sample is removed with pump P4 into a waste reservoir. The vial is rinsed with more diluent (P3) and the diluent removed again with P4.

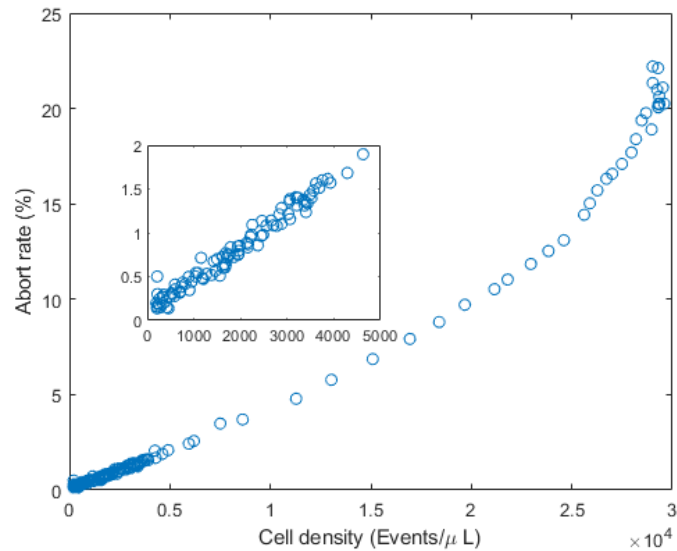


Figure 2.6: The abort rate as given by the flow cytometer as a function of the measured cell density. The inset shows the data for abort rates below 2%.

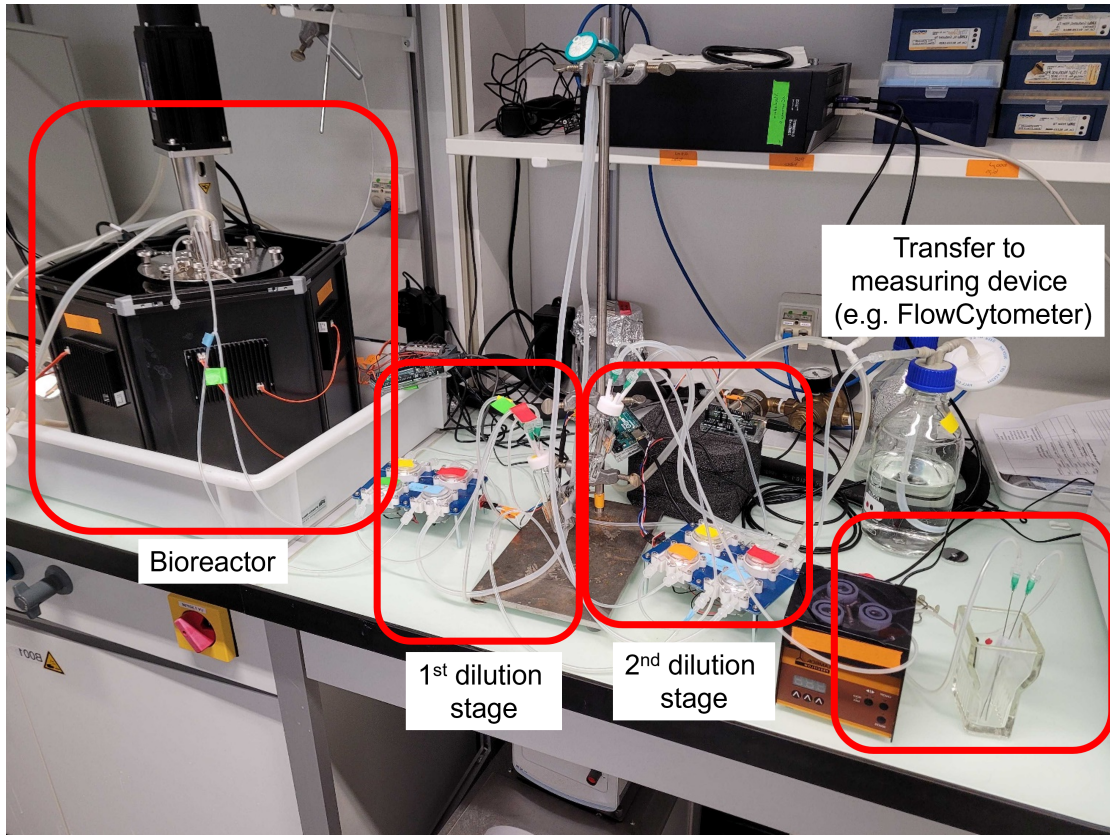


Figure 2.7: Picture of the photobioreactor platform with attached automatic sampling and dilution stages. Samples are transferred from the bioreactor through the first and second dilution stage and are finally pumped to the measuring device (in our case a flow cytometer, not shown).

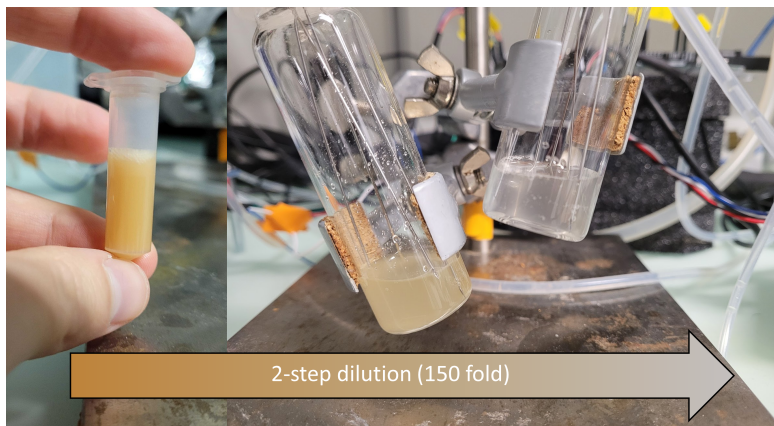


Figure 2.8: Example of dilutions achieved with the automatic dilution device. The sample shown on the left in the Eppendorf tube is sampled manually from the bioreactor. The vial in the center shows the sample after the first dilution (12-fold) and the vial on the right the dilution after the second dilution (150-fold total dilution). Samples are pumped from that second vial to the measuring device.

2.2.2 Beckman Coulter CytFLEX S automation workflow

In order to fully automate the sampling, dilution and measurement with a flow cytometer the procedure for measurement also needs automation.

The CytFLEX S flow cytometer has two different sample injection modes; tube and plate. For convenience we used the tube injection mode, as bigger volumes can be supplied in comparison to standard 96-well plates. Beckman Coulter does not provide a simple interface to communicate with the flow cytometer software CytExpert. For a typical measurement, a new and empty tube sample needs to be created, the machine needs to be initialized, the sample line backflushed and boosted, the sample measurement needs to be recorded for 30 seconds, followed by another cleaning step and finally export of the data to a csv file.

To automate these steps, we initially leveraged robotics process automation (RPA) with the RPA software UiPath. RPA tries to mimic the software workflow that is otherwise implemented by a human. One can record the software workflow, by clicking on required fields and play this workflow again upon request. Old software implementations of RPA clicked on defined pixels on a defined screen. In contrast, the newer software packages like UiPath record the field that is clicked on rather than the pixel position, making it more robust to changes in e.g. screen resolution. One such workflow can be seen in figure 2.9 which automates the clicking for a measurement. For maximal robustness, additional steps and cases need to be introduced, so that UiPath can always control CytExpert, even when it is e.g. running in the background.

Two such workflows were implemented, one for measurement automation and one for automation of the export of raw data into a csv file. The workflow could be run through the command line of Matlab and allowed the successful automation of measurements. These scripts were improved over time, incorporating more and more special cases and were used for experiments until August 2021.

For longterm experiments (> 24 hours), some limitations arise from the CytExpert software. The graphical user interface requires a manual system startup of the machine every 24 hours which requires manual intervention. In parallel to this PhD project a similar setup was developed in the Control Theory and Systems Biology Laboratory, which also automates the measurement with a CytFLEX S [21], though without the need of dilutions. This project leveraged an application programming interface (API) from CytExpert, which overruled the need for daily system startups and daily cleans. Thus, the RPA was replaced with the API module to automate measurements with the CytFLEX S. The code of the API module was developed by Sant Kumar.

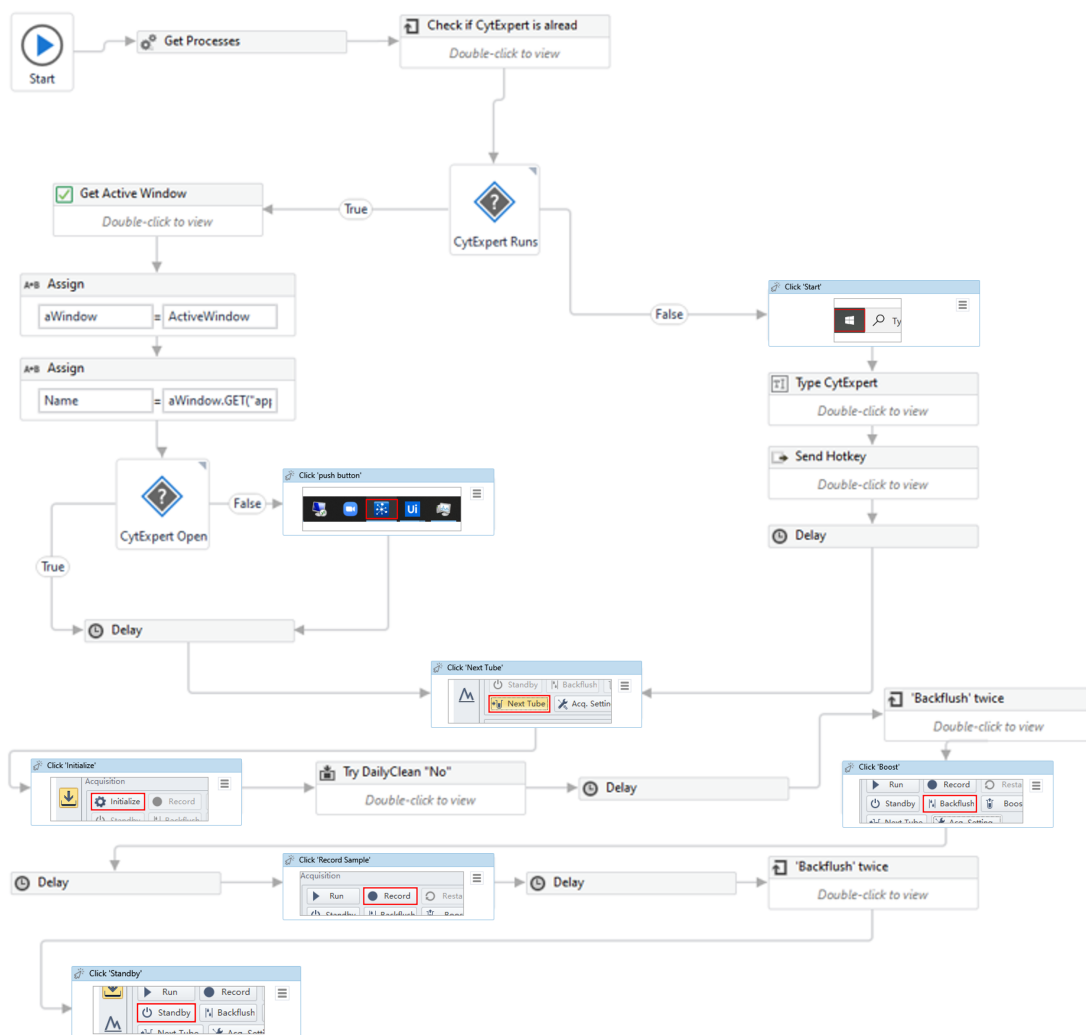


Figure 2.9: Robotic process automation (RPA) workflow for automation of the measurement procedure in CytExpert. Initially, the script checks whether CytExpert is running. If it is running, we need to identify from the active windows, whether CytExpert is selected and in the foreground or whether it needs to be selected first. If it is not running, CytExpert is started from the Windows Start menu. After this process, the actual measurement steps are implemented (lower half). Note that between some steps an additional delay is implemented, so that the machine can perform the previous step (e.g. between record sample, which takes 30 seconds and backflush).

2.3 Amylase sampling and automation

We decided to collect samples of cells with supernatant separately in order to measure the concentration of amylase offline.

2.3.1 Automation of amylase sampling

We custom-built an amylase sampling device that can collect up to 14 individual samples for later offline quantification. The sampler is based around a set of pneumatic manifolds with fitted pneumatic valves that allow the distribution of liquids through several channels. In our case, we used two manifolds with each one inlet and eight outlets. Figure 2.10 shows the sample origin (here the bioreactor), a pump P, two manifolds (grey box) and a total of 16 pneumatic valves (stacked triangles).

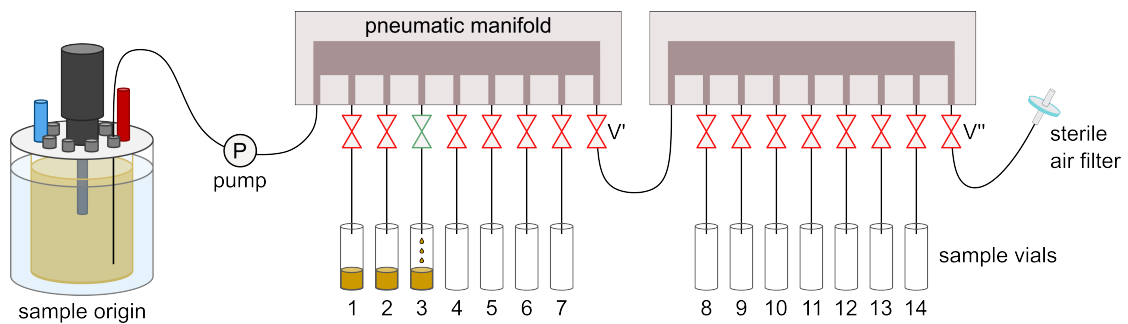


Figure 2.10: Graphic depiction of the amylase sampling station with two pneumatic manifolds, each with one inlet and 8 outlets and a pump P driving the liquid from sample origin to sample vials. Valves (stacked triangles) control the flow of liquid. Samples are filled in order from left (1) to right (14). In the depicted example, sample vials 1 and 2 have already been filled in previous steps. The valve for sample vial 3 is opened while all other valves are closed resulting in liquid flowing into sample vial 3.

For implementation we used the pneumatic manifold pneumadyne msv10-8 with pneumadyne S10MM-30 24-3 valves to control the flow. Each outlet can be fitted with a pneumatic valve stopping any liquid from passing through the output. Sample from the bioreactor is pumped with a pump P to the inlet of manifold one. For samples that should be stored in sample vials 1-7, only this specific valve is opened. For samples that should be stored in vials 8-14, the valve V' is additionally opened to allow the liquid to pass through manifold 1. A final outlet V'' is used to flush back any remaining liquid in the tube back to the sample origin, by switching the flow direction of the pump while keeping valves V' and V'' open. The outlet at V'' is equipped with a sterile filter, so that air is filtered before being pushed back to the bioreactor. The valves are controlled by an arduino and can be opened by sending a digital "HIGH" signal to the pin to which the valve is connected.

To stop microbial growth, 10 μL of 11mM sodium azide are placed in each sample vial, sufficient to stop microbial growth in a 1mL sample. The sample is kept at room temperature until measurement.

2.3.2 Automation of the amylase offline quantification

We used the Megazyme alpha-amylase assay kit to measure the concentration of alpha-amylase in samples. Commercial α -amylase (Sigma-Aldrich) from *Aspergillus oryzae* (Cat. Nr. 10065) was used as a standard. The sample was centrifuged at 4'000 rcf for 5 minutes to separate cells from the amylase containing supernatant. The Megazyme alpha-amylase assay reaction volume was scaled-down from 3.4 ml total volume to 250 μ l total volume to enable automation and quantification in 96-well plates. The assay was further performed at room temperature, to reduce the amount of evaporation in the 96-well plate and the incubation time with HR reagent was extended to 15 minutes.

The following steps for quantification were performed. 50 μ l of biologically inactivated supernatant was transferred into a 96-well plate and incubated at room temperature for 5 minutes. 50 μ l of HR solution was added to each sample to start the reaction. The plate was shaken at 1'000 rpm for 15 minutes before adding 150 μ l of 40 g/l Na_3PO_4 (pH=11). The stop buffer was prepared by dissolving 92.768 g of $\text{Na}_3\text{PO}_4 \cdot 12 \text{H}_2\text{O}$ in 1l water. The pipetting was performed on a Hamilton microlab star and measurement of the absorbance at 400nm with a Tecan Infinite M200Pro. The Hamilton automation protocol was developed together with Gregor Schmidt from the ETH, D-BSSE automation facility.

2.3.3 Maintaining a monoseptic fermentation

Initially, the platform with installed automatic sampling was running well with no observed contamination of the monoseptic culture. After the installment of the amylase sampling device significant and persistent contaminations of the bioreactor with *E. coli* and hyphae were observed (see figure 2.11). The pneumatic valves can not be autoclaved and cleaning of the setup with 70% ethanol did not remove the contaminations.

Four alterations were performed that in combination removed any further contamination. First, we identified that hyphae could survive in the O-rings of the bioreactor even after autoclaving and keeping the rings in 70% ethanol (see figure 2.11). We thus replaced all O-rings in the setup in an effort to remove any remaining spores.

Secondly, we observed that small amounts of liquid could travel back from the first dilution vial into the bioreactor potentially contaminating the fermentation. Thus, we replaced the needle used to transfer the liquid into the dilution vial with a shorter one. Thus, the liquid is now dropped into the mixture rather than directly added at the bottom of the vial. Previously, the sampling apparatus was sterilized *in situ* by rinsing the apparatus with 70% ethanol. This is improved by autoclaving the entire

2.3 Amylase sampling and automation

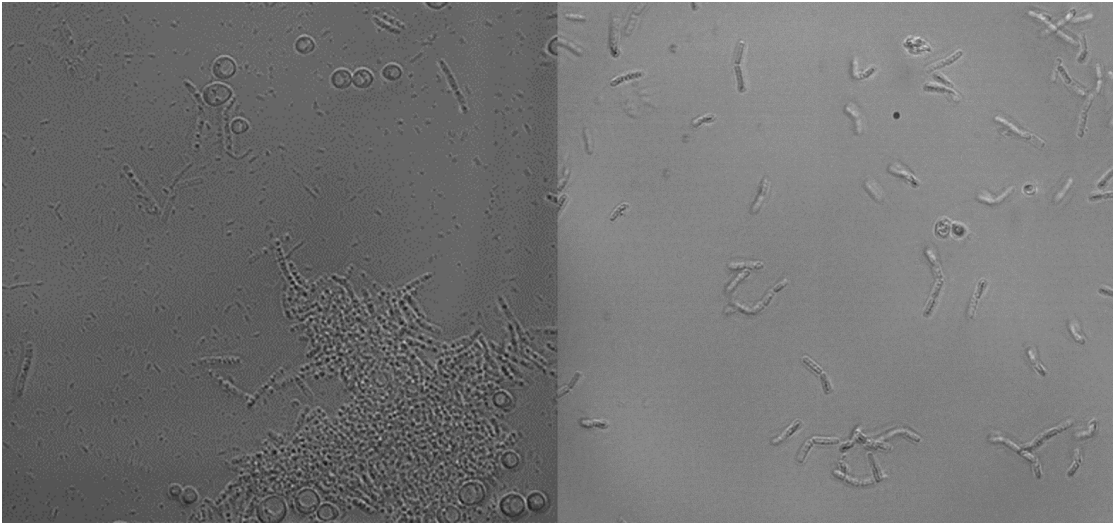


Figure 2.11: (left) Microscope picture of the bioreactor liquid after an experimental run. Visible are some yeast cells at the top of the image, a colony of hyphae and small dots representing *E. coli*. (right) Culture of hyphae after culturing YPD at 30°C for 2 days with an O-ring placed in the tube.

tube apparatus with tubes passing through the pump heads P1 and P2 together with the bioreactor.

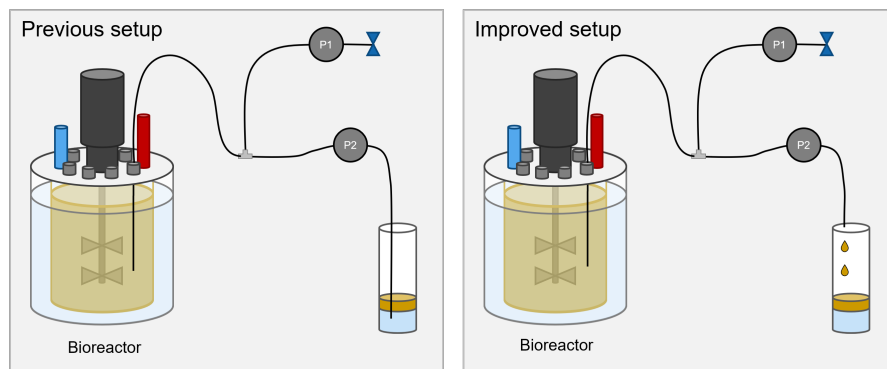


Figure 2.12: Improvement to the sampling setup by replacing the long needle at dilution vial 1 (left) with a shorter one (right). This results in a monoseptic environment containing the bioreactor and all tubes until the tip of the needle.

Sheath fluid used to run flow cytometry experiments contains a variety of biocides that prevents microbial growth inside the machine. Taking inspiration from this, the diluent buffer was supplemented with the biocide sodium azide at a concentration of $5.5\mu\text{M}$ to prevent extensive accumulation of bacteria and fungi in the sampling apparatus. We tested, whether the measurement of fluorescent reporters was impacted by this addition of sodium azide but found no difference between treated and untreated cells.

Lastly, the amylase autosampler is not directly connected to the bioreactor as shown in figure 2.10 but is connected to the first dilution vial (see figure 2.14). This removes any backspilling from the pneumatic manifold or valves into the bioreactor.

2.4 Matlab-bioreactor interfacing

Matlab® is used as the leader software for the photobioreactor platform. The code is designed in a modular way, allowing to weave different tasks into each other. Matlab is designed in a way that only one function can be executed at a time. This extends to functions such as “pause” that could be used to time different steps. We can circumvent this by using timers. A timer can be set for a specific task, e.g. initiating the sampling sequence every 30 minutes, but allows simultaneous other functions to run, until the initiation sequence needs triggering. Thus, each module is controlled by an internal timer, which allows the user to schedule the command execution of individual modules. Using timers allows the user to still interact with code and data outputs, as in the breaks between timers, no processing from Matlab is required. Six modules were coded in Matlab, spanning tasks from data communication with the bioreactor, data visualization, light control, automatic sampling, data analysis and closed-loop control (illustrated in figure 2.13).

A first module allows the connection and communication with the bioreactor control unit. This connection is opened using the open platform communication (OPC) standard. After creating the OPC object, all process parameters can be read from the bioreactor (temperature, pH, heating/cooling rate, stirring rate, gas flow rates, dissolved oxygen (DO), and process control parameters). This allows easy access to data for further analysis. Data is read once every second, for maximum information. For example, we were able to improve the speed of the temperature control, so that minute changes in the temperature profile could be observed and extracted.

A second module controls the light intensity of the light panels of the photobioreactor. The arduino adafruit neopixel library is used to create a connection from Matlab through an arduino to the light panels. The module allows the quick and easy change of the light intensity of all 384 LEDs in a single line of code. We built a module for the setup and updating of data plots. Plots show the temporal evolution of temperature, pH, DO, light intensity, pump flow rates, and flow cytometry measurements and are updated every three seconds.

A big portion of the development work happened in the module for automatic sampling and measurement, implementing the steps described in section 2.2.1. The sampling,

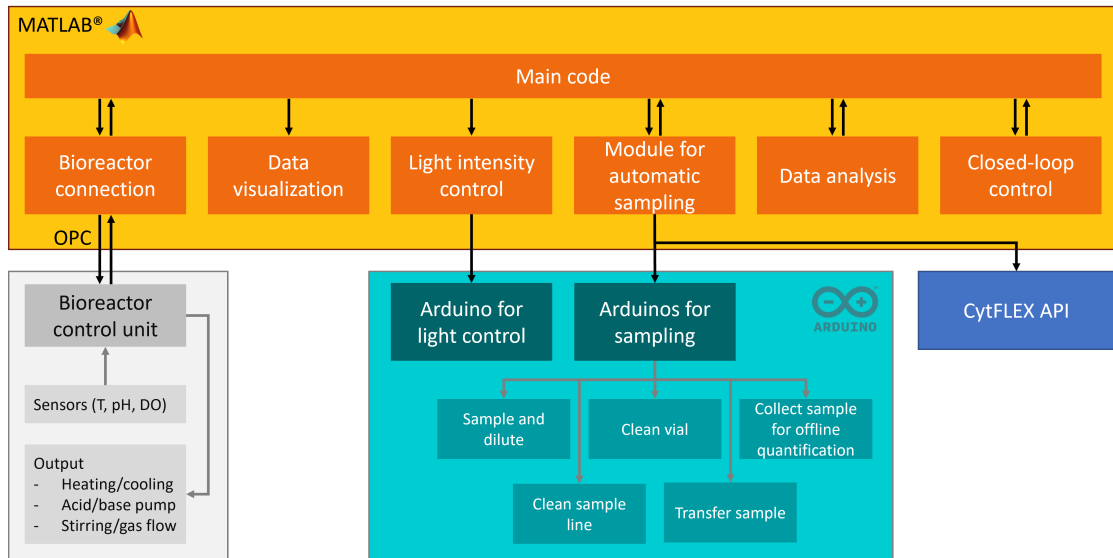


Figure 2.13: Chart illustrating the data and information flows required to run the photobioreactor platform. Unidirectional arrows indicate a flow of data or information in one direction only, whereas a bidirectional arrow indicates flow of information in both directions. Matlab acts as the leader software with a main code orchestrating the different core units. The bioreactor module connects through the OPC interface to the bioreactor control unit. The unit receives sensor data about temperature (T), pH and dissolve oxygen (DO) and regulates the heating and cooling rate, the acid and base pump rate and the stirring and gas flow rate accordingly, without interaction with the main Matlab code. Both, light intensity and automatic sampling are executed on arduino microcontrollers. Three arduinos are dedicated to sampling, one for each pump board and a third for valve switching of the amylase sampler. Finally, the Matlab module for automatic sampling additionally connects to the flow cytometry application processing interface (CytFLEX API).

dilution, measurement and cleaning procedure requires ≈ 12 minutes. During the sampling procedure 62 individual actions of pump turn on/off are triggered followed by 74 pump turn on/off actions for cleaning the vials. In order to allow user inputs during this time, the sequence of steps is coded in five submodules in the arduino itself. As an example, one such submodule takes the sample and dilution rate as input from Matlab and performs all sampling steps (figure 2.5a-d) whilst other submodule are designed for sample transfer (figure 2.5e) or cleaning vials (figure 2.5f). With all substeps executed on the arduinos, Matlab only needs to initiate these submodules in a single line. The physical wiring of the computer running Matlab to the different arduinos is shown in figure 2.14.

The data analysis module imports data from the flow cytometry data, saves them in data structs and calculates the next dilution rate for the next measurement including the actual pulse lengths for sample pump and diluent pump. Finally, the closed-loop control module takes information of the cell state and calculates the next light input.

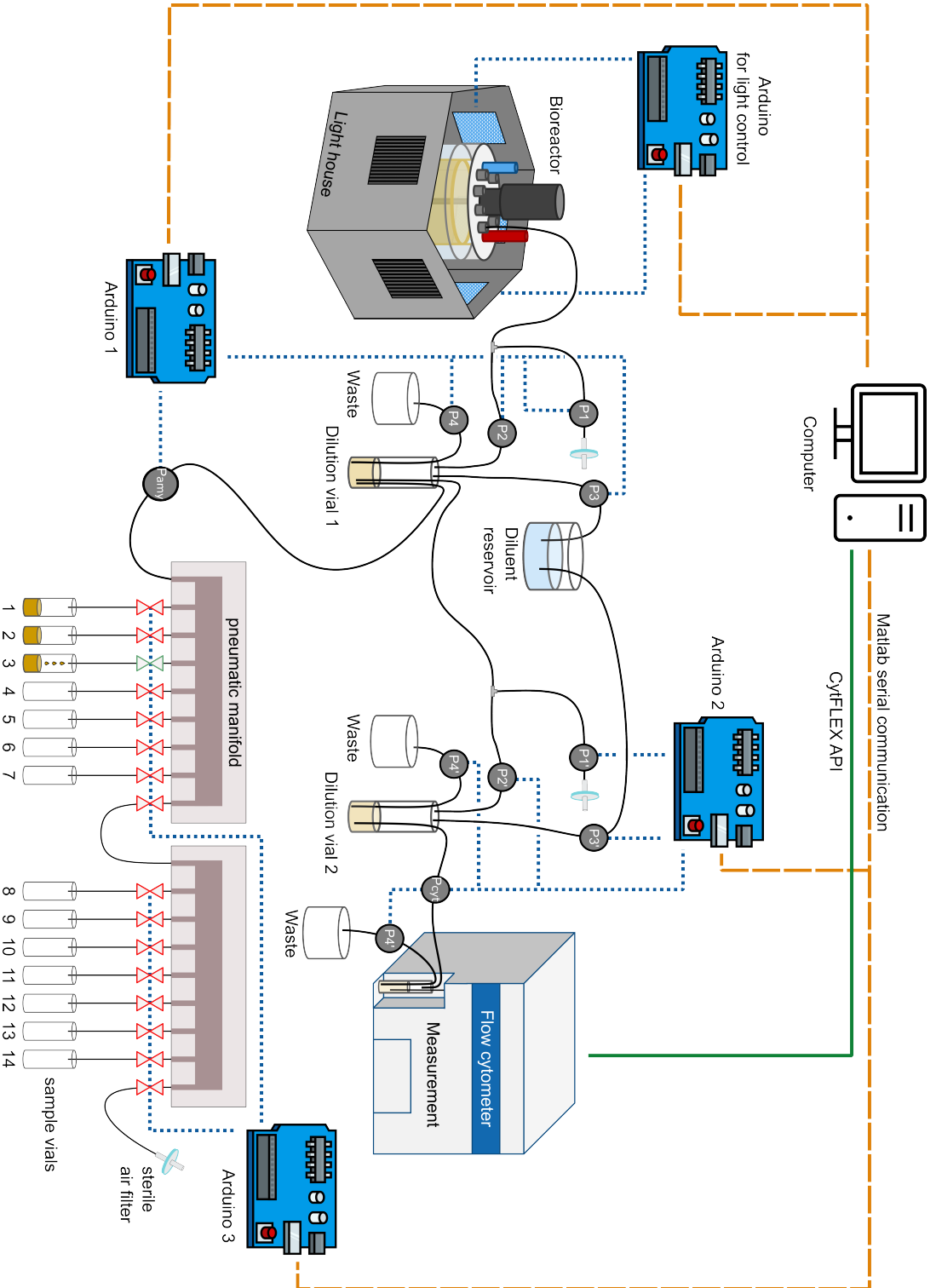


Figure 2.14: Full photobioreactor platform. Matlab communication to the four arduinos is shown in orange dashed lines. Arduino communication to pumps, valves and light panels is shown in blue dotted lines. Tube connections are drawn in solid black curved lines. Lastly the CytFLEX API is drawn in solid green. Arduino 1 is responsible for the communication to all pumps in the first dilution stage P1-P4 and the amylase sampler pump Pamy. Arduino 2 is responsible for communication to the set of pumps of the second dilutions stage P1-P4 and the pump for sample transfer to the flow cytometer Pcyt. Arduino 3 controls the pneumatic valves.

2.5 Conclusion

We have successfully developed a photobioreactor platform with added automatic sampling, dilution and quantification. This platform allows us to run closed-loop control experiments on a volume scale not demonstrated so far. The photobioreactor allows easy access to all hardware, while simultaneously shielding the bioreactor contents from any ambient light source. The light illumination device can also be equipped onto any glass-walled bioreactor, quickly transforming it into a photobioreactor.

The mechanical module for automatic sampling, dilution and measurements allows quantification of cellular parameters over a whole bioreactor run. While it is practically infeasible to manually sample and dilute every 30 minutes for 72 hours, this task can be achieved through automation. Standard sampling frequencies for fermentations in the academic research are in the order of once every 12 to 24 hours. Many biological processes can happen faster than these frequencies and not be picked up on. Our sampling system alone could thus be a useful tool for studying biological process during fermentation. After setup and inoculation of the bioreactor no user interactions are required until the end of the run, making this economically efficient and decreasing required labor costs.

The modular structure of the control code allows future modifications to the platform. This project implemented closed loop control using proportional integral controllers. Future projects could e.g. implement more advanced algorithms, such as model predictive control. Also changes in the light source, different pumps or measuring devices can be adjusted readily in the code.

2.6 Outlook

During the course of the project, some technical issues were observed that could be improved on in the future. The pump heads used for sampling, mixing, diluting and transferring liquids are cheap and are unreliable for long duration use. Specifically, the motor running the pumps broke in several instances with no prior indication of a failure. This results in some failed runs, when e.g. the waste pumps are not operating properly forcing an overflow of the liquid in the sample vials. In order to improve this, more reliable peristaltic pumps could be used. Furthermore, the implementation of automatic dilution was performed with peristaltic pumps. With the rise of affordable pipetting robots (e.g. Opentron), these steps could be replaced by a pipetting routine. While not necessary for closed-loop feedback, this could improve the accuracy of the cell density measurements.

Bibliography

- [1] Brigitte Gasser and Diethard Mattanovich. *Recombinant protein production in Yeast*. 2019. ISBN 9781493990238. doi: 10.1002/9783527678679.dg10657.
- [2] Olena P. Ishchuk, Jose L. Martinez, and Dina Petranovic. Improving the Production of Cofactor-Containing Proteins: Production of Human Hemoglobin in Yeast. In *Recombinant Protein Production in Yeast*, pages 243–264. Methods in edition, 2019.
- [3] Andreas Miliias-Argeitis, Sean Summers, Jacob Stewart-Ornstein, Ignacio Zuleta, David Pincus, Hana El-Samad, Mustafa Khammash, and John Lygeros. In silico feedback for in vivo regulation of a gene expression circuit. *Nature Biotechnology*, 29(12):1114–1116, 2011. ISSN 10870156. doi: 10.1038/nbt.2018.
- [4] Evan J. Olson, Lucas A. Hartsough, Brian P. Landry, Raghav Shroff, and Jeffrey J. Tabor. Characterizing bacterial gene circuit dynamics with optically programmed gene expression signals. *Nature Methods*, 11(4):449–455, 2014. ISSN 15487105. doi: 10.1038/nmeth.2884.
- [5] Andreas Miliias-Argeitis, Marc Rullan, Stephanie K. Aoki, Peter Buchmann, and Mustafa Khammash. Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth. *Nature Communications*, 7(May):1–11, 2016. ISSN 20411723. doi: 10.1038/ncomms12546. URL <http://dx.doi.org/10.1038/ncomms12546>.
- [6] Chetan Aditya, François Bertaux, Gregory Batt, and Jakob Ruess. A light tunable differentiation system for the creation and control of consortia in yeast. *Nature Communications*, 12(1), 2021. ISSN 20411723. doi: 10.1038/s41467-021-26129-7. URL <http://dx.doi.org/10.1038/s41467-021-26129-7>.
- [7] Sant Kumar and Mustafa Khammash. Platforms for Optogenetic Stimulation and Feedback Control. *Frontiers in Bioengineering and Biotechnology*, 10(June):1–15, 2022. doi: 10.3389/fbioe.2022.918917.
- [8] Harrison Steel, Robert Habgood, Ciarán Kelly, and Antonis Papachristodoulou. In situ characterisation and manipulation of biological systems with Chi.Bio. *PLoS Biology*, 18(7):1–12, 2020. ISSN 15457885. doi: 10.1371/journal.pbio.3000794.
- [9] Sebastian Tommi Tandar, Sachie Senoo, Yoshihiro Toya, and Hiroshi Shimizu. Optogenetic switch for controlling the central metabolic flux of *Escherichia coli*. *Metabolic Engineering*, 55(June):68–75, 2019. ISSN 10967184. doi: 10.1016/j.ymben.2019.06.002. URL <https://doi.org/10.1016/j.ymben.2019.06.002>.

-
- [10] R. N. Singh and Shaishav Sharma. Development of suitable photobioreactor for algae production - A review. *Renewable and Sustainable Energy Reviews*, 16(4): 2347–2353, 2012. ISSN 13640321. doi: 10.1016/j.rser.2012.01.026. URL <http://dx.doi.org/10.1016/j.rser.2012.01.026>.
- [11] Xianjun Chen, Renmei Liu, Zhengcai Ma, Xiaopei Xu, Haoqian Zhang, Jianhe Xu, Qi Ouyang, and Yi Yang. An extraordinary stringent and sensitive light-switchable gene expression system for bacterial cells. *Cell Research*, 26(7):854–857, 2016. ISSN 17487838. doi: 10.1038/cr.2016.74.
- [12] Fei Chang, Xianbing Zhang, Yu Pan, Youxue Lu, Wei Fang, Zemin Fang, and Yazhong Xiao. Light induced expression of β -glucosidase in *Escherichia coli* with autolysis of cell. *BMC Biotechnology*, 17(1):1–11, 2017. ISSN 14726750. doi: 10.1186/s12896-017-0402-1.
- [13] Evan M. Zhao, Yanfei Zhang, Justin Mehl, Helen Park, Makoto A. Lalwani, Jared E. Toettcher, and José L. Avalos. Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature*, 555(7698):683–687, 2018. ISSN 14764687. doi: 10.1038/nature26141.
- [14] Evan M. Zhao, Makoto A. Lalwani, Jhong Min Chen, Paulina Orillac, Jared E. Toettcher, and José L. Avalos. Optogenetic Amplification Circuits for Light-Induced Metabolic Control. *ACS Synthetic Biology*, 10(5):1143–1154, 2021. ISSN 21615063. doi: 10.1021/acssynbio.0c00642.
- [15] Makoto A. Lalwani, Hinako Kawabe, Rebecca L. Mays, Shannon M. Hoffman, and José L. Avalos. Optogenetic Control of Microbial Consortia Populations for Chemical Production. *ACS Synthetic Biology*, 10(8):2015–2029, 2021. ISSN 21615063. doi: 10.1021/acssynbio.1c00182.
- [16] D. Riesenberger and R. Guthke. High-cell-density cultivation of microorganisms. *Applied Microbiology and Biotechnology*, 51(4):422–430, 1999. ISSN 01757598. doi: 10.1007/s002530051412.
- [17] U. Schaefer, W. Boos, R. Takors, and D. Weuster-Botz. Automated sampling device for monitoring intracellular metabolite dynamics. *Analytical Biochemistry*, 270(1): 88–96, 1999. ISSN 00032697. doi: 10.1006/abio.1999.4048.
- [18] Justin Melendez, Michael Patel, Benjamin L. Oakes, Ping Xu, Patrick Morton, and Megan N. McClean. Real-time optogenetic control of intracellular protein concentration in microbial cell cultures. *Integrative Biology (United Kingdom)*, 6(3): 366–372, 2014. ISSN 17579708. doi: 10.1039/c3ib40102b.

Chapter 2. Platform development

- [19] Matthew J Maurer, Jeffrey M Skerker, Adam P Arkin, William Miller, Michael Biksacky, and Claudia M Huether-franken. Automated Bioreactor Sampling – Process Trigger Sampling for Enhancing Microbial Strain Characterization, 2015.
- [20] John P. Efromson, Shuai Li, and Michael D. Lynch. BioSamplr: An open source, low cost automated sampling system for bioreactors. *HardwareX*, 9:e00177, 2021. ISSN 24680672. doi: 10.1016/j.ohx.2021.e00177. URL <https://doi.org/10.1016/j.ohx.2021.e00177>.
- [21] Joaquín Gutiérrez Mena, Sant Kumar, and Mustafa Khammash. Dynamic cybergenetic control of bacterial co-culture composition via optogenetic feedback. *Nature Communications*, 13(1):1–16, 2022. ISSN 20411723. doi: 10.1038/s41467-022-32392-z.
- [22] Eppendorf. <https://bioflo120.eppendorf.com/>, 2023.
- [23] Laura B. Motta-Mena, Anna Reade, Michael J. Mallory, Spencer Glantz, Orion D. Weiner, Kristen W. Lynch, and Kevin H. Gardner. An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nature Chemical Biology*, 10(3):196–202, 2014. ISSN 15524469. doi: 10.1038/nchembio.1430.

3 Development of a *S. cerevisiae* strain for optogenetic production of α -amylase and UPR sensing

The biopharmaceutical market has been evergrowing and is expected to hit a project size of 900 billion USD in 2030 [1]. Currently, *S. cerevisiae* is one of the most commonly used production systems for fine chemicals and proteins and there is a continuous need to improve production efficiency [2]. Protein production consists of a multitude of intermediate steps similar to a conveyor belt manufacturing line, starting from transcription from the DNA template, running through translation, translocation, post-translational modifications and folding in the endoplasmic reticulum (ER) and finally being processed for delivery through cleavage of peptides, sorting and secretion [3].

When expressing a biopharmaceutical protein, limitations or bottlenecks can be encountered on a process level (e.g. insufficient aeration [4]) or along the previously described cellular conveyor belt production line [3]. During maximum gene overexpression one often observes that proteins will accumulate in a cell compartment with negative effects for the cell [5]. Engineering of the protein expression [6], folding [7] or secretory system [8] can remove those bottlenecks in the production line and increase product titers.

A limitation on the cellular level is generally present when the rate of incoming molecules into a specific cell compartment is higher than the rate of outgoing molecules. For optimal usage of the cellular resources, the rate of in- and out-going molecules in each cell compartment is equal. This can be achieved by finely tuning the reaction rates (e.g. promoter strength) of each step. Xu et al. [9] tuned the transcription and translation rates of three modules in the fatty acid synthesis pathway resulting in higher yield of the product. While this approach works nicely with given process and biological conditions, changes in the process conditions could result in sub-optimal production.

Introducing genetic feedback (i.e. dynamic pathway regulation) is a way to ensure

Chapter 3. Development of a *S. cerevisiae* strain for optogenetic production of α -amylase and UPR sensing

more robustness over a wider variety of process conditions [10]. In simple terms, a negative feedback from a downstream cell machinery to a more upstream flux is implemented, so that upon overburdening of the downstream machinery, new influx is reduced. Zhang et al. [11] engineered a promoter to dynamically sense the presence of a biodiesel intermediate and repress upstream metabolic fluxes, allowing the cell to reach homeostasis by converting the intermediate to the final biodiesel product (fatty acid ethyl ester) before derepressing the ingoing flux. We want to generalize the approach of genetic feedback by implementing the feedback outside the cell *in silico*. An *in silico* feedback control loop comprises the measurement of a cell state, computation of the next control input and application of a new stimuli.

We chose optogenetics as the stimuli of choice as it allows time-varying and tunable control over the gene expression of cells, while light is an inexpensive inducer [12, 13]. A variety of different optogenetic systems exist for *S. cerevisiae* [14–16]. Kennedy et al. [17] for example reconstituted the function of a split Gal4 transcription factor consisting of the Gal4 DNA-binding domain (GalBD) and Gal4 activation domain (GalAD). This was achieved by fusing GalBD to the cryptochrome 2 (Cry2) from *A. thaliana* and GalAD to its blue-light induced binding partner CIB1. Upon blue light illumination, Cry2 and Cib1 dimerized and increased expression from a Gal-responsive promoter was observed in comparison to the dark state control. We chose to work with EL222 as optogenetic actuator due to the vast expertise of our group in this area [18–20]. Our optogenetic transcription factor consists of the blue-light-sensitive DNA-binding domain EL222 fused to the Msn2 activation domain (Msn2AD-EL222) [20, 21]. Additionally, EL222 has also been used in high cell density culture [22], validating our choice for this project.

As a target protein, we chose the model protein α -amylase. Amylase is a relatively large protein with a size of 49-51 kDa and requires the formation of four disulfide bonds between pairs of cysteine, making it a hard-to-fold protein [23, 24]. Amylase production is subject to the bottlenecks discussed previously [8, 23], making it an ideal target protein. It has been used in a variety of protein production screens [8, 23–25] and commercial quantification assays are available.

The last part for the *in silico* feedback control loop is the measurement of a cellular state to feedback on. One option is to develop a biosensor that is specific for the product of interest as shown for small molecules such as muconic acid [26], xylose [27] or branched chain amino acid metabolism [28] as well as for some proteins such as glucose dehydrogenase [29], β -lactamase [30], anti-apoptosis protein BCL-2, IgG1 Fc domain and HER2 receptor [31]. The development of a biosensor with good sensitivity,

specificity and a high dynamic range properties is challenging [32] and could limit the applicability of this approach to a large selection of proteins of interest. Thus, we focus our attention on the more general target that is the unfolded protein response (UPR). The UPR coordinates the cellular response in yeast to elevated levels of unfolded proteins in the ER and can trigger the expression of chaperones, cofactors and the ER-associated degradation (ERAD) pathway [33]. It is thus central to the proper folding of proteins and is often tuned to enhance protein production [25, 34] hinting towards an ideal UPR induction level. The UPR also integrates over disturbances stemming from growth kinetics, nutrient availability and protein folding state. We thus believe that implementing the sensing channel through the UPR allows this approach to be widely generalizable beyond the expression of α -amylase. As the basis for sensing, we use the promoter topology from Merksamer et al. [35].

3.1 Methods

3.1.1 Media

YPD and SD-URA medium for transformations was prepared according to the Clontech Yeast Protocols Handbook [36]. The YPE medium contains 10 g/L yeast extract, 20 g/L peptone and 25 ml/L ethanol. Just before usage, acetaldehyde is added to YPE at a concentration of 0.01 % to reduce the lag time in growth observed otherwise. The SD-2xSCAA medium [37] (designed for heterologous protein secretion) is used for all precultures and experiments and contains 20 g/L D-glucose, 6.9 g/L yeast nitrogen base without amino acids, folic acid and riboflavin, 1 g/L bovine serum albumin (BSA), amino acids as listed in supplementary table S4, 5.4 g/L Na_2PO_4 and 8.56 g/L $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (pH = 6.0 by H_2SO_4). Initial experiments for characterization were performed without inositol supplementation. Final experiments were performed in media supplemented with *myo*-inositol to a final inositol concentration of 20mg/l. To prepare SD or YPD agar plates, 20 g/L agar is added to the medium. For SD-2xSCAA plates, BSA was not added.

For precultures and characterization experiments, the antibiotics Hygromycin B and Geneticin G418 were added at a final concentration of $200 \frac{\mu\text{g}}{\text{ml}}$. For transformations with Zeocin resistance YPD plates were poured with Zeocin at a concentration of $100 \frac{\mu\text{g}}{\text{ml}}$. Zeocin was not added during experiments as it is highly unstable when exposed to light. Amino acids, α -amylase, D-glucose, phosphate buffers, *myo*-inositol and BSA were purchased from Sigma-Aldrich. Geneticin was purchased from Gibco, Zeocin and Hygromycin B from Invitrogen, agar, yeast extract, peptone from BD, yeast nitrogen

Chapter 3. Development of a *S. cerevisiae* strain for optogenetic production of α -amylase and UPR sensing

base from Formedium and ammonium sulfate from Roth.

3.1.2 Plasmid construction and strains

The genotype of the background, all further strains and plasmids is shown in supplementary table S1 and S3. The phototoxicity experiments (figure 3.8) were performed with yMB9, all other experiments with yMB44. The yeast production strain *S. cerevisiae* Cen.PK 530-1C (kindly provided by Jens Nielsen, Chalmers University of Technology, Sweden) was used for all work in this publication. Cen.PK 530-1C has a triose phosphate isomerase (*tpi1*) deletion and is thus unable to metabolize glucose and can only grow on ethanol. The high-copy production plasmid pAlphaAmyCPOT (kindly provided by Jens Nielsen, Chalmers University of Technology, Sweden [24]) contains a *S. pombe* POT1 expression cassette, enabling metabolization of glucose in the *tpi1*-deletion background, thus allowing for selection of plasmid retention by using glucose as carbon source.

We created pYMB10, a plasmid with amylase expression under optogenetic EL222 control, by replacing the original *tpi1* promoter of the pAlphaAmyCPOT plasmid with the optogenetic P_{EL222} promoter. To do this, the pAlphaAmyCPOT plasmid was cut with FseI and KpnI and the 8.5 kbp fragment gel-extracted. The optogenetic P_{EL222} promoter was prepared by PCR amplifying pYTKmk216 with the oligos oMB020 and oMB021 and cutting with FseI and KpnI. Finally both parts were ligated. This results in the plasmid pYMB10 with the α -amylase expression cassette consisting of the P_{EL222} promoter, the amylase gene and the TPI1 terminator under POT1 selection.

The Msn2AD-EL222 expression cassette was expressed from the constitutive ScrPL18B promoter and integrated in the URA3 locus by uracil selection (pYTKmk48). The transcriptional reporter pYMB12v consists of the yellow fluorescent protein Venus [38] tagged with an N-degron Ubiquitin (Ubi-Y, [39]), under the control of the P_{EL222} promoter and was genomically integrated in the HO locus and selected for by Hygromycin B. It was constructed by combining the MoClo parts pYTK042, pYTK045, pYTK054, pYTKmk110 and pYTKmk217 using the Molecular cloning protocol [39].

We implemented a UPR sensor (UPRS) by placing mScarlet-I [40] tagged with an N-degron Ubiquitin (Ubi-Y, [39]) under the control P_{4xUPRE}, resulting in pYMB15m. The UPRS was integrated in the YIRC Δ 6 locus [41] and selected for by Zeocin. It was constructed by combining the MoClo parts pYTK042, pYTK054, pYTKmk227, pYTKmk110 and pYTKmk123.

Top 10 competent *E. coli* was used for all molecular cloning. Plasmid construction was

done using standard molecular cloning techniques or using the yeast molecular cloning toolkit (YTK) from Lee et al. [39]. Yeast transformations were performed according to the protocol from Gietz and Schiestl [42]. For genomic integrations plasmids were cut with Not-I HF (New England Biolabs) and purified using with a DNA Clean & Concentrator Kit (Zymo Research). Successful transformation was checked with the genotyping protocol from Lööke et al. [43].

3.1.3 Culture conditions

All precultures and experiments were grown at 30°C in SD-2xSCAA. Experiments for characterization were performed at two different volume scales. Experiments for initial characterization of the UPRS and experiments to determine the ideal concentration of inositol were performed in 10ml culture tubes in 3ml media in a shaking incubator (New Brunswick). Optogenetic characterization experiments were performed in a water bath setup [18] in 4mL cultures with individual blue light illumination ($\lambda = 450nm$) and stirred at 900 rpm.

For all experiments, yeast strains were streaked from the frozen glycerol stock on a SD-2xSCAA plate and incubated at 30°C for 3 days. Precultures were inoculated from single colonies into 10 ml of SD-2xSCAA media in a 125 baffled flask and incubated overnight to an $OD_{600} \approx 1.25$. The preculture cell density was measured by flow cytometry and experiments started at a cell density of $200 \frac{cells}{\mu l}$ ($OD_{600} \approx 0.02$). All experiments were shielded from ambient light to avoid unwanted optogenetic activation of the circuit.

3.1.4 Flow cytometry

Fluorescence and cell density measurements were performed on a Beckman Coulter CytoFLEX S. The internal quality control program was run before experiments using CytoFLEX Daily QC Fluorospheres. Events were gated in the FSC-A and SSC-A channels to remove measurements of debris (see figure 3.1). Samples were diluted with PBS to cell densities of $100-4'000 \frac{cells}{\mu l}$. All fluorescence measurements are reported as the mean of the gated population in arbitrary units (a.u.).

3.2 Final design

We engineered *S. cerevisiae* to express the hard-to-fold, secreted protein α -amylase under control of an optogenetic transcription factor (TF), consisting of the blue-light-sensitive DNA-binding domain EL222 fused to the Msn2 activation domain (Msn2AD-EL222)

Chapter 3. Development of a *S. cerevisiae* strain for optogenetic production of α -amylase and UPR sensing

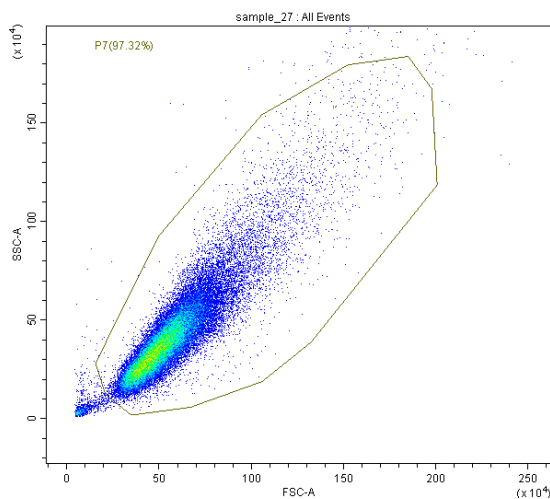


Figure 3.1: Gating strategy for all experiments. An elliptical gate is drawn in the forward scatter (FSC-A) - side scatter (SSC-A) plane. The gate is chosen sufficiently large to remove any debris that is present in the sample and measured by the flow cytometer.

[20, 21]. Msn2AD-EL222 is expressed constitutively from pScRPL18B and results in EL222 monomers. Under blue light, EL222 dimerizes, enabling it to bind to its target promoter P_{EL222} and thus activate gene expression. In the dark, Msn2AD-EL222 reverts to its transcriptionally inactive state within minutes [20, 21], enabling highly dynamic expression regulation. To dynamically measure α -amylase transcription, we further constructed a transcriptional reporter (referred to as TR) by placing the fluorescent protein Venus under control of P_{EL222} .

The exact mechanism for UPR activation in yeast is reported in Bernales et al. [33]. In the inactive state, Hac1 mRNA is constitutively transcribed. Unspliced Hac1 mRNA contains an intron forming a stem loop to the 5' UTR, preventing full translation and additionally reducing translation rate. Ire1 (Inositol requiring enzyme 1) is a sensor in the endoplasmic reticulum. The cytosolic side of Ire1 consists of a kinase and an endoribonuclease. Upon presence of unfolded proteins, those unfolded proteins bind to Ire1 and make them cluster on the ER-lumen-cytosol surface. Due to the proximity of the kinases of Ire1, the kinases transautophosphorylate each other and thus result in an activation of the endoribonucleases, which excise the Hac1mRNA intron. Ribosomes are able to now efficiently translate the spliced Hac1mRNA to functional Hac1p transcription factor. Downstream targets of Hac1p are chaperones (e.g. Kar2), oxidoreductases, glycosylating enzymes, and ER degradation components.

We leverage this mechanism to implement a UPR sensor (UPRS) by incorporating Hac1p responsive elements [35] in a CYC1 minimal promoter [44] (P_{4xUPRE}) that drives

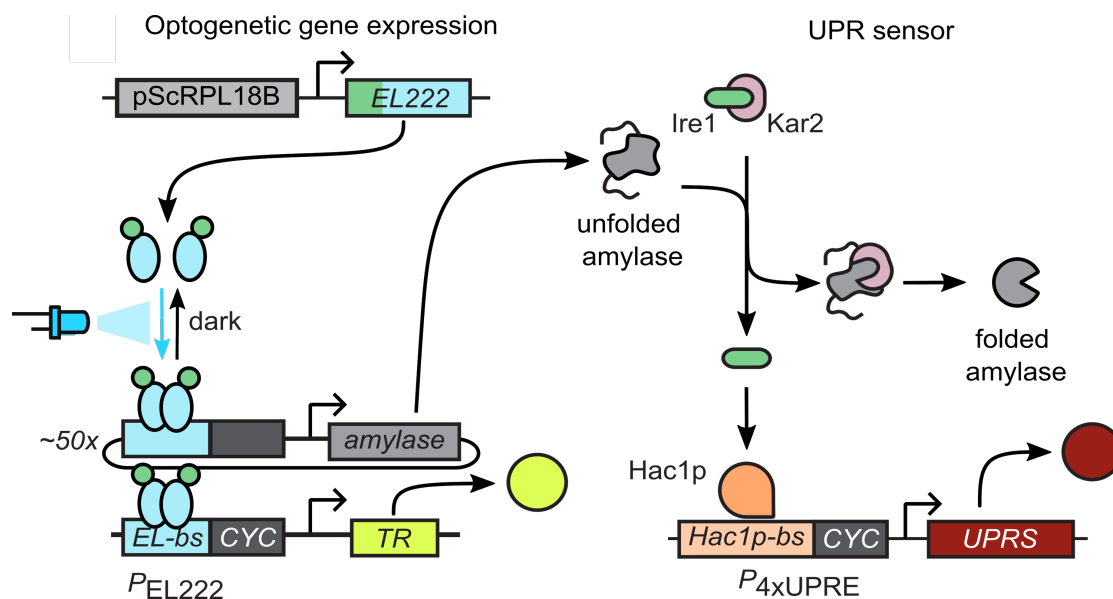


Figure 3.2: Illustration of the optogenetic expression system and UPR sensor. Constitutive expression of Msn2AD-EL222 from the genome results in Msn2AD-EL222 monomers that reversibly dimerize under blue light and allows binding to the Msn2AD-EL222 binding site (EL-bs) activating transcription from the P_{EL222} promoter. P_{EL222} drives the expression of α -amylase and transcriptional reporter (TR), from a high-copy plasmid and chromosomal integration respectively. Unfolded proteins trigger the unfolded protein response (UPR), which leads to expression of the Hac1p transcription factor, subsequent binding to the P_{4xUPRE} and expression of a UPR sensor (UPRS) from a chromosomal integration.

the expression of the fluorescent protein mScarlet-I.

The strain containing the constitutive EL222 expression cassette, integrated transcriptional reporter, integrated UPR sensor and the CPOT plasmid with α -amylase under EL222 control is named yMB44.

3.3 Results

The following section details the testing of the biological parts required for cybergenetic control.

Figure 3.3 shows the mean fluorescence level of the TR over time for cells grown in the dark, at intermediate duty cycles 15 min/6 h and with constant blue light illumination. We observed a noticeable difference between dark and light conditions after 30 minutes. Additionally, the promoter shows a high dynamic range with a 300-fold increase in the level of TR after 12 hours. Cells exposed to light pulses, on the other hand, reach an intermediate expression level with a step increase due to optogenetic gene expression and a gradual decline pattern corresponding to degradation and dilution of the reporter.

Chapter 3. Development of a *S. cerevisiae* strain for optogenetic production of α -amylase and UPR sensing

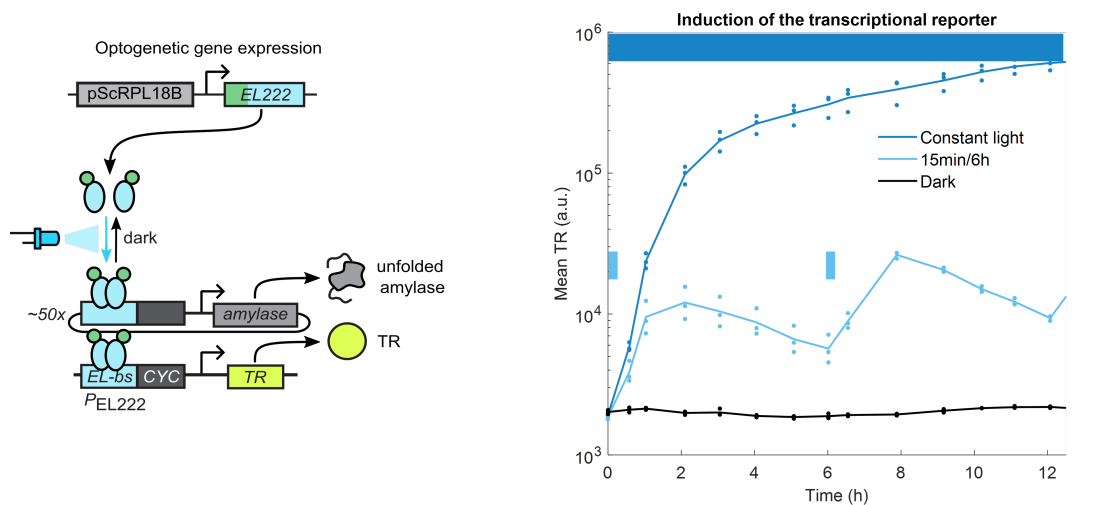


Figure 3.3: **left** Optogenetic gene expression cassette. **right** Dynamic response of the mean fluorescence level of TR to no light, 15 minutes of light every 6 hours (as indicated in small blue bars) and to constant light measured in the water bath setup (n=3, lines represent means of replicates)

We tested the functionality of the UPRS with dithiothreitol (DTT). DTT is a global reducing agent that breaks intramolecular disulfide bonds, leading to protein unfolding and activation of the UPR. Figure 3.4 shows a graded response of the UPRS to DTT concentrations between 0.2 and 2 mM and changes in the reporter levels can be observed 1 hour after the addition of DTT.

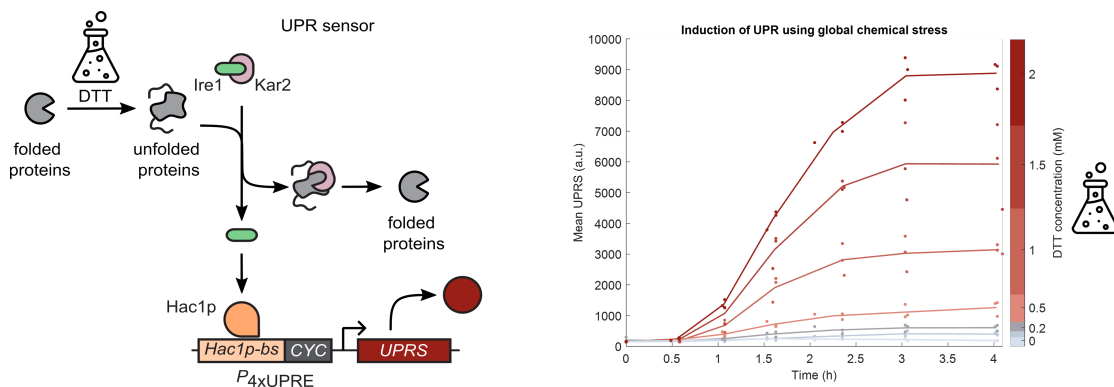


Figure 3.4: **left** Unfolded protein response sensor (UPRS) in response to global unfolding of proteins triggered by DTT. **right** Dynamic response of the mean fluorescence level of UPRS in the dark to different concentrations of the global chemical stressor dithiothreitol (DTT) in culture tubes (n=3, lines represent means of replicates)

Leber et al. [45] reported that the unfolded protein response could be activated upon inositol starvation. The SD-2xSCAA media that is used for all experiments already contains 2 mg/L inositol, which stems from the yeast nitrogen base (YNB), but the

additional availability of amino acids could mean that an inositol deficiency is encountered by the cells after some growth. We thus decided to track the level of UPRS over time in the SD-2xSCAA media. Figure 3.5 shows that cells grown in normal, unsupplemented SD-2xSCAA media ($c_{inositol} = 2 \text{ mg/L}$) start expressing the unfolded protein response sensor after 17 hours indicating a UPR induction and inositol deficiency. To counteract this inositol-dependant induction, we supplemented the media with additional *myo*-inositol to final concentrations of inositol of 4-200 mg/L and tracked the UPRS response again. Supplementation of the media with inositol delayed the UPRS response and decreased the induction fold-change. Given these results we supplemented our media to a final inositol concentration of 20 mg/l for all subsequent experiments. Even with supplementation, a slight induction of the UPRS is observed after 20 hours. This is due to the onset of stationary phase, where cells are dividing slower, resulting in decreased apparent dilution rates and higher total levels of UPRS.

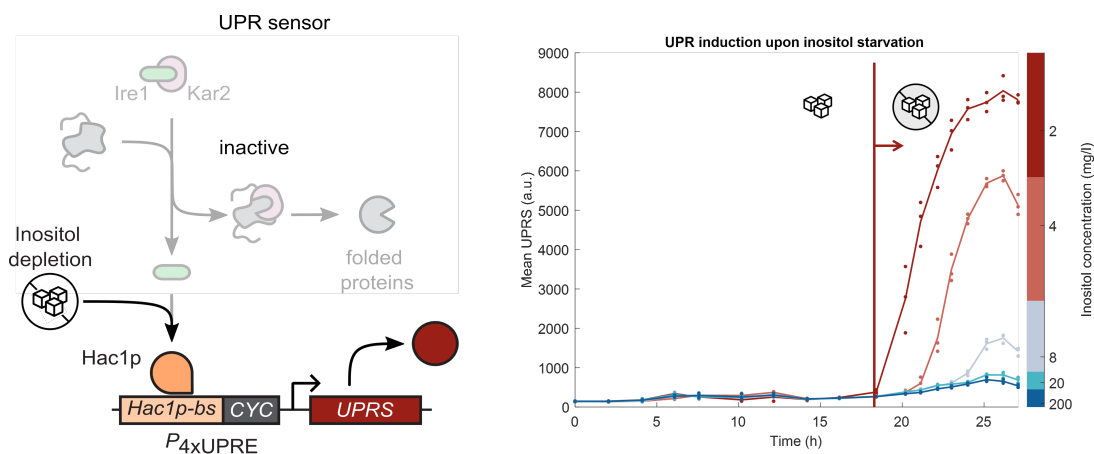


Figure 3.5: **left** Hac1p activation is modulated by inositol availability and can trigger the UPRS independently of the presence of unfolded proteins. **right** Dynamic response of the mean fluorescence level of UPRS in the dark to different concentrations of *myo*-inositol in culture tubes ($n=3$, lines represent means of replicates). UPR induction upon inositol starvation is observed at different times for different initial concentrations of inositol. The initial SD-2xSCAA media contains 2mg/l of inositol and was supplemented with *myo*-inositol to reach the inositol concentrations given on the right.

In this supplemented media, we tested whether the optogenetic expression of α -amylase is working and further, whether its expression would trigger the UPR. Figure 3.6 show that upon blue light illumination, the α -amylase proxy TR is expressed and the UPRS levels follow this rise. Elevated levels of TR also correspond to increased amounts of amylase (see figure 3.7). This indicates that expression of α -amylase induces the UPR. Turning off the blue light stopped the transcription of amylase, reducing the influx of new unfolded proteins to the ER and decreasing the level of UPRS. Light pulses as short as 15 minutes were able to induce the UPRS five-fold. A decrease in UPRS was observed roughly 1 hour after turning the light off, highlighting the reversibility and fast dynamics

Chapter 3. Development of a *S. cerevisiae* strain for optogenetic production of α -amylase and UPR sensing

of the optogenetic expression system and UPR. Light-mediated activation was graded and could be adjusted by increasing the duration of blue light illumination.

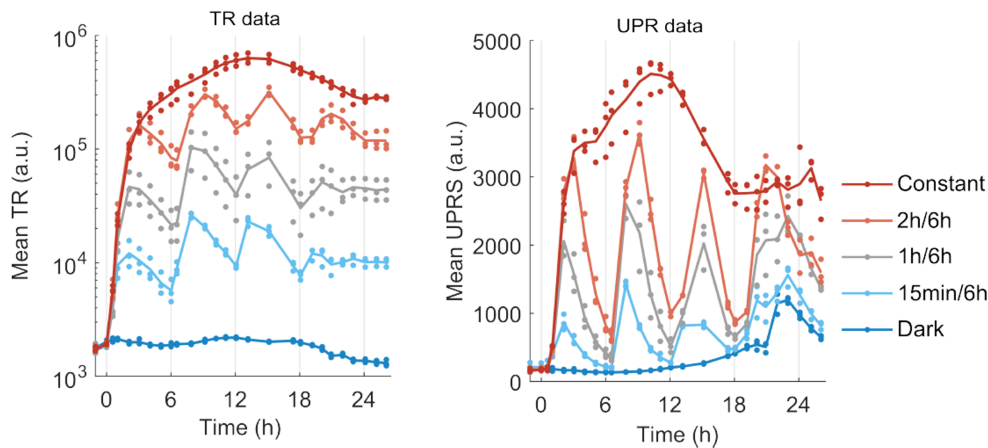


Figure 3.6: Dynamic response of the mean fluorescence level of TR **left** and UPRS **right** to no light, pulses of 15 minutes, 1 hour and 2 hours of light and constant light measured in the water bath setup (n=3, lines represent means of replicates).

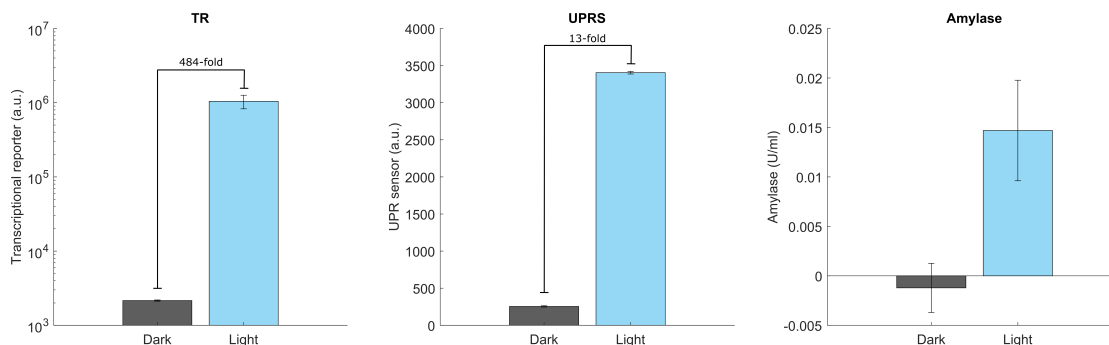


Figure 3.7: Comparison of the mean fluorescence level of transcriptional reporter (TR), unfolded protein response sensor (UPRS) and amylase concentration in tube cultures of inositol-supplemented SD-2xSCAA media measured after 17 hours of growth. Cells were grown either in the dark or with constant light illumination (n=2).

We performed further controls to ensure that UPRS is not triggered through phototoxicity. To this end, we integrated the UPRS in a strain lacking the amylase expression plasmid and tested whether the UPRS would sense induction by DTT and light. Figure 3.8 shows that only induction with DTT triggered a detectable change in UPRS expression. Shining blue light on cells results in no detectable difference in UPRS compared to cells grown in the dark. The maximum used irradiance of $660 \frac{\mu W}{cm^2}$ also poses an upper limit of light irradiance that can be supplied in the bioreactor (see figure 2.4), so that light toxicity is also not experienced in the bioreactor at maximum irradiance.

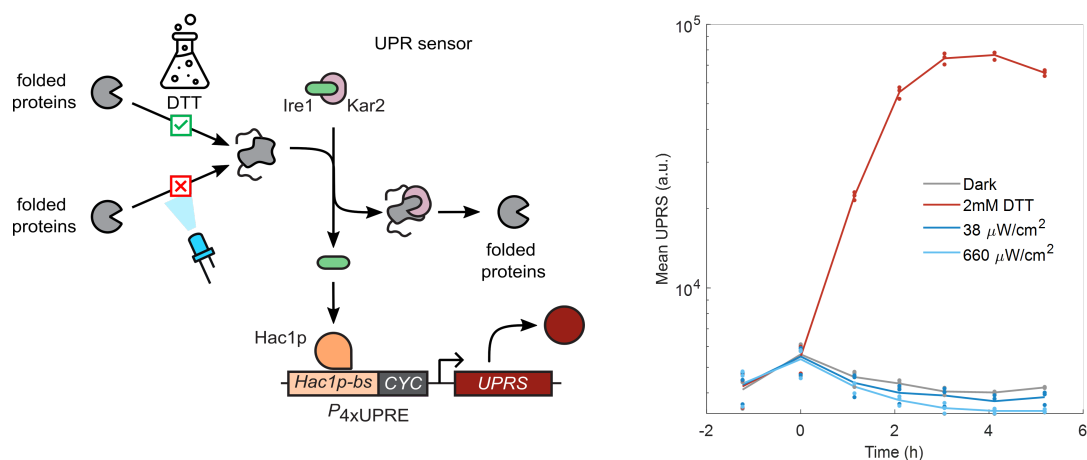


Figure 3.8: **left** UPRS in response to DTT and blue light. **right** Progression of the unfolded protein response sensor over time in a strain lacking the α -amylase gene expression cassette (yMB9) in response to different inducers. Cells were induced with 2mM DTT or blue light (38 and 660 $\mu\text{W}/\text{cm}^2$) at $t_0 = 0\text{h}$ (n=3)

3.4 Conclusion

In this chapter, we described the development of *S. cerevisiae* to optogenetically express α -amylase, a hard-to-fold model protein [23, 24], as well as a fluorescent sensor of UPR activity. We characterized this strain and showed reversible, tunable, optogenetic activation and relaxation of the UPR with blue light. We optimized the growth conditions to ensure the measurability of the reporters over a longer time horizon.

This biological circuit alone could be a valuable tool to study the dynamics of the UPR *in vivo*. Most experimental studies of the UPR use knockout strains or global chemical stressors like tunicamycin and DTT to trigger a cellular response [35, 46, 47]. While inducing the UPR efficiently, those chemicals do not alter the influx of new unfolded proteins into the ER but rather lead to global protein unfolding. In order to understand microbial cell factories for protein production better, a UPR induction mechanism mimicking the normal process could generate novel biological insights. We believe that the usage of optogenetic expression of hard-to-fold proteins are ideal for this purpose.

Bibliography

- [1] Next Move Strategy Consulting and Matej Mikulic. Biopharmaceutical Market by Product (Monoclonal Antibody, Interferon, Insulin, Growth and Coagulation Factor, Erythropoietin, Vaccine, Hormone, And Others) and by Application (Oncology, Blood Disorder, Metabolic Disease, Infectious Disease, Cardiovascular. Available at <https://www.statista.com/statistics/1293077/global-biopharmaceuticals-market-size/>. Accessed July 14, 2022., 2022.
- [2] Mingtao Huang, Jichen Bao, and Jens Nielsen. Biopharmaceutical protein production by *Saccharomyces cerevisiae* : current state and future prospects. *Pharmaceutical Bioprocessing*, 2(2):167–182, 2014. ISSN 2048-9145. doi: 10.4155/pbp.14.8.
- [3] Carissa L. Young and Anne S. Robinson. Protein folding and secretion: Mechanistic insights advancing recombinant protein production in *S. cerevisiae*. *Current Opinion in Biotechnology*, 30:168–177, 2014. ISSN 18790429. doi: 10.1016/j.copbio.2014.06.018. URL <http://dx.doi.org/10.1016/j.copbio.2014.06.018>.
- [4] A. Kapat, J. K. Jung, and Y. H. Park. Enhancement of glucose oxidase production in batch cultivation of recombinant *Saccharomyces cerevisiae*: Optimization of oxygen transfer condition. *Journal of Applied Microbiology*, 90(2):216–222, 2001. ISSN 13645072. doi: 10.1046/j.1365-2672.2001.01233.x.
- [5] Claudia Rubio, David Pincus, Alexei Korennykh, Sebastian Schuck, Hana El-Samad, and Peter Walter. Homeostatic adaptation to endoplasmic reticulum stress depends on Ire1 kinase activity. *Journal of Cell Biology*, 193(1):171–184, 2011. ISSN 00219525. doi: 10.1083/jcb.201007077.
- [6] Yoichiro Ito, Takao Kitagawa, Mamoru Yamanishi, Satoshi Katahira, Shingo Izawa, Kenji Irie, Makoto Furutani-Seiki, and Takashi Matsuyama. Enhancement of protein production via the strong DIT1 terminator and two RNA-binding proteins in *Saccharomyces cerevisiae*. *Scientific Reports*, 6(July):1–9, 2016. ISSN 20452322. doi: 10.1038/srep36997.
- [7] Jorg C. de Ruijter, Essi V. Koskela, and Alexander D. Frey. Enhancing antibody folding and secretion by tailoring the *Saccharomyces cerevisiae* endoplasmic reticulum. *Microbial Cell Factories*, 15(1):1–18, 2016. ISSN 14752859. doi: 10.1186/s12934-016-0488-5.
- [8] Mingtao Huang, Guokun Wang, Jiufu Qin, Dina Petranovic, and Jens Nielsen. Engineering the protein secretory pathway of *Saccharomyces cerevisiae* enables im-

- proved protein production. *PNAS*, 115(47):E11025–E11032, 2018. doi: 10.1073/pnas.1809921115.
- [9] Peng Xu, Qin Gu, Wenya Wang, Lynn Wong, Adam G.W. Bower, Cynthia H. Collins, and Mattheos A.G. Koffas. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nature Communications*, 4:1408–1409, 2013. ISSN 20411723. doi: 10.1038/ncomms2425. URL <http://dx.doi.org/10.1038/ncomms2425>.
- [10] Francesca Ceroni, Alice Boo, Simone Furini, Thomas E. Goroehowski, Olivier Borkowski, Yaseen N. Ladak, Ali R. Awan, Charlie Gilbert, Guy Bart Stan, and Tom Ellis. Burden-driven feedback control of gene expression. *Nature Methods*, 15(5):387–393, 2018. ISSN 15487105. doi: 10.1038/nmeth.4635.
- [11] Fuzhong Zhang, James M. Carothers, and Jay D. Keasling. Design of a dynamic sensor-regulator system for production of chemicals and fuels derived from fatty acids. *Nature Biotechnology*, 30(4):354–359, 2012. ISSN 10870156. doi: 10.1038/nbt.2149.
- [12] Andreas Miliadis-Argeitis, Marc Rullan, Stephanie K. Aoki, Peter Buchmann, and Mustafa Khammash. Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth. *Nature Communications*, 7(May):1–11, 2016. ISSN 20411723. doi: 10.1038/ncomms12546. URL <http://dx.doi.org/10.1038/ncomms12546>.
- [13] Sebastian Tommi Tandar, Sachie Senoo, Yoshihiro Toya, and Hiroshi Shimizu. Optogenetic switch for controlling the central metabolic flux of *Escherichia coli*. *Metabolic Engineering*, 55(June):68–75, 2019. ISSN 10967184. doi: 10.1016/j.ymben.2019.06.002. URL <https://doi.org/10.1016/j.ymben.2019.06.002>.
- [14] Sae Shimizu-Sato, Enamul Huq, James M. Tepperman, and Peter H. Quail. A light-switchable gene promoter system. *Nature Biotechnology*, 20(10):1041–1044, 2002. ISSN 10870156. doi: 10.1038/nbt734.
- [15] Devin Strickland, Yuan Lin, Elizabeth Wagner, C. Matthew Hope, Josiah Zayner, Chloe Antoniou, Tobin R. Sosnick, Eric L. Weiss, and Michael Glotzer. TULIPs: Tunable, light-controlled interacting protein tags for cell biology. *Nature Methods*, 9(4):379–384, 2012. ISSN 15487091. doi: 10.1038/nmeth.1904.
- [16] Oana I. Lungu, Ryan A. Hallett, Eun Jung Choi, Mary J. Aiken, Klaus M. Hahn, and Brian Kuhlman. Designing Photoswitchable Peptides Using the AsLOV2 Domain. *Chemistry and Biology*, 19(4):507–517, 2012. ISSN 10745521. doi: 10.1016/j.chembiol.2012.02.006. URL <http://dx.doi.org/10.1016/j.chembiol.2012.02.006>.

Chapter 3. Development of a *S. cerevisiae* strain for optogenetic production of α -amylase and UPR sensing

- [17] Matthew J. Kennedy, Robert M. Hughes, Leslie A. Peteya, Joel W. Schwartz, Michael D. Ehlers, and Chandra L. Tucker. Rapid blue-light-mediated induction of protein interactions in living cells. *Nature Methods*, 7(12):973–975, 2010. ISSN 15487091. doi: 10.1038/nmeth.1524.
- [18] Dirk Benzinger and Mustafa Khammash. Pulsatile inputs achieve tunable attenuation of gene expression variability and graded multi-gene regulation. *Nature Communications*, 9(2018):1–38, 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05882-2. URL <http://dx.doi.org/10.1038/s41467-018-05882-2>.
- [19] Sant Kumar, Marc Rullan, and Mustafa Khammash. Rapid prototyping and design of cybergenetic single-cell controllers. *Nature Communications*, 12(1), 2021. ISSN 20411723. doi: 10.1038/s41467-021-25754-6. URL <http://dx.doi.org/10.1038/s41467-021-25754-6>.
- [20] Dirk Benzinger, Serguei Ovinnikov, and Mustafa Khammash. Synthetic gene networks recapitulate dynamic signal decoding and differential gene expression. *Cell Systems*, 13(5):353–364.e6, 2022. ISSN 24054720. doi: 10.1016/j.cels.2022.02.004. URL <https://doi.org/10.1016/j.cels.2022.02.004>.
- [21] Laura B. Motta-Mena, Anna Reade, Michael J. Mallory, Spencer Glantz, Orion D. Weiner, Kristen W. Lynch, and Kevin H. Gardner. An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nature Chemical Biology*, 10(3):196–202, 2014. ISSN 15524469. doi: 10.1038/nchembio.1430.
- [22] Evan M. Zhao, Yanfei Zhang, Justin Mehl, Helen Park, Makoto A. Lalwani, Jared E. Toettcher, and José L. Avalos. Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature*, 555(7698):683–687, 2018. ISSN 14764687. doi: 10.1038/nature26141.
- [23] Keith E.J. Tyo, Zihe Liu, Dina Petranovic, and Jens Nielsen. Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress. *BMC Biology*, 10, 2012. ISSN 17417007. doi: 10.1186/1741-7007-10-16.
- [24] Zihe Liu, Keith E.J. Tyo, Jose L. Martinez, Dina Petranovic, and Jens Nielsen. Different Expression Systems for Production of Recombinant Proteins in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering*, 2012. doi: 10.1002/bit.24409.
- [25] Mari Valkonen, Merja Penttilä, and Markku Saloheimo. Effects of inactivation and constitutive expression of the unfolded-protein response pathway on protein production in the yeast *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 69(4):2065–2072, 2003. ISSN 00992240. doi: 10.1128/AEM.69.4.2065-2072.2003.

- [26] John M. Leavitt, James M. Wagner, Cuong C. Tu, Alice Tong, Yanyi Liu, and Hal S. Alper. Biosensor-Enabled Directed Evolution to Improve Muconic Acid Production in *Saccharomyces cerevisiae*. *Biotechnology Journal*, 12, 2017.
- [27] Wei Suong Teo and Matthew Wook Chang. Bacterial XylRs and synthetic promoters function as genetically encoded xylose biosensors in *Saccharomyces cerevisiae*. *Biotechnology Journal*, 10(2):315–322, 2015. ISSN 18607314. doi: 10.1002/biot.201400159.
- [28] Yanfei Zhang, Jeremy D. Cortez, Sarah K. Hammer, César Carrasco-López, Sergio García Echaury, Jessica B. Wiggins, Wei Wang, and José L. Avalos. Biosensor for branched-chain amino acid metabolism in yeast and applications in isobutanol and isopentanol production. *Nature Communications*, 13(1):1–14, 2022. ISSN 20411723. doi: 10.1038/s41467-021-27852-x.
- [29] Zhong Guo, Wayne A. Johnston, Jason Whitfield, Patricia Walden, Zhenling Cui, Elvira Wijker, Selvakumar Edwardraja, Ignacio Retamal Lantadilla, Fernanda Ely, Claudia Vickers, Jacobus P.J. Ungerer, and Kirill Alexandrov. Generalizable Protein Biosensors Based on Synthetic Switch Modules. *Journal of the American Chemical Society*, 141(20):8128–8135, 2019. ISSN 15205126. doi: 10.1021/jacs.8b12298.
- [30] Saurabh R. Nirantar, Kun Song Yeo, Sharon Chee, David P. Lane, and Farid J. Ghadessy. A generic scaffold for conversion of peptide ligands into homogenous biosensors. *Biosensors and Bioelectronics*, 47:421–428, 2013. ISSN 09565663. doi: 10.1016/j.bios.2013.03.049. URL <http://dx.doi.org/10.1016/j.bios.2013.03.049>.
- [31] Alfredo Quijano-Rubio, Hsien Wei Yeh, Jooyoung Park, Hansol Lee, Robert A. Langan, Scott E. Boyken, Marc J. Lajoie, Longxing Cao, Cameron M. Chow, Marcos C. Miranda, Jimin Wi, Hyo Jeong Hong, Lance Stewart, Byung Ha Oh, and David Baker. De novo design of modular and tunable protein biosensors. *Nature*, 591(7850):482–487, 2021. ISSN 14764687. doi: 10.1038/s41586-021-03258-z. URL <http://dx.doi.org/10.1038/s41586-021-03258-z>.
- [32] Colin Jackson, Alisha Anderson, and Kirill Alexandrov. The present and the future of protein biosensor engineering. *Current Opinion in Structural Biology*, 75(July): 102424, 2022. ISSN 1879033X. doi: 10.1016/j.sbi.2022.102424. URL <https://doi.org/10.1016/j.sbi.2022.102424>.
- [33] Sebastián Bernales, Feroz R. Papa, and Peter Walter. Intracellular signaling by the unfolded protein response. *Annual Review of Cell and Developmental Biology*, 22: 487–508, 2006. ISSN 10810706. doi: 10.1146/annurev.cellbio.21.122303.120200.

Chapter 3. Development of a *S. cerevisiae* strain for optogenetic production of α -amylase and UPR sensing

- [34] Mengmeng Huang, Yanyun Gao, Xiangshan Zhou, Yuanxing Zhang, and Menghao Cai. Regulating unfolded protein response activator HAC1p for production of thermostable raw-starch hydrolyzing alpha-amylase in *Pichia pastoris*. *Bio-process and Biosystems Engineering*, 40(3):341–350, 2017. ISSN 16157605. doi: 10.1007/s00449-016-1701-y.
- [35] Philip I. Merksamer, Ala Trusina, and Feroz R. Papa. Real-Time Redox Measurements during Endoplasmic Reticulum Stress Reveal Interlinked Protein Folding Functions. *Cell*, 135(5):933–947, 2008. ISSN 00928674. doi: 10.1016/j.cell.2008.10.011. URL <http://dx.doi.org/10.1016/j.cell.2008.10.011>.
- [36] Clontech Laboratories. Yeast Protocols Handbook. *Yeast*, 1(July):1–66, 2008. ISSN 13227130. URL <papers://f9646dff-0d84-4514-b2e6-8a168e77928c/Paper/p3846>.
- [37] K. D. Wittrup and V. Benig. Optimization of amino acid supplements for heterologous protein secretion in *Saccharomyces cerevisiae*. *Biotechnology Techniques*, 8(3):161–166, 1994. ISSN 0951208X. doi: 10.1007/BF00161582.
- [38] T Nagai, K Ibata, E Park, M Kubota, K Mikoshiba, and Atsushi Miyawaki. A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. *Genetics*, 20:1585–1588, 1989.
- [39] Michael E. Lee, William C. DeLoache, Bernardo Cervantes, and John E. Dueber. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synthetic Biology*, 4(9):975–986, 2015. ISSN 21615063. doi: 10.1021/sb500366v.
- [40] Daphne S. Bindels, Lindsay Haarbosch, Laura Van Weeren, Marten Postma, Katrin E. Wiese, Marieke Mastop, Sylvain Aumonier, Guillaume Gotthard, Antoine Royant, Mark A. Hink, and Theodorus W.J. Gadella. MScarlet: A bright monomeric red fluorescent protein for cellular imaging. *Nature Methods*, 14(1):53–56, 2016. ISSN 15487105. doi: 10.1038/nmeth.4074.
- [41] Dongmei Flagfeldt, Verena Siewers, Le Huang, and Jens Nielsen. Characterization of chromosomal integration sites for heterologous gene expression in *Saccharomyces cerevisiae*. *Yeast*, 2009.
- [42] R. Daniel Gietz and Robert H. Schiestl. High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nature Protocols*, 2(1):31–34, 2007. ISSN 17542189. doi: 10.1038/nprot.2007.13.
- [43] Marko Lõoke, Kersti Kristjuhan, and Arnold Kristjuhan. Extraction of genomic DNA from yeasts for PCR-based applications. *BioTechniques*, 50(5):325–328, 2011. ISSN 07366205. doi: 10.2144/000113672.

- [44] Diana S.M. Ottoz, Fabian Rudolf, and Joerg Stelling. Inducible, tightly regulated and growth condition-independent transcription factor in *Saccharomyces cerevisiae*. *Nucleic Acids Research*, 42(17), 2014. ISSN 13624962. doi: 10.1093/nar/gku616.
- [45] Jess H. Leber, Sebastián Bernales, and Peter Walter. IRE1-independent gain control of the unfolded protein response. *PLoS Biology*, 2(8), 2004. ISSN 15449173. doi: 10.1371/journal.pbio.0020235.
- [46] Ignacio A Zuleta, Andrés Aranda-díaz, Hao Li, and Hana El-samad. Dynamic characterization of growth and gene expression using high-throughput automated flow cytometry. *Nature Methods*, 11(4), 2014. doi: 10.1038/nmeth.2879.
- [47] Martin C. Jonikas, Sean R. Collins, Vladimir Denic, Eugene Oh, Erin M. Quan, Volker Schmid, Jimena Weibezahn, Blanche Schwappach, Peter Walter, Jonathan S. Weissman, and Maya Schuldiner. Comprehensive Characterization of Genes Required for Protein Folding in the Endoplasmic Reticulum. *Science*, 323(March): 1693–1697, 2009. doi: 10.4324/9781315657455.

4 Mathematical modelling

Mathematical modeling has become an essential tool in many fields of science, including synthetic biology. It allows researchers to distil complex biological systems and observations into a set of defined equations and parameters, providing a framework for understanding the underlying mechanisms and predicting the behavior of these systems [1]. In the context of this project, mathematical modeling can bridge the gap between small-scale characterization experiments and complex fermentation runs.

One of the key advantages of mathematical modeling is that it provides a way to test and refine hypotheses *in silico* before carrying out costly and time-consuming experimental studies. This is particularly important in synthetic biology, where it can be difficult to predict the behavior of complex biological systems from first principles. Mathematical modeling can help identify key parameters and design constraints, guide experimental design, and provide insight into the underlying mechanisms of biological systems. For example, Alvarez-Vasquez et al. [2] used mathematical modelling to predict the metabolic behaviour of *S. cerevisiae* in the ergosterol pathway to drug treatment. Similarly, Dessauges et al. [3] evaluated different model topologies to unravel MAPK network nodes that shape ERK dynamics.

The unfolded protein response (UPR) in particular has been a center for the development of mathematical models, trying to understand the complex structure and feedback channels that are present [4]. Briefly, the main sensor of the UPR in *S. cerevisiae* is inositol-requiring enzyme 1 (Ire1). Upon sensing of unfolded proteins in the endoplasmic reticulum (ER), it splices unspliced Hac1 mRNA, allowing translation to the transcription factor Hac1p. Hac1p regulates the expression of chaperones, glycosylating enzymes and ER associated degradation (ERAD) [5].

Axelsen and Sneppen [6] developed a minimal model of the UPR in yeast, assum-

ing a rapid equilibrium between unfolded proteins and Ire1. Four ordinary differential equations (ODEs) capture the dynamics of unspliced and spliced mRNA, translation of Hac1p and transcription of the chaperone Kar2. They find that the cost of degraded proteins is minimized with a feedback on translation rather than on transcription. Notably, the main focus of modelling is to understand the dynamics of the UPR actors (Ire1, Hac1 and Kar2) and not the dynamics of unfolded proteins, although the authors briefly touch on the unfolded protein dynamics. Trusina et al. [8] build on this minimal modelling framework. They approximate the steps between Ire1 sensing and Hac1 activation (namely dynamics of Hac1 unspliced and spliced mRNA), into one differential equation. In the model for mammalian UPR, they take into account the two additional UPR sensing branches (PERK and ATF6) to propose an explanation of the advantages of translational attenuation specifically for pancreatic β -cells. They find that translational attenuation reduces the amount of unfolded proteins in the ER during stress and decreases the accumulation of excess chaperones between pulses of ER stress. We use this second model as a basis for the description of UPR.

Another area where mathematical modeling is increasingly important is in upscaling. As biological systems move from the laboratory to large-scale production, it becomes more challenging to predict how the system will behave. Mathematical modeling can help address this challenge by providing a way to predict the behavior of the system under different conditions and optimize production processes [9, 10].

Overall, mathematical modeling is an essential tool for advancing synthetic biology and developing new applications in fields such as medicine, biotechnology, and environmental engineering. By distilling complex observations into a set of defined equations and parameters, mathematical modeling provides a way to understand and predict the behavior of complex biological systems. This has the potential to revolutionize our understanding of the natural world and help us develop new and more efficient ways of harnessing its power.

In this chapter, we describe the development of a deterministic mathematical model that allows us to understand the biological processes and reduce the experimental effort to scale up our experiments from initial characterization scale (4 mL) to production scale (in a 1L-bioreactor). Rather than developing these modules from scratch, we decided to use the model structure of two publications as base for the modeling work. The modelling of the upstream optogenetic circuit with activation of the transcription factor Msn2-EL222 and EL222 dependent mRNA expression is captured in the first module with three ODEs (previously described in Benzinger and Khammash [11]). The second module describes the dynamics of the unfolded protein response and is based on the

4.1 Detailed description of the model

minimal model of the UPR by Trusina et al. [7]. The interaction between the basic components of these two modules can be seen in figure 4.1.

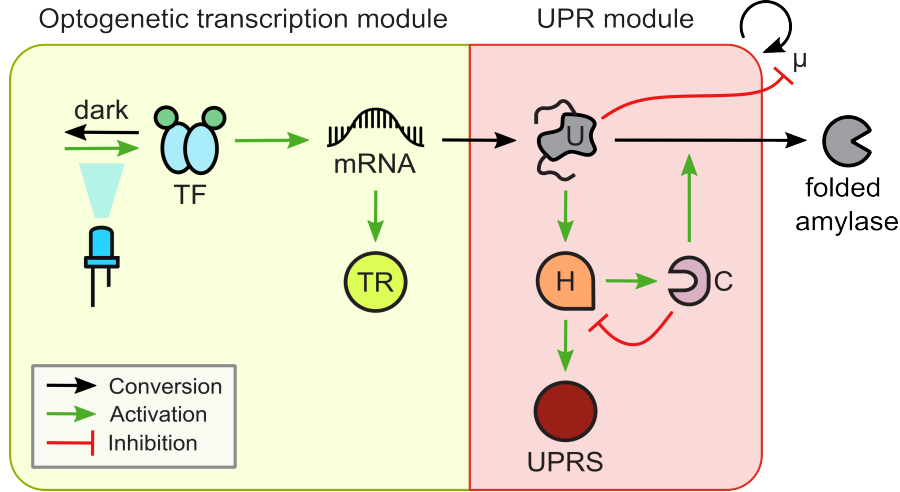


Figure 4.1: Structure of the mathematical model, encompassing the description of the optogenetics [11] (first module, *left*) with activation of the transcription factor TF by light and subsequent transcription to mRNA and translation to TR and unfolded proteins (U). The second module (*right*) describes the UPR dynamics [7] with activation of Hac1p (H) in the presence of U. H triggers the upregulation of chaperones and related enzymes (C) and UPRS, while C reduces the amount of U by folding, secretion or degradation. In section 4.1.2 we describe in more detail the equations governing TR, UPRS and growth rate (μ).

4.1 Detailed description of the model

4.1.1 Description of previous modelling work

The first module consists of three ODEs (previously described in Benzinger and Khamash [11]) modelling the upstream optogenetic circuit with activation of the transcription factor Msn2-EL222 (eq. 4.1) and EL222 dependent mRNA expression (eq. 4.2). Additionally, the translation of TR from mRNA is modelled (eq. 4.12). In more detail, blue light input (I_{eff}) triggers structural changes and homodimerization of the transcription factor Msn2-EL222 (TF) (eq. 4.1). The intensity between 0% and 100% is scaled by k_{on} . We assume that the total amount of transcription factor TF_{tot} and the rate of monomerization k_{off} are both constant. Promoter binding of the transcription factor is fast, so that the rate of mRNA transcription is a direct function of TF, n and K_m with k_m as a scaling factor (eq. 4.2). The degradation rate of mRNA is the basal degradation rate plus dilution from growth (eq. 4.4). After initial fitting the values for $k_{m,basal}$ and n in equation 4.2 were close to 0 resp. 1. To avoid overfitting, they were fixed to 0 and 1, resulting in equation 4.3.

$$\frac{dTF}{dt} = k_{on}I_{eff}(TF_{tot} - TF) - k_{off}TF \quad (4.1)$$

$$\frac{dmRNA}{dt} = k_{m,basal} + k_m \frac{TF^n}{K_m^n + TF^n} - k_{m,deg}mRNA \quad (4.2)$$

$$\frac{dmRNA}{dt} = k_m \frac{TF}{K_m + TF} - k_{m,deg}mRNA \quad (4.3)$$

$$k_{m,deg} = k_{m,deg,b} + \mu \quad (4.4)$$

The second module describes the dynamics of the unfolded protein response and is based on the minimal model of the UPR by Trusina et al. [7]. It consists of mRNA translation to unfolded proteins (U, eq. 4.5a), activation of Hac1p (H) by U (4.6) and the resulting upregulation of chaperones, oxidoreductases, glycosylating enzymes and ER degradation components (all combined in node C, eq. 4.7). C reduces the amount of unfolded proteins in the ER, by either folding, secreting the proteins or degrading it and at the same time inhibits H (eq. 4.5b). The UPRS is transcribed as H binds to the Hac1p specific promoter (P_{4xUPRE} , eq. 4.13).

In more detail, upon presence of U, Ire1 clusters on the ER surface. Due to the proximity of Ire1, they transautophosphorylates each other resulting in an activation of the endoribonuclease function of Ire1 ($Ire1_{act}$, eq. 4.8). This endoribonuclease excises an intron of the unspliced Hac1 mRNA [5]. The translation process of now spliced Hac1 mRNA is then initiated to produce the potent transcription factor Hac1p (H, eq. 4.6). The minimal model reduces the splicing and translation steps of Hac1 mRNA into one ODE (eq. 4.6), assuming that at any time point Hac1p protein concentration is proportional to Hac1 mRNA. The assumption is valid because the half-life of Hac1 protein (~ 1.5 min, [12]) is much shorter than Hac1 mRNA (~ 30 min, [5, 13]). The basal degradation rate of H was assumed to follow mRNA kinetics, because of the short half-life. The model describes the regulation of the UPR with strength parameters α , β , γ and δ (eq. 4.5, 4.6 and 4.7).

$$\frac{dU}{dt} = k_{rib}k_{u,trans}mRNA - \delta \frac{CU}{U + K_{CU}} - \mu U \quad (4.5)$$

$$\frac{dH}{dt} = k_{rib}\beta \frac{IRE1_{act}}{IRE1_{total}} - k_{m,deg}H \quad (4.6)$$

$$\frac{dC}{dt} = k_{rib}(\gamma + \alpha H) - \mu C \quad (4.7)$$

$$\frac{IRE1_{act}}{IRE1_{total}} = \frac{U}{U + K_{IU} + \frac{C}{1+U/K_{CU}} \cdot \frac{K_{IU}}{K_{CI}}} \quad (4.8)$$

4.1.2 Novel modelling performed in this PhD thesis

So far, we have explained the modelling performed by Benzinger and Khammash [11] and Trusina et al. [7]. The main change in the mathematical modelling lies in the source of unfolded proteins for the UPR module in equation 4.5. In the original UPR model [7] the formation of new unfolded proteins U is modelled through a switch-like stress term or a constant translation rate of new polypeptides into the ER. Such a switch could be a drastic environmental change or the addition of a global unfolding agent (e.g. DTT or tunicamycin). With the biological circuit described in chapter 3 we are able to control the creation of new unfolded proteins much more gradually and temporally. Thus, a simple source pulse is not sufficient to capture the circuit behaviour. We connect these two modules by using the time-varying translation rate of mRNA ($k_{u,trans}mRNA$) from our optogenetic module as input to the second module (i.e. the production rate of unfolded proteins (U , eq. 4.5a).

Our experiments are performed in batch cultures. In such cultures, cells are initially growing exponentially, until some limiting substrate S is running out (in our case the main sugar D-glucose). As the substrate S is running out, the growth rate is slowing down, until eventually all substrate is consumed and cells are resting in the stationary phase. In our experiments, cells are growing exponentially for 16-20 hours before experiencing substrate limitations and finally reaching stationarity around 24 hours after inoculation. We thus want to ensure, that these growth stages are accurately captured, as e.g. the dilution rate is heavily impacted by altered growth rates.

Thus, the progression of the cell density X and limiting substrate concentration S is modelled with equations 4.9 and 4.10. Production of biomass X leads to the depletion of the substrate S with a yield factor $Y_{X/S}$. Figure 4.3 shows the progression of substrate concentration and biomass over time.

$$\frac{dX}{dt} = \mu X \quad (4.9)$$

$$\frac{dS}{dt} = -\mu \frac{X}{Y_{X/S}} \quad (4.10)$$

The growth rate μ is described through Monods law [14] with a limiting substrate S. Additionally, We observed the impact of a high transcriptional load on growth rate (figure 4.2). Cells which are illuminated with constant light and light pulses of 2h duration every 6h grew slower than cells in the dark. We thus chose to use an inhibitory factor in the growth law of μ that incorporates the dose-dependent growth reduction caused by the presence of unfolded proteins (eq. 4.11). The hypothesis is that the cell growth rate is primarily affected by metabolic burden [15] and correlates with unfolded proteins U. Equation 4.11 describes the burden to cell growth, where μ_{max} and U_{max} correspond to the maximal cell growth rate and maximal amount of unfolded protein that the optogenetic circuit can induce. Further factors influencing growth rate such as possible toxicity of fluorescent reporters are not considered here.

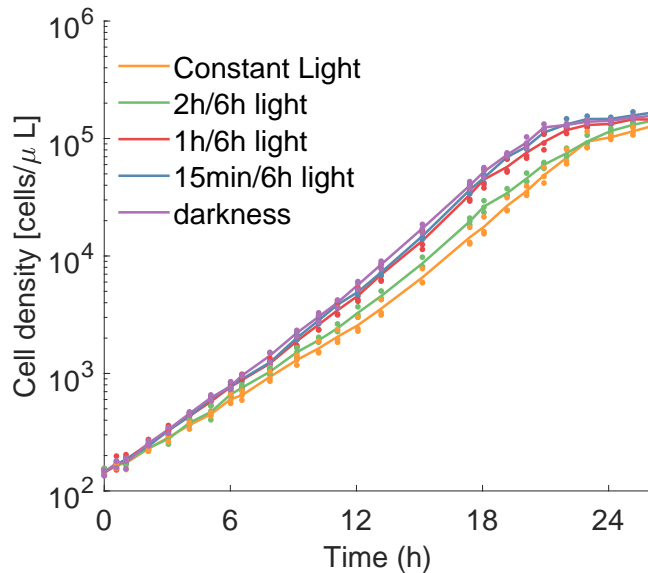


Figure 4.2: Cell density as measured by flow cytometry in cells/ μ L over time for the five conditions shown in figure 3.6. The impact of a high translational load on growth rate can be observed. Cells which are illuminated with constant light and light pulses of 2h duration every 6h grew slower than cells in the dark. (n=3)

$$\mu = \mu_{max} \left(\frac{S}{K_S + S} \right) \left(1 - \frac{U}{U + U_{max}} \right) \quad (4.11)$$

Our biological circuit contains two fluorescent reporters that are modelled for later fitting. The fluorescent transcriptional reporter (TR, eq. 4.12) is translated from mRNA with translation rate $k_{u,trans}k_{rib}$. The contribution of k_{rib} is explained later in equation 4.16. For the UPRS, we assume a linear activation from H to UPRS with strength $k_{UPR,trans}$ and degradation rate $k_{UPRS,deg}$ (eq. 4.13).

$$\frac{dTR}{dt} = k_{rib}k_{u,trans}\text{mRNA} - \mu TR \quad (4.12)$$

$$\frac{dUPRS}{dt} = k_{rib}(k_{UPRS,trans}H + k_{UPRS,stat}) - k_{UPRS,deg}UPRS \quad (4.13)$$

$$k_{UPRS,stat} = k_{UPRS,0} \frac{(S_0 - S)^{n_S}}{K_{d,UPRS} + (S_0 - S)^{n_S}} \quad (4.14)$$

$$k_{UPRS,deg} = k_{UPRS,deg,b} + \mu \quad (4.15)$$

While supplementation of the media with additional inositol reduced the amount of induction of the UPRS (see figure 3.5), still some activation of the UPRS at the end of exponential growth (around 24h) was observed. To capture this induction, a substrate dependant production term of UPRS was introduced ($k_{UPRS,stat}$, eq. 4.14). Figure 4.3 shows that for substrate concentrations close to the initial concentration S_0 resulting in no additional UPRS activation. As substrate S is reaching the hill constant $K_{d,UPRS}$ this stationary induced UPRS activation is triggered. Figure 4.3 also shows when this activation is present in a batch culture. As cells are inoculated at a very low cell density not consuming a lot of substrate, no activation is observed in the first 15 hours. Only as significant biomass starts consuming significant substrate concentrations, can we see the stationary induced induction of the UPRS.

Metzl-Raz et al. [16] report that the ribosomal fraction correlates with the growth rate. While the cells are growing exponentially the fraction of ribosomes is higher, than when cells are going into stationary phase. We see this effect as well, where exponentially growing cells are highly expressing the transcriptional reporter, whereas the amount of TR reduces as cells are in stationarity (see e.g. figure S1). To capture this behaviour all translation rates are multiplied with the factor k_{rib} (eq. 4.16) making the translation rates growth rate dependent.

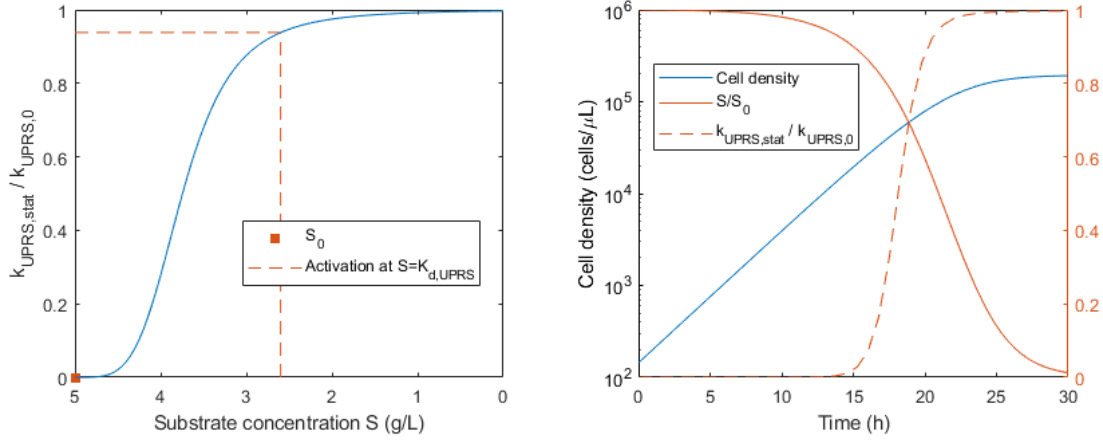


Figure 4.3: **left** Normalized UPRS activation ($\frac{k_{UPRS,stat}}{k_{UPRS,0}}$) as a function of substrate concentration S with parameters from table S5. **right** Simulation of a reduced ODE system modelling the progression of cell density X and substrate concentration S with no growth rate reduction through U .

$$k_{rib} = k_{rib,max} \frac{\mu}{\mu_{max}} \quad (4.16)$$

Lastly, the model can be extended to capture the secretion of amylase. The folding term of the equation describing unfolded protein dynamics (eq. 4.5) can be used as production term for intracellular amylase A_{in} . Assuming a secretion rate k_{sec} and simple dilution of amylase, equation 4.17 can be derived. The concentration of amylase in the supernatant A_{SN} is then a direct function of the secretion rate of the individual cells and the cell density X (eq. 4.18).

$$\frac{dA_{in}}{dt} = \delta \frac{CU}{U + K_{CU}} - \mu A_{in} - k_{sec} A_{in} \quad (4.17)$$

$$\frac{dA_{in}}{dt} = k_{sec} A_{in} X \quad (4.18)$$

Overall, the mathematical model consists of the following eleven states; EL222 transcription factor (TF), mRNA (M), unfolded proteins (U), Hac1p (H), Chaperones (C), transcriptional reporter (TR), UPR sensor (UPRS), cell density (X), substrate concentration (S), intracellular amylase concentration (A_{in}), and amylase concentration in the supernatant (A_{SN}).

4.2 Fitting of the model to the data

For simulating the set of eleven ODEs an initial condition has to be provided. In line with the actual physical experiment, we take as initial condition the steady state value for exponentially growing cells in the dark. Consequently, the light intensity is $I_{eff} = 0$ and the growth rate $\mu = \mu_{max}$. With these conditions all states are zero except for $C_0 = C_{ss} = \frac{\gamma k_{rib}}{\mu}$. Additionally, the initial substrate concentration and initial cell density are parameters from the experiment: $S_0 = 5 \frac{g}{L}$ and $X_0 = 144 \frac{cells}{\mu L}$.

To fit the model to data, experiments for 5 different light conditions were recorded in triplicates in the waterbath. In these experiments, we recorded three different channels; the fluorescence of TR (fig. 3.6a), UPRS (fig. 3.6b), as well as the cell density in Events/ μl (fig. 4.2) with a CytFLEX S. Briefly, cells were plated on an agar plate, a single colony inoculated into a liquid culture and grown overnight, while ensuring that they are always exponentially growing, and not reaching stationarity. Cells from these liquid cultures are inoculated into fresh media and exposed to various light induction profiles, measuring their cellular response. Specifically, cells were exposed to no light, pulses of 15 minutes, 1 hour and 2 hours every 6 hours and to constant light. The fluorescence data was preprocessed, by subtracting the background fluorescence from the data. The mean of the triplicates was used for fitting. As not enough measurements of amylase concentration are available, the parameter k_{sec} is set.

The objective function is the normalized sum of squares of the error between data and model. Specifically, for each of the five conditions (subscript i), the sum of squares of the difference between data (subscript *data*) and model (subscript *sim*) is calculated over all timepoints (subscript j) separately for all three channels (TR, UPRS and cell density X , eqns. 4.19-4.21).

$$SS_{i,TR} = \sum_j (TR_{data,i,j} - TR_{sim,i,j})^2 \quad (4.19)$$

$$SS_{i,UPRS} = \sum_j (UPRS_{data,i,j} - UPRS_{sim,i,j})^2 \quad (4.20)$$

$$SS_{i,X} = \sum_j (X_{data,i,j} - X_{sim,i,j})^2 \quad (4.21)$$

The magnitude of those sums ($SS_{i,TR}$, $SS_{i,UPRS}$, and $SS_{i,X}$) is vastly different and so they are weighted by the square of the mean of the experimental data to give rise to the normalized sum of squares (\overline{SS}), done exemplarily in equation 4.22 for the TR (J is

Chapter 4. Mathematical modelling

the total number of all timepoints in a condition). This ensures that all measurements contribute equally to the fitting objective function. The normalization is also performed within each condition rather than for all. Additionally, a penalty is applied to the fitting of the initial point ($j = 0$).

$$\overline{SS}_{i,TR} = \frac{SS_{i,TR}}{\left(\frac{\sum_{j=1}^J (TR_{data,i,j})}{J}\right)^2} + \frac{(TR_{data,i,0} - TR_{sim,i,0})^2}{TR_{data,i,0}^2} \quad (4.22)$$

Figure 4.4 shows the contribution of the data of each condition and channel to the total sum of squares without (left) and with normalization (right). Without normalization close to 97% of the fitting sum of squares is determined by SS_{TR} for the two conditions "constant light" and "2h of light every 6h". After normalization, the contribution of each channel and condition is distributed more equally. The resulting fit now fits UPRS and cell density well. The error metric SS_{total} is calculated by taking the sum over all conditions and channels (equation 4.23).

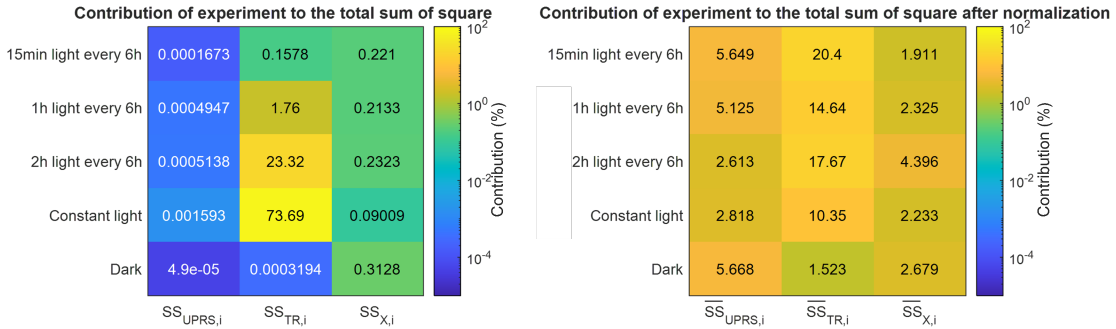


Figure 4.4: Percentage contribution of experimental data to the total sum of squares for fitting either without normalization (left) or with normalization as done in eqn. 4.22.

$$SS_{total} = \sum_i (\overline{SS}_{i,TR} + \overline{SS}_{i,UPRS} + \overline{SS}_{i,X}) \quad (4.23)$$

All simulations and the model fitting were performed using Matlab (R2020b, Mathworks). For optimal computational performance when solving the system of ODEs, we use the built-in variable order ode-solver ode113 for all simulations. To fit the mathematical model to the data we chose an interior-point gradient based algorithm (implemented through fmincon), minimizing SS_{total} and repeated the optimization for a variety of ini-

tial conditions.

Figure 4.5 shows the results of fitting the model to the experimental data with final model parameters listed in table S5. The mathematical model is able to capture the dynamics of UPRS, TR and cell density well. Both data and model show a quicker dark-state reversion in TR levels than in the UPR, highlighting the longevity of the UPR. A discussion about model mismatches can be found in section 4.4.

4.3 Sensitivity analysis

We performed a sensitivity analysis of the fitted parameters to understand what the critical parameters are in the model. To do this we perturbed each parameter by the perturbation h_j and computed the total sum of squares (eq. 4.23, SS_{total}). Figure 4.6 shows the objective function SS_{total} as a function of the perturbation for each parameter. Generally, one can observe, that the parameters are all optimized within numerical accuracy. The only non-smooth behaviour is obtained with γ , which however has a negligible effect on SS_{total} .

Sensitivity measures how the change in one parameter i changes the value returned by the objective function. The sensitivity of each parameter (S_i) is quantified with equation 4.24, where $SS_{i,total}$ is the total sum of squares when the parameter i is perturbed and $SS_{o,total}$ is the optimized (minimized) sum of squares from the fitting results. h_j with $j = 1 \dots M$ is the level of perturbation applied to each parameter. The sensitivity is the average change in SS_{total} of M perturbations. Parameter sensitivities are shown in the title of each panel in figure 4.6.

$$S_i = \frac{1}{M} \sum_j^M \left| \frac{1}{h_j} \right| \cdot \left| \frac{SS_{i,total}}{SS_{o,total}} - 1 \right| \quad (4.24)$$

If we consider parameters with sensitivities >2 as sensitive parameters, 7 parameters are classified as sensitive, namely k_{on} , k_{off} , $k_{m,deg,b}$, K_m , μ_{max} , $k_{rib,max}$, and K_S . These parameters can be grouped in two classes, one being the immediate optogenetic activation of transcription (k_{on} , k_{off} , $k_{m,deg,b}$ and K_m) and the other class holding parameters related to the growth kinetics (μ_{max} , $k_{rib,max}$, and K_S). Surprisingly, none of the parameters related to the actual unfolded protein response are present in the set of sensitive parameters.

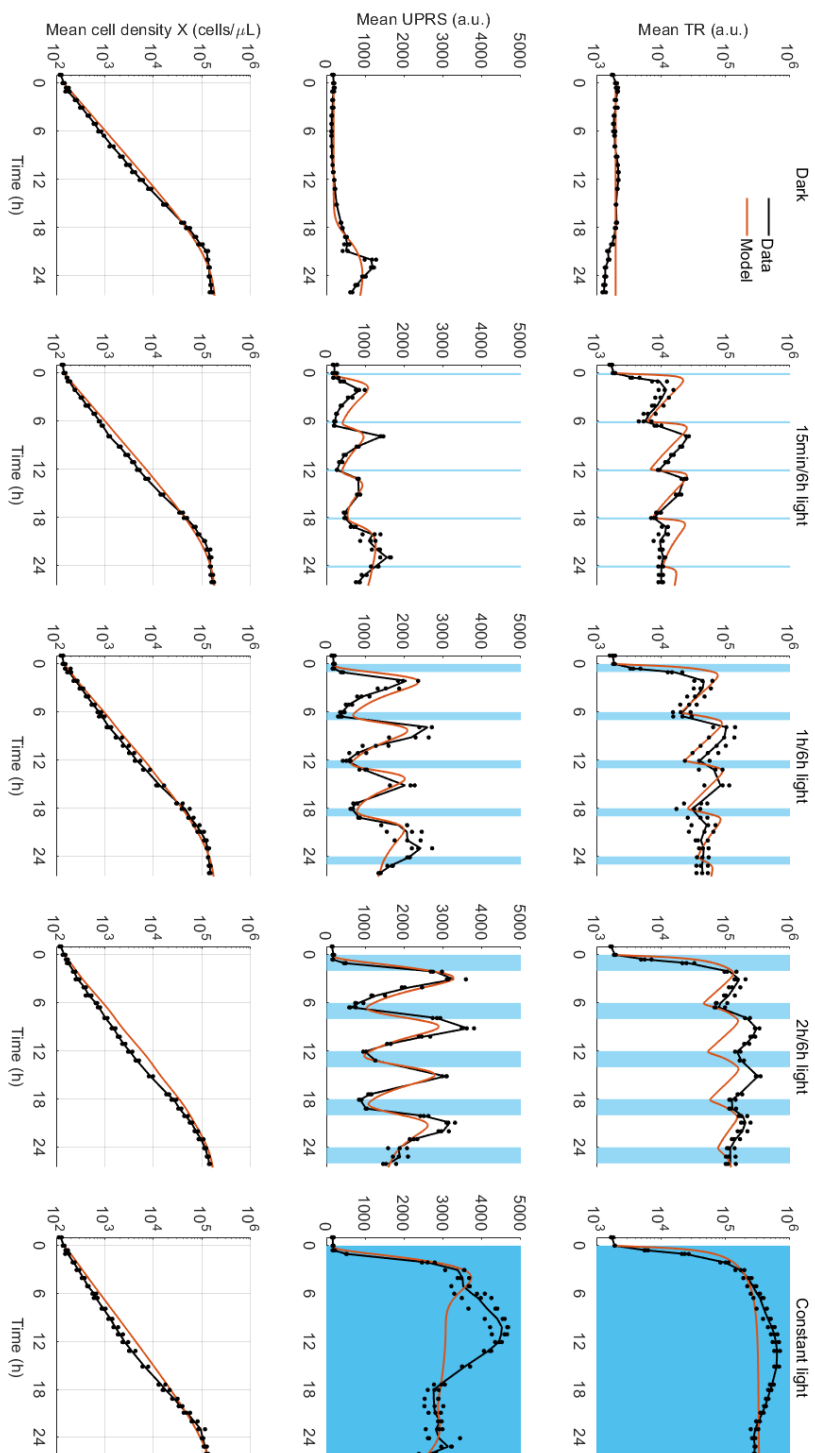


Figure 4.5: Experimental data (single replicate with dots, mean data in black solid line) plotted with mathematical model fit (red) for transcriptional reporter (TR, first row), UPR sensor (UPRS, second row) and cell density (X, third row). Blue shaded areas indicate blue light illumination.

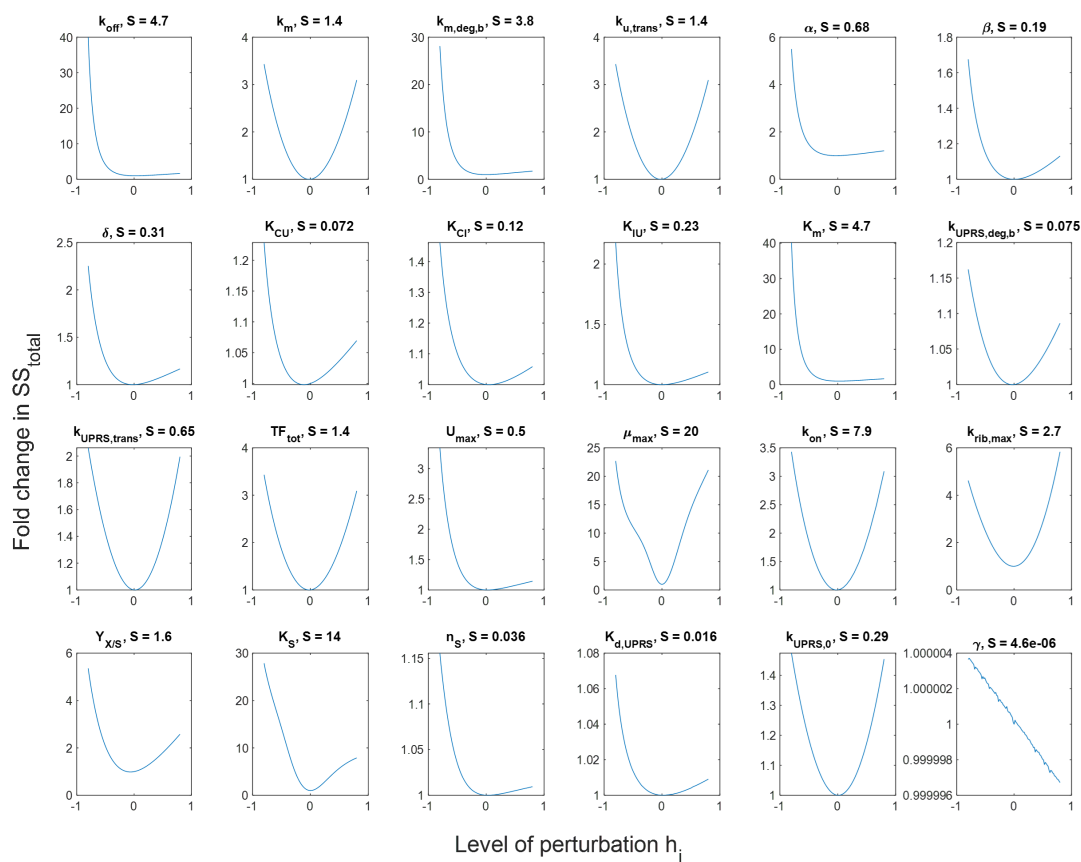


Figure 4.6: Fold change in the total sum of squares (SS_{total}) when single parameter is perturbed (from -80% to +80%). The sensitivity of each individual parameter is shown in the title. The majority of the parameters are optimized (at 0 perturbation). The only parameter with a non-smooth curve and where the fitting has not converged, γ has a negligible effect on SS_{total} .

4.4 Model inaccuracies

This section goes into further details about the model inaccuracies, where the model fit is unable to capture the dynamics of the data. These discussions should highlight potential improvements leading to a deeper understanding of the circuit. At the same time, the final usage of this mathematical model as a quick predictor for control parameters for closed-loop control (see section 4.5) is still fulfilled.

4.4.1 Jump behaviour after induction

Highly dynamic induction of the circuit can be observed when cells are exposed to a new optogenetic input. Figure 4.7 shows the behaviour of transcriptional reporter and UPR sensor of cells that are growing exponentially in the dark and are then exposed to blue

light at time $t=0h$ for 60 minutes. The observed data shows a delay of more than 30 minutes before a significant increase in the reporter levels is observed. The mathematical model predicts a much quicker response in both reporters. It is noteworthy, that this quick response in the states is especially present in the model fit of the TR. The fitting algorithm is compromising between accurately capturing the positive induction time dynamics and negative degradation and dilution time dynamics.

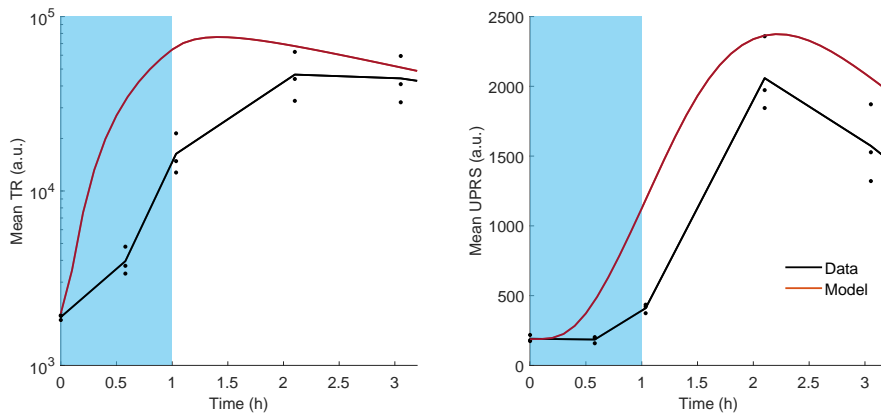


Figure 4.7: Progression of TR (left) and UPRS (right) as a function of time as measured (Data) and fitted by the mathematical model (Model).

In order to understand how the model predicts reporter expression, we looked at the hidden states of the mathematical model. Figure 4.8 shows the predicted response of transcription factor, mRNA, unfolded proteins and the fluorescent reporters. In general, the further away from the optogenetic input, the later the response of the state, i.e. activation of TF, then mRNA, followed by U and TR and finally activation of UPRS. In order to better capture the jump behaviour of the circuit, one approach could be to introduce more intermediate states that “delay” the activation of a reporter (e.g. applied in Korsbo and Jönsson [17]). These states could be TF-binding to the promoter, polymerase binding to DNA, Hac1 mRNA transcription and splicing [6] or the two-step process of transcribing and translating the UPRS. Alternatively, one could work with delay differential equations to capture the multiple intermediate states, as for example applied in Glass et al. [18].

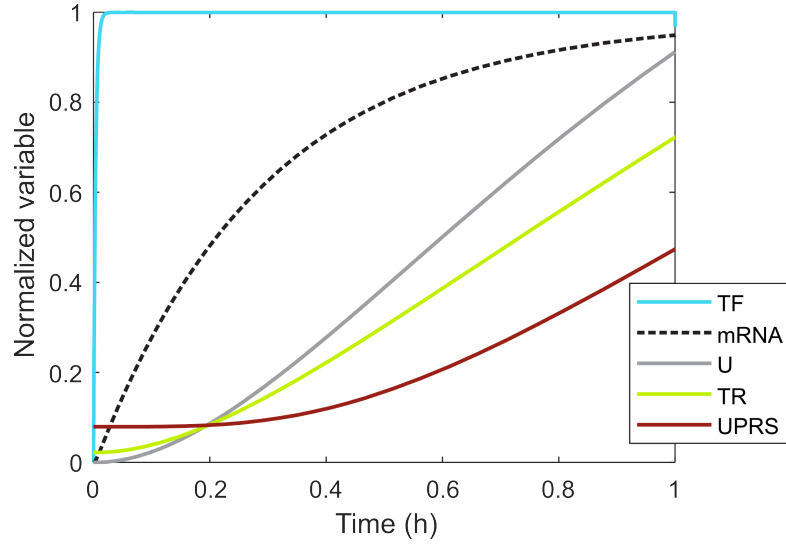


Figure 4.8: (top) Model prediction of the progression of transcription factor (TF), mRNA, unfolded proteins (U), transcriptional reporter (TR) and UPR sensor (UPRS) to a pulse of light. The states are normalized to the maximum that is experienced during the jump, to better compare the time dynamics.

4.4.2 Model inaccuracies in capturing the growth dynamics

When observing the cell density measurements in comparison to the fitted mathematical model (figure 4.5), one can observe that a model mismatch is present in the time window around 12 hours. The simulations yield higher cell densities than what is actually measured. In order to analyze this effect more clearly, we calculated the instantaneous growth rate μ_{inst} from the cell density data X of each replicate between two measurement points X_t and X_{t+1} (see eq. 4.25, [19]). Because the data is very noisy, we additionally, smoothed over μ_{inst} using a moving average filter with $n=7$.

$$\mu_{inst} = \frac{\ln \frac{X_{t+1}}{X_t}}{\Delta t} \quad (4.25)$$

We can observe in figure 4.9 that the instantaneous growth rate is increasing over time, reaching a maximum before the cells go into stationarity and μ_{inst} tends to zero. In the first two hours, the measured growth rate is increasing from $\approx 0.15h^{-1}$ to an initial steady state between $0.25 - 0.35h^{-1}$. This lag-phase effect is usually observed, as cells are transferred to fresh media and require some adjustment time for sensing the new environment. The mathematical model does not incorporate this effect (see eq. 4.11).

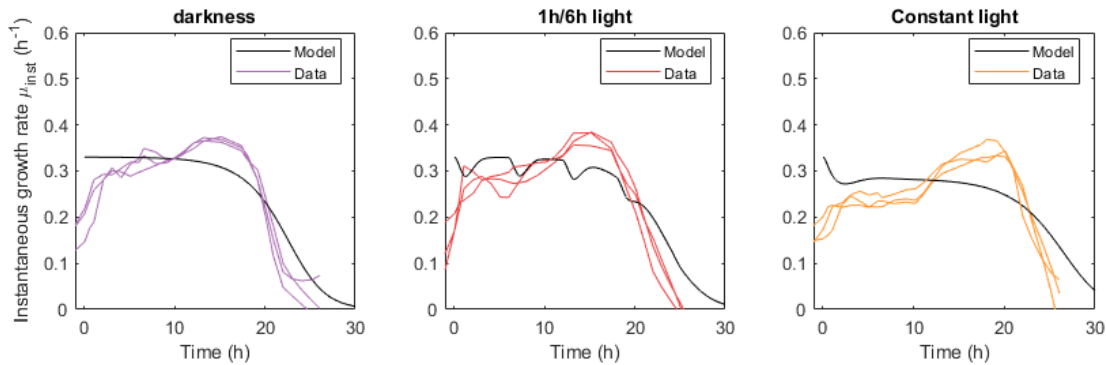


Figure 4.9: The instantaneous growth rate μ_{inst} plotted as a function of time (shown are individual replicates, $n=3$). Shown is the model fit as well as the data.

We can further observe a slow increase of the growth rate after the lag phase between 3 and 18 hours after inoculation. This is an initially surprising result, as we expect cells to adapt rather quickly to growth conditions and grow exponentially with a constant maximum growth rate μ_{max} . We developed several hypotheses that could explain this increase in the growth rate over time.

- Metabolic adaptation

The growth media is composed of different sources of growth, including glucose, a variety of amino acids and the complex bovine serum albumin. Because of the multitude of potential carbon sources, the cells may be growing on different carbon sources throughout the experiment, leading to different growth rates. Alternatively, the cells may be producing a different intermediate metabolite over time, which enhances the growth rate as well.

- Morphology differences

We observe, that the cell size (as measured by the forward-scatter of the flow cytometer) is increasing up to 20% in the first 12 hours, before decreasing as time progresses. Possibly, cells are initially favoring cell size growth, before preferentially favoring cell division.

- Genomic mutation resulting in growth advantages

Another possible explanation is the appearance of a mutation in one cell, which has a growth advantage, and out-competing the mother population. This would result in an increased apparent population growth rate, as the mutated strain is taking over the culture. We can exclude this possibility, as the onset and increase follows very similar behaviour over all replicates and conditions.

A way to further understand this observation, would be to perform metabolic studies on cells sampled at different time points. This would allow us to more clearly delineate where the growth rate increase is coming from. From a modeling perspective, we could capture this behaviour phenomenologically, which would surely improve the fit. For our application however, the fit is satisfactory, capturing the main driving force of growth rate reduction as a substrate is used up ($t > 20h$).

4.4.3 Constant light inaccuracies

The mathematical model is able to capture the general increase and decrease of TR and UPRS for most conditions shown in figure 4.5. The constant light case is an exception where both TR and UPRS response are underestimated and the mathematical model does not follow the trends that are observed in the experimental data. The fitted model shows a saturation of the level of TR with constant light where in reality, TR accumulates in the cells until ribosomal activity decreases (see equation 4.16), resulting in a drop of TR (after 14 hours). The model also predicts a UPRS overshoot and subsequent return to a high homeostatic state. In reality, the cells remain on a highly-stressed state.

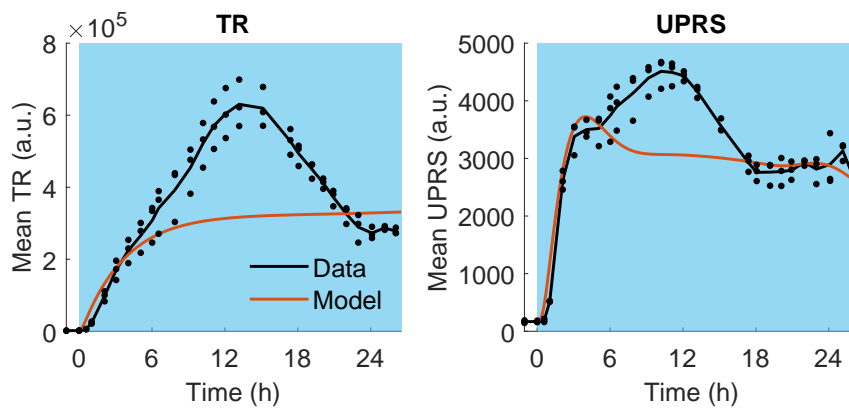


Figure 4.10: Experimental data (single replicate with dots, mean data in black solid line) plotted with mathematical model fit (red) for transcriptional reporter (TR, left) and UPR sensor (UPRS, right). Blue shaded areas indicate blue light illumination.

Two possible explanations can explain the model mismatch in UPRS.

- The underestimation of the transcriptional reporter expression (TR) by the mathematical model could directly explain the lower levels of UPRS.
- Secondly, the UPR model from Trusina et al. [7], which we have adapted here,

excludes unmitigated stress and the onset of a super-UPR as previously reported in Bernales et al. [5]. A constant high UPR can result in a second signalling level termed super-UPR, which triggers a new set of genomic targets, such as the endoplasmic associated degradation (ERAD) pathway and further chaperones. Including these driving forces in the model could further improve the model fit and explain the experimental observation.

4.5 Prediction of control parameters

We can leverage the fitted mathematical model to prototype controllers and predict parameter regimes that will result in robust behaviour.

4.5.1 Choice of controller input

Importantly, before running any controller simulations, we need to ensure that the automatic sampling implemented in chapter 2 can sample fast enough to capture the dynamics of the UPRS without aliasing. Aliasing is an artifact in signal processing, where the recorded or reconstructed signal does not accurately represent the original, continuous signal. It occurs when the sampling rate of a continuous signal is too low relative to the highest frequency component in the signal, causing high-frequency components to be represented as lower frequency components. This results in an "aliased" signal that has a different frequency content than the original.

To avoid this, we fitted a linear first order plus delay model to our experimental data and calculated the Nyquist frequency f_{Nyq} to be $0.6 h^{-1}$. The minimum sampling rate should be at least twice as large as the Nyquist frequency. We had initially run experiments with a sampling rate of $0.667 h^{-1}$, which is lower than $2 \cdot f_{Nyq}$. After this calculation we increased the sampling frequency to $2 h^{-1}$ (=sampling interval of 30min). This is also in agreement with experimental findings. Bennett et al. [20] performed an experiment where they were exposing *S. cerevisiae* strains to periodic changes in the carbon source and tracked the metabolic response of the cells. They find, that cells are acting as a low-pass filter with a maximum response frequency of $0.88 h^{-1}$. This suggests that a sampling rate of $2 h^{-1}$ is sufficient for yeast.

Due to the non-linearity of light intensity to gene expression, we chose to modulate the light input by altering the duty cycle at constant light intensity [11]. Given the analysis of the sampling frequencies in the previous paragraph, we varied the duty cycle within a 30-minute period. For a duty cycle of e.g. 20%, light is turned on for 6 minutes

after measurement and turned off for the remaining 24 minutes.

4.5.2 Choice of controller for initial proof of concept

In this initial proof of concept study, we want to control the expression of the unfolded protein sensor UPRS to a constant. We chose to use a simple proportional-integral controller architecture. Initially, preliminary experiments were yielding quite noisy measurement data. As the derivative portion in a possible PID controller is known to be sensitive to measurement noise, we opted for the simpler PI controller. For future applications, the addition of the derivative term could lead to better control.

4.5.3 Method to predict PI parameters

In order to predict the performance of control parameters, we simulate the system of equations for 30 minutes ($= \Delta t$) resulting in a predicted measurement of UPRS levels (y_{pred}). The difference between target UPRS (y_{UPRS}^*) and the sum of predicted measurement of UPRS (y_{pred}) and background fluorescence ($y_{bg,UPRS}$) is the new error e_n (eq. 4.26). The controller with control parameters K_P and K_I is then calculating the next light input u_{n+1} from the error terms $e_0, e_1, \dots, e_{n-1}, e_n$ according to equation 4.27. The next light input u_{n+1} is bounded to stay between zero and the sampling time (30 minutes).

$$e_n = y_{UPRS}^* - (y_{pred} + y_{bg,UPRS}) \quad (4.26)$$

$$u_{n+1} = K_p e_n + \sum_{j=0}^n K_i e_j \Delta t \quad (4.27)$$

Figure 4.11 shows an example of using the mathematical model to predict the performance of a PI control parameter set. Rather than manually tuning the control parameters, we ran simulations for a set of 50 K_p and 50 K_i parameters. To quantify the performance of each parameter combination, we calculate the area between prediction and target UPRS by forming the Riemann integral.

Figure 4.12 shows the mean error in the first 24 hours between UPRS setpoint and UPRS as predicted by the model on a grid of 2'500 K_P and K_I parameter combinations. This PI parameter tuning is repeated for 3 setpoints of the UPRS. Only a limited set of parameter combinations is predicted to achieve tracking of the setpoint (blue shaded

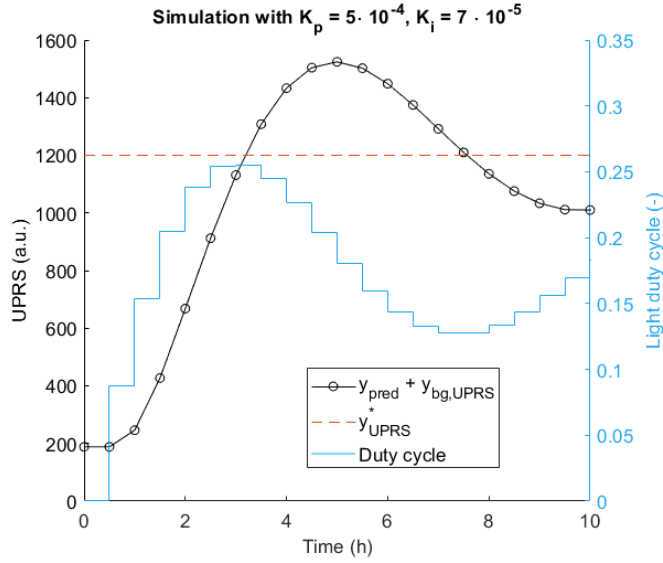


Figure 4.11: Model prediction of the UPRS as a function of time with a proportional-integral control with parameters $K_p = 5 \cdot 10^{-4}$ and $K_i = 7 \cdot 10^{-4}$.

areas). One can observe that the higher the setpoint, the steeper the resulting surface response and thus the smaller the operating range.

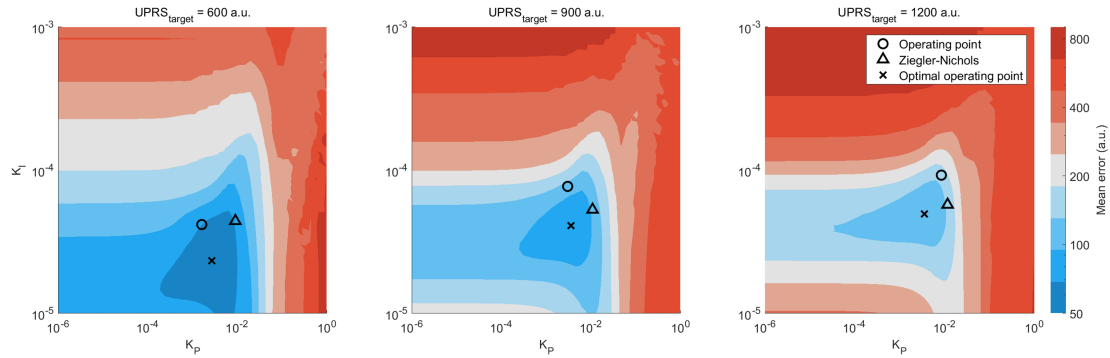


Figure 4.12: Surface response of the mean error in a.u. to varying values of proportional (K_P) and integral (K_I) controller gain for three different target values of the UPRS. Shown is additionally, the operating point, predicted optimal point, as well as the parameter with Ziegler-Nichols tuning.

The initial parameter choice was based off preliminary experiments, data and model fits. The experiments that are described in later chapters are performed with these parameter sets (“Operating point” in figure 4.12). Since then, new experimental data and further characterization refined the model yielding the heatmaps shown in the figure. Still, the chosen operating points are in parameter regions which are predicted to yield adaptation.

4.5.4 Comparison of different controller architectures

We also calculated the PI parameter set using the Ziegler-Nichols heuristic method. Following general practice, the control loop is closed using a proportional gain, increasing the proportional gain until at $K_p = K_{crit}$ sustained oscillations with period (T_{crit}) are obtained. The Ziegler-Nichols PI parameters ($K_{p,ZN}$ and $K_{i,ZN}$) follow directly from these constants (eqns. 4.28-4.29). The derived model never achieves sustained oscillations, as cells are going into stationarity, limiting the expression of new proteins and fluorescent reporters. Thus, we adapted our model, by setting the initial cell density X_0 to zero, effectively allowing cells to grow exponentially indefinitely. After these adjustments, we can see that the predicted parameters fall into a similarly robust regime on the heatmaps in figure 4.12, highlighting the effectiveness of these parameter tuning methods. All control parameters are shown in table S6.

$$K_{p,ZN} = 0.45 \cdot K_{crit} \quad (4.28)$$

$$K_{i,ZN} = \frac{K_{p,ZN}}{0.85 \cdot T_{crit}} \quad (4.29)$$

4.6 Conclusion

In this chapter, we have developed a deterministic mathematical model that captures the response of engineered *S. cerevisiae* to pulses of light. The model describes the dynamics of components of the UPR, cell density, as well as the progression of reporter dynamics in stationarity. More states describing the dynamics (e.g. Hac1 mRNA transcription and splicing, [6]) and a better understanding of the growth dynamics could further improve the model fit. We used the model to predict the performance of proportional-integral control parameters. In chapter 5, we demonstrate, that we indeed achieve robust control with these parameters, highlighting the versatility of the model. More advanced control schemes such as proportional-integral-derivative (PID), model predictive control (MPC) or lead-lag controllers could be tuned very quickly with this model and generate even more robust transient behaviour.

The model part for amylase secretion is in its infancy and not fitted to experimental data. Thus, only limited information can be gained from this part. In the future, with the availability of more amylase measurements, this part of the mathematical model should be characterized in more detail. Once present, this would then further allow to predict more ideal production schemes.

Bibliography

- [1] Yuting Zheng and Ganesh Sriram. Mathematical modeling: Bridging the gap between concept and realization in synthetic biology. *Journal of Biomedicine and Biotechnology*, 2010(Figure 1), 2010. ISSN 11107243. doi: 10.1155/2010/541609.
- [2] Fernando Alvarez-Vasquez, Howard Riezman, Yusuf A. Hannun, and Eberhard O. Voit. Mathematical modeling and validation of the ergosterol pathway in *saccharomyces cerevisiae*. *PLoS ONE*, 6(12), 2011. ISSN 19326203. doi: 10.1371/journal.pone.0028344.
- [3] Coralie Dessauges, Jan Mikelson, Maciej Dobrzyński, Marc-Antoine Jacques, Agne Frismantiene, Paolo Armando Gagliardi, Mustafa Khammash, and Olivier Pertz. Optogenetic actuator – ERK biosensor circuits identify MAPK network nodes that shape ERK dynamics. *Molecular Systems Biology*, 18(6):1–22, 2022. ISSN 1744-4292. doi: 10.15252/msb.202110670.
- [4] Ilaria Pontisso, Roberto Ornelas Guevara, Laurent Combettes, and Genevieve Dupont. A journey in UPR modeling. *Biology of the Cell*, 2023. doi: 10.1111/boc.202200111.
- [5] Sebastián Bernales, Feroz R. Papa, and Peter Walter. Intracellular signaling by the unfolded protein response. *Annual Review of Cell and Developmental Biology*, 22: 487–508, 2006. ISSN 10810706. doi: 10.1146/annurev.cellbio.21.122303.120200.
- [6] Jacob Bock Axelsen and Kim Sneppen. Quantifying the benefits of translation regulation in the unfolded protein response. *Physical Biology*, 1(3):159–165, 2004. ISSN 14783975. doi: 10.1088/1478-3967/1/3/003.
- [7] Ala Trusina, Feroz R. Papa, and Chao Tang. Rationalizing translation attenuation in the network architecture of the unfolded protein response. *PNAS*, 105(51):20280–20285, 2008. ISSN 00278424. doi: 10.1073/pnas.0803476105.
- [8] A. Trusina and C. Tang. The unfolded protein response and translation attenuation: a modelling approach. *Diabetes, Obesity and Metabolism*, 12(SUPPL. 2):27–31, 2010. ISSN 14631326. doi: 10.1111/j.1463-1326.2010.01274.x.
- [9] Keith EJ Tyo, Kanokarn Kocharin, and Jens Nielsen. Toward design-based engineering of industrial microbes. *Current Opinion in Microbiology*, 13(3):255–262, 2010. ISSN 13695274. doi: 10.1016/j.mib.2010.02.001. URL <http://dx.doi.org/10.1016/j.mib.2010.02.001>.

-
- [10] Joachim Almquist, Marija Cvijovic, Vassily Hatzimanikatis, Jens Nielsen, and Mats Jirstrand. Kinetic models in industrial biotechnology - Improving cell factory performance. *Metabolic Engineering*, 24:38–60, 2014. ISSN 10967184. doi: 10.1016/j.ymben.2014.03.007. URL <http://dx.doi.org/10.1016/j.ymben.2014.03.007>.
- [11] Dirk Benzinger and Mustafa Khammash. Pulsatile inputs achieve tunable attenuation of gene expression variability and graded multi-gene regulation. *Nature Communications*, 9(2018):1–38, 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05882-2. URL <http://dx.doi.org/10.1038/s41467-018-05882-2>.
- [12] Bhupinder Pal, Nickie C. Chan, Leon Helfenbaum, Kaeling Tan, William P. Tansey, and Mary-Jane Gething. SCFCdc4-mediated Degradation of the Hac1p Transcription Factor Regulates the Unfolded Protein Response in *Saccharomyces cerevisiae*. *Molecular biology of the cell*, 18:426–440, 2007. ISSN 1059-1524. doi: 10.1091/mbc.E06.
- [13] Ursula Rügsegger, Jess H Leber, Peter Walter, and San Francisco. Block of HAC1 mRNA Translation by Long-Range Base Pairing Is Released by Cytoplasmic Splicing upon Induction of the Unfolded Protein Response. *Cell*, 107:103–114, 2001.
- [14] Jacques Monod. The Growth of Bacterial Cultures. *Annual Reviews in M*, 3(XI): 371–394, 1949.
- [15] Ashty S. Karim, Kathleen A. Curran, and Hal S. Alper. Characterization of plasmid burden and copy number in *Saccharomyces cerevisiae* for optimization of metabolic engineering applications. *FEMS Yeast Research*, 13(1):107–116, 2013. ISSN 15671356. doi: 10.1111/1567-1364.12016.
- [16] Eyal Metzl-Raz, Moshe Kafri, Gilad Yaakov, Ilya Soifer, Yonat Gurvich, and Naama Barkai. Principles of cellular resource allocation revealed by condition-dependent proteome profiling. *eLife*, 6:1–21, 2017. ISSN 2050084X. doi: 10.7554/eLife.28034.
- [17] Niklas Korsbo and Henrik Jönsson. It’s about time: Analysing simplifying assumptions for modelling multi-step pathways in systems biology. *PLoS Computational Biology*, 16(6):1–29, 2020. ISSN 15537358. doi: 10.1371/journal.pcbi.1007982.
- [18] David S. Glass, Xiaofan Jin, and Ingmar H. Riedel-Kruse. Nonlinear delay differential equations and their application to modeling biological network motifs. *Nature Communications*, 12(1), 2021. ISSN 20411723. doi: 10.1038/s41467-021-21700-8. URL <http://dx.doi.org/10.1038/s41467-021-21700-8>.
- [19] Stefan Bruder, Mara Reifenrath, Thomas Thomik, Eckhard Boles, and Konrad Herzog. Parallelised online biomass monitoring in shake flasks enables efficient

Chapter 4. Mathematical modelling

strain and carbon source dependent growth characterisation of *Saccharomyces cerevisiae*. *Microbial Cell Factories*, 15(1):1–14, 2016. ISSN 14752859. doi: 10.1186/s12934-016-0526-3.

- [20] Matthew R Bennett, Wyming Lee Pang, Natalie A Ostroff, L Bridget, Sujata Nayak, Lev S Tsimring, and Jeff Hasty. Metabolic gene regulation in a dynamically changing environment. *Nature*, 454(7208):1119–1122, 2008. doi: 10.1038/nature07211.Metabolic.

5 Closed-loop bioreactor experiments

The goal of this PhD thesis is to study how we could use closed-loop feedback control to improve heterologous protein expression. Poor heterologous protein expression can stem from several sources including but not limited to transcription, translation, folding, post-translational modification and secretion [1]. Intuitively, increasing the rate of gene expression should result in more secreted protein. While this is true at low protein expression rates, maximal expression does not necessarily lead to maximal protein production, but can even reduce protein yield [2, 3]. Consequently, there exists an optimal gene expression rate that maximizes production. The diminishing rate of protein production can be attributed to an overwhelmed cell machinery, resulting in oxidative stress, product misfolding, inclusion bodies, upregulated endoplasmic reticulum-associated degradation, and stress-induced genomic instability [4–7].

The mapping between optimal gene expression rate and optimal protein production is influenced by many different factors, such as growth rate, process stage, and complexity of the protein of interest [8]. Thus, a constant, fine-tuned gene expression level [9] is generally only optimal during a limited process window and for a single product. One possible solution is to adjust protein expression automatically based on stress levels using burden-driven, genetic feedback circuits [10]. However, for this approach to function optimally, the synthetic circuits themselves need to be fine-tuned to a given product and production condition. Additionally, hard-wired feedback cannot adapt when process requirements change over time.

These challenges can be solved by applying direct *in silico* control on cell internal states similar to what has been established in the industry for process parameters like pH and dissolved oxygen. This allows for the adjustment of feedback strength to the design criteria, ultimately improving protein production (figure 5.1).

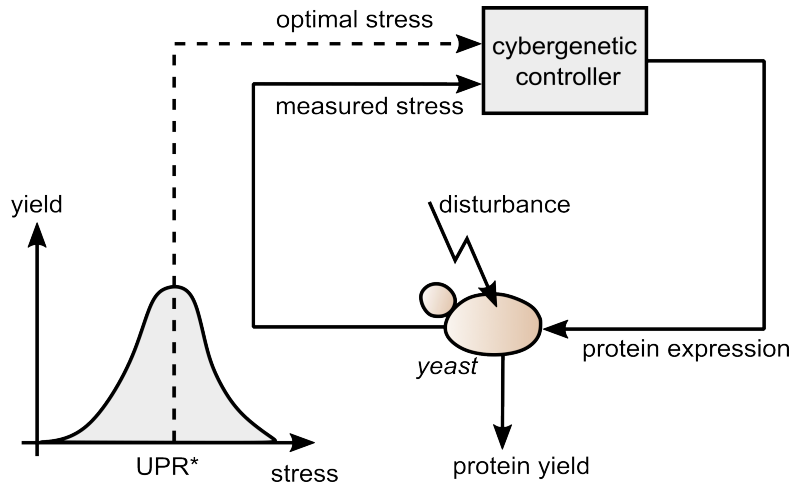


Figure 5.1: Cybergenetic control of the UPR optimizes protein yield. We postulate that an optimal UPR stress (UPR^*) exists which maximizes protein yield. A cybergenetic controller takes this optimal stress input and the measured UPR stress and computes the optimal next protein expression strength. The engineered *S. cerevisiae* produce protein given the protein expression strength and its growth parameters (disturbance).

Bulk feedback of yeast liquid cultures was first demonstrated by Miliias-Argeitis et al. [11]. The authors used the light-responsive Phy/PIF red/far-red optogenetic system to achieve robust regulation of gene expression fold change. The feedback was implemented by manual sampling of the cultures. Melendez et al. [12] improved this setup by removing the need for manual sampling, with automatic fluorescence imaging attached to a microfluidic device. With this setup, tight control over gene expression was achieved for 2 days. Miliias-Argeitis et al. 2016 used the green/red CcaS/CcaR system in *E. coli* to track the GFP level at different set points leveraging model predictive control to achieve the target levels. They additionally, controlled *MetE* a key metabolic enzyme for growth in *E. coli* and were able to control the growth rate with this setup. Building on this idea, Harrigan et al. [14] studied internal feedback structures in yeast, by deleting feedback components in the pheromone response pathway and re-introducing this feedback with closed loop optogenetic control. With this platform, they were able to identify not only which nodes are involved in biological signaling pathways, but additionally what dynamic requirements are computed in each biological node. Steel et al. more recently developed a cheap custom-built 25mL optogenetic device and demonstrated closed-loop feedback control of GFP in *E. coli*. A different approach was used by Aditya et al. [16] to maintain a coculture system. The Cre recombinase was placed under optogenetic control, so that upon light induction the strain would irreversibly switch from mCrim production to mNeon production. Operating in a turbidostat and constantly re-introducing the mCrim producer, a coculture at different ratios of mCrim/mNeon producer, could be

maintained for up to 100 hours. These papers have in common that they use reversible light inputs, and are focused on development of optogenetic regulators and regulation strategies in general. However, these proof of principles are operated at small volumes (<20mL) and neglect the challenges related to bioprocessing such as aeration, high cell densities and upscaling. Lastly, none of the above described papers focus on the optimization of the production of a protein or metabolite of interest. The idea of using closed-loop optogenetic feedback control for optimal bioproduction has been extensively discussed in the community [17, 18] and this work should fill this gap.

In this chapter, we optimize the production of α -amylase by using *in silico* feedback control of the unfolded protein response (UPR). We have chosen the UPR as target control as it is central to the proper folding of proteins and is often engineered to enhance protein production [19, 20] hinting towards an ideal UPR induction level. The UPR also integrates over disturbances stemming from growth kinetics, nutrient availability and protein folding state. Chapter 3 described the engineering of *S. cerevisiae* to optogenetically express α -amylase, a hard-to-fold model protein [21, 22], as well as a fluorescent sensor of UPR activity. In chapter 2, we describe the development of a cybergenetic bioreactor platform that automates sampling, dilution, measurement, and dynamic real-time feedback of our culture. Lastly, we describe the development of a mathematical model in chapter 4, which predicts control parameters for closed-loop feedback control. This chapter brings together the work of the previous three and demonstrates closed-loop feedback of the UPR in exponentially growing *S. cerevisiae*. Ultimately, we study how clamping the mean population stress level to different constant levels of UPR affects the final yield of the model protein α -amylase. We show that intermediate unfolded protein response levels achieved by cybergenetics result in 60% increased protein production in comparison to maximal UPR levels.

5.1 Overview and mode of operation

Figure 5.2 shows the full photobioreactor operated in closed-loop. Cells are automatically sampled and diluted as shown in figure 2.14. The sample is delivered to a flow cytometer (CytoFLEX S), where an API allows the automatic measurement, and export of single-cell fluorescence data. Here, we measure the level of unfolded protein response sensor (UPRS) as produced by the individual cells. The mean of the single-cell UPRS measurements is taken and fed into the controller. As discussed in chapter 4, the controller is a proportional-integral controller (PI). It computes the next duty cycle light input which is implemented by a microcontroller and activates the blue light LEDs in the photobioreactor LED housing. Samples for α -amylase measurements are collected

Chapter 5. Closed-loop bioreactor experiments

automatically every 6 hours in the first 24 hours and then once every 12 hours. Additionally, an end-point manual sample for amylase quantification is drawn at the end of the run.

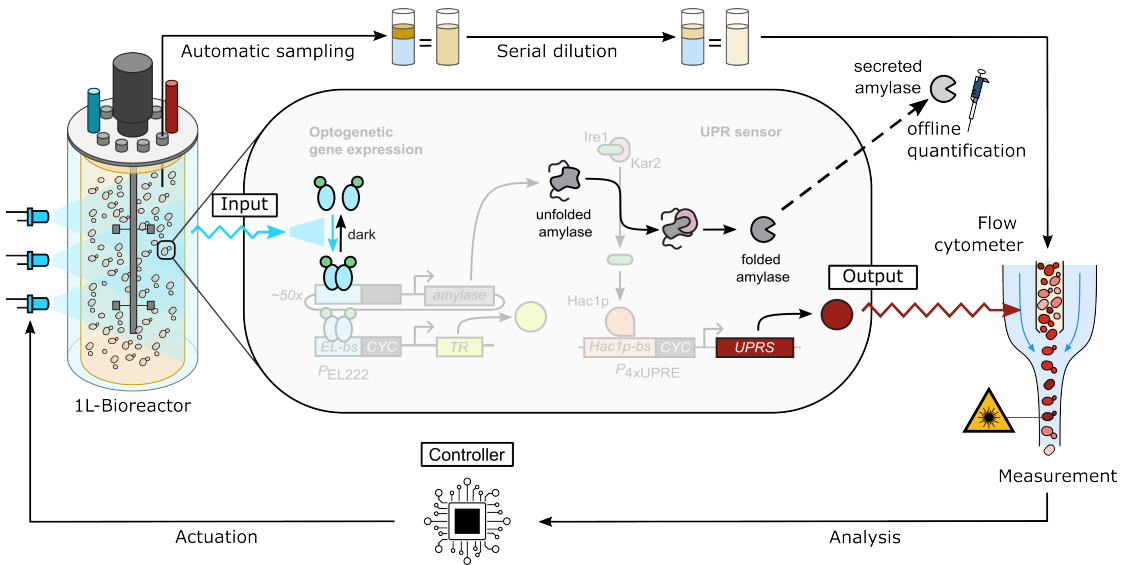


Figure 5.2: Closed-loop optogenetic 1L-bioreactor platform and engineered *S. cerevisiae*. The bioreactor is placed in a light house to allow illumination of the culture broth. The culture broth is automatically sampled by peristaltic pumps at regular intervals, diluted in two stages with diluent and finally transferred to a flow cytometer for measurement of the cell density and UPRs. The measurements of the UPRs are automatically analyzed, and the PI controller calculates the next light input, closing the feedback loop.

The bioreactor is operated in batch mode, mimicking industrial conditions and allowing cell densities to reach high levels. This means, that cells are initially exponentially growing (approximately for the first 24 hours) until they reach stationarity. The bioreactor run is stopped when a rise in dissolved oxygen signals the end of fermentation (≈ 70 h, fig. S2). Batch fermentation resembles the bioprocessing operating conditions more closely as systems that are operated in turbidostat mode (constant optical density OD₆₀₀), but makes the control harder. In production facilities, the batch phase is followed by a nutrient feeding phase, allowing the cell density to reach even higher levels and increasing productivity further. To reduce complexity, a feeding phase was not carried out.

5.2 Methods

5.2.1 Media

The SD-2xSCAA medium [23] (designed for heterologous protein secretion) is used for all precultures and experiments and contains 20 g/L D-glucose, 6.9 g/L yeast nitrogen base without amino acids, folic acid and riboflavin, 1 g/L bovine serum albumin (BSA), amino acids as listed in supplementary table S4, 5.4 g/L Na_2PO_4 and 8.56 g/L $\text{NaH}_2\text{PO}_4 \cdot \text{H}_2\text{O}$ (pH = 6.0 by H_2SO_4). For bioreactor experiments the phosphate buffer is replaced with 2 g/L KH_2PO_4 (pH = 6.0 by H_2SO_4). To inhibit foaming, we added 250 μL of 20% polypropylenglycol (PPG) to the bioreactor media at the beginning of the run. All experiments were performed in media supplemented with *myo*-inositol to a final inositol concentration of 20mg/l. To prepare SD-2xSCAA plates (containing the higher concentration of phosphate buffer), BSA is not added while 20 g/L agar is added to the medium. For precultures, the antibiotics Hygromycin B and Geneticin G418 were added at a final concentration of 200 $\frac{\mu\text{g}}{\text{ml}}$.

Amino acids, α -amylase, D-glucose, phosphate buffers, *myo*-inositol and BSA were purchased from Sigma-Aldrich. Geneticin was purchased from Gibco, Hygromycin B from Invitrogen, yeast nitrogen base from Formedium, ammonium sulfate from Roth and PPG from Fluka.

5.2.2 Culture conditions

For all experiments, yeast strains were streaked from the frozen glycerol stock on a SD-2xSCAA plate and incubated at 30°C for 3 days. Precultures were inoculated from single colonies into 25 ml of SD-2xSCAA media in a 250 baffled flask and incubated overnight to an $OD_{600} \approx 1.25$. The preculture cell density was measured by flow cytometry and experiments started at a cell density of 200 $\frac{\text{cells}}{\mu\text{l}}$ ($OD_{600} \approx 0.02$). All experiments were shielded from ambient light to avoid unwanted optogenetic activation of the circuit.

For the final amylase production, cells were grown in a 1.0L Eppendorf BioFlo120 water-jacketed bioreactor with a working volume of 500 ml, 300 rpm agitation and 2.0 SLPM (standard liter per minute). Before autoclaving, the bioreactor was fully assembled and YNB and water added in the bioreactor. The bioreactor plus the tubing leading up to the first dilution stage are autoclaved with the bioreactor autoclaving program at 120°C for 120 minutes with a 2L reference volume. Autoclaving is usually performed in the evening before an experiment. The bioreactor is cooled down and stirring and air flow are turned on to 50 rpm and 0.5 SLPM as soon as the temperature

Chapter 5. Closed-loop bioreactor experiments

decreases below 40°C. The remaining components for the media are added in the morning in the following order PPG, Glucose, Amino Acids, KH Buffer, BSA, Inositol. pH and DO control are switched on sequentially, and the lights of the light housing are tested.

The dissolved oxygen was controlled above 20% using the InPro 6800 Series O₂ Sensor (Mettler Toledo, Switzerland) by automatic adjustment of the impeller speed (between 300 and 500 rpm) and air flow rate (between 2.0 and 3.0 SLPM). The pH was maintained between 5.95 and 6.05 by a Type 405-DPAS-SC pH sensor (Mettler Toledo, Switzerland) using 4 M NaOH and 10% H₂SO₄. All fermentations were done in biological duplicates (figure S5 and S6).

5.2.3 Flow cytometry

Fluorescence and cell density measurements were performed on a Beckman Coulter CytoFLEX S. The internal quality control program was run before experiments using CytoFLEX Daily QC Fluorospheres. Events were gated in the FSC-A and SSC-A channels to remove measurements of debris (see figure 3.1). Samples were diluted by the automatic dilution device to cell densities of 100-4'000 $\frac{cells}{\mu l}$. All fluorescence measurements are reported as the mean of the gated population in arbitrary units (a.u.). To run the CytFLEX S API successfully, MATLAB needs to be run in “admin” mode.

5.3 First closed-loop proof of principle

As a first test we wanted to show that the individual components of the platform could work together and allow closed-loop control. Figure 5.3 on the left shows a first test of closed-loop control. The controller implemented was a PID with initial guesses of the PID parameters. One can observe clear sustained oscillations with significant over- and undershoots. The right part of the figure shows a closed-loop run with the predicted PI control parameters from section 4.5. The control is much improved with clear signs of damping of the UPRS response over time. The control is not perfect and could be further improved with the usage of e.g. PID or MPC control architecture.

5.3.1 Error rejection and anti-windup

After a couple of runs, we observed that in some rare instances, the sampling or dilution did not work and the flow cytometer would produce negative values for cell density, TR or UPRS. This is likely hardware related, where the final pump does not transfer liquid to the flow cytometer sampling port. The flow cytometer in these cases “measured” air

5.3 First closed-loop proof of principle

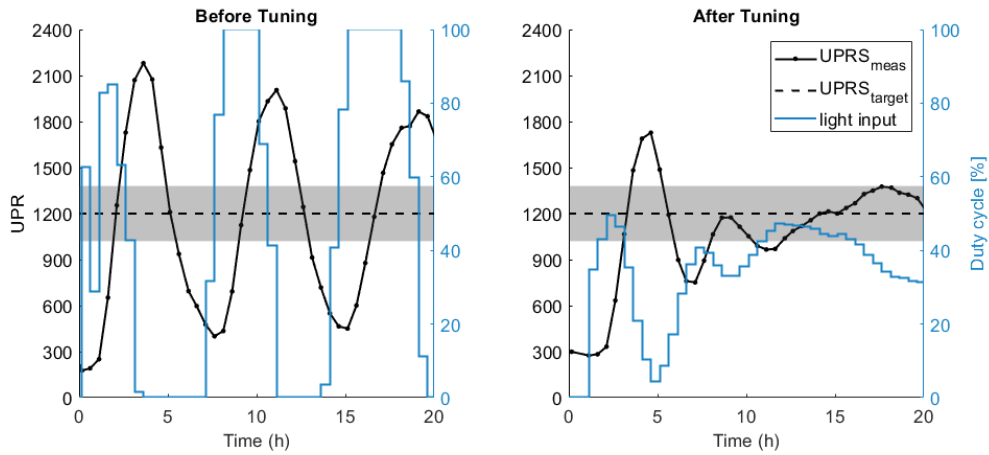


Figure 5.3: Closed loop control of the UPRS with a target ($UPRS_{target}=1200$ a.u.). **left** For initial demonstration we leveraged a PID controller with guessed parameters $K_P = 4.6 \cdot 10^{-6}$, $K_I = 1.8 \cdot 10^{-2}$ and $K_D = 9.9 \cdot 10^1$. **right** Control with a PI controller with parameters $K_P = 8.5 \cdot 10^{-3}$ and $K_I = 9.2 \cdot 10^{-5}$.

and thus produced negative fluorescence data (see figure 5.4). This of course negatively impacted the performance of the closed-loop control. This effect is relatively rare (1.4% of all measurements), but can have negative effects on a whole bioreactor run, especially, if it happens early on. Rather than running a filter over the measured data, we put in a conditional statement that flagged measurements with negative values in either one of the three channels (TR, UPRS or cell density). For the error calculation, the previously measured value was taken to calculate the next controller input. With measurement error rejection, the control input is improved.

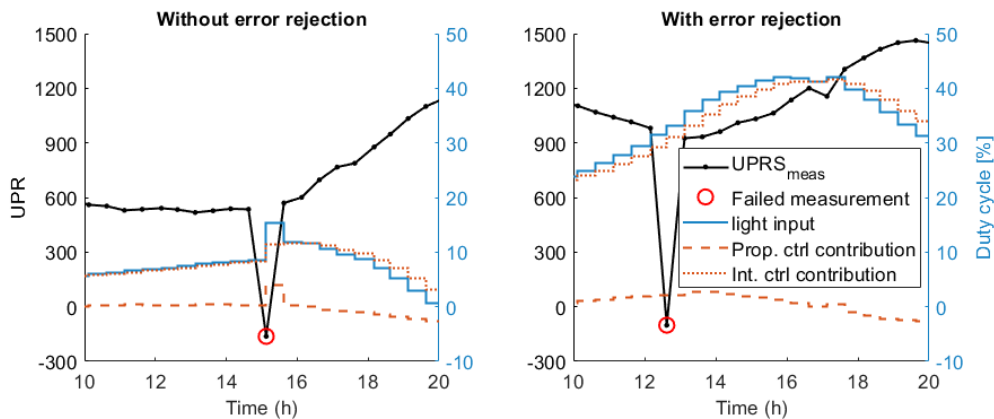


Figure 5.4: Examples of closed loop control of the UPRS without error rejection **left** and with error rejection **right**. Marked with a red circle is the measurement producing negative fluorescence values. Additionally, we plot the control contribution of the proportional and integral gain.

Integrator windup in PI feedback control is well known and present if a physical

variable limit is reached. In this case, the integral term of the controller can accumulate error over time, and if the process being controlled cannot respond quickly enough, the integral term can continue to increase, causing the controller to wind up [24]. This can happen in our system as well, as the light duty cycle can not take negative values or be above 100%. We use a conditional anti-windup scheme to counteract this effect. The control output is constantly monitored and compared to these two limits. If saturation occurs, the integral action is turned off.

5.4 Closed-loop control of the unfolded protein response optimizes the production of a model, hard-to-fold protein

We first evaluated UPR activity during a batch process without and with constant, maximum blue-light induction, respectively (figure 5.5a). Without light, cells express a basal level of UPRS throughout the run, with a transient, stationary-phase-induced increase after 18 hours. Blue-light induction results in a fast and sharp increase of UPR activity which reaches its maximum at 6 h post-induction, after which the UPR adapts to an intermediate activity. The dynamic behaviour of the UPR highlights the need for closed-loop control to achieve constant levels of activity during a run. The experiment is stopped after 70 hours as the dissolved oxygen plateaus (figure S2).

To evaluate the relationship between UPR activity and protein production, we aimed to clamp the UPR to three different levels by optogenetic feedback control. As explained in chapter 4, we chose to modulate the light input by altering the duty cycle at constant light intensity, because of the non-linearity of light intensity to gene expression [25]. Due to the onset of the stationary phase we decided to limit the control to the first 24h and set the light duty cycle to 100% for the next 46 hours (see figure S4 and S6 for full timeline) [26]. Three setpoints were chosen based off the range of UPRS achieved by constant light and in the dark.

Generally, the maximum UPRS that can be achieved by maximum light is 80% lower than in the waterbath setup (compare figures 4.10 and 5.5a). As the transcriptional reporter expression reaches similar levels (compare figures 4.10 and S3) this hints towards more stressful conditions in the waterbath setup in comparison to the bioreactor. Reasons for this could be lower dissolved oxygen levels in the waterbath system, as the culture is only stirred with a small magnetic stirrer and no air bubbling through the cultures. Increased sensitivity of hypoxic cultures to UPR stress is observed compared to well aerated cultures [27]. Similarly, pH control in the waterbath is only provided

5.4 Closed-loop control of the unfolded protein response optimizes the production of a model, hard-to-fold protein

through the addition of a phosphate buffer, so that a pH of 6.0 is not guaranteed throughout the waterbath run. Kawazoe et al. [28] found that non-physiological pH can also trigger the UPR.

We predicted the performance of proportional-integral (PI) control parameters for three setpoints (see section 4.5). For closed-loop control, the operating point in figure 4.12 is chosen for the three setpoints. Performing closed-loop feedback experiments with these parameters, we achieved tracking of all three setpoints (600, 900, and 1200 a.u. UPRS). This tracking is maintained, after an over- and undershoot, while cells are growing exponentially (100-fold range of cell density, fig. S7). To our knowledge, this is the first implementation of closed-loop feedback of a cell state in a high cell density culture.

We collected the supernatant after 70 hours and measured the α -amylase concentration. Figure 5.6 shows the final amylase concentration vs. the mean controlled UPRS in the first 24 hours after inoculation. One can observe that at higher levels of UPRS (constant illumination), the final α -amylase concentration is lower than for the intermediate UPRS setpoints that were achieved using the closed-loop control setup. This follows our hypothesis from figure 5.1. For optimal protein production, the cells should be kept at 900 a.u. of UPRS to improve the titer of α -amylase by 60% over maximally inducing the cells as well as an increase of 159% over cells grown in the dark. We hypothesize that the cells are running into a limitation of the protein production machinery, most likely in the folding machinery. There is additional evidence, that strong and consistent induction of the UPR over a long time, will lead to the formation of inclusion bodies and trigger the ER-associated degradation machinery [29].

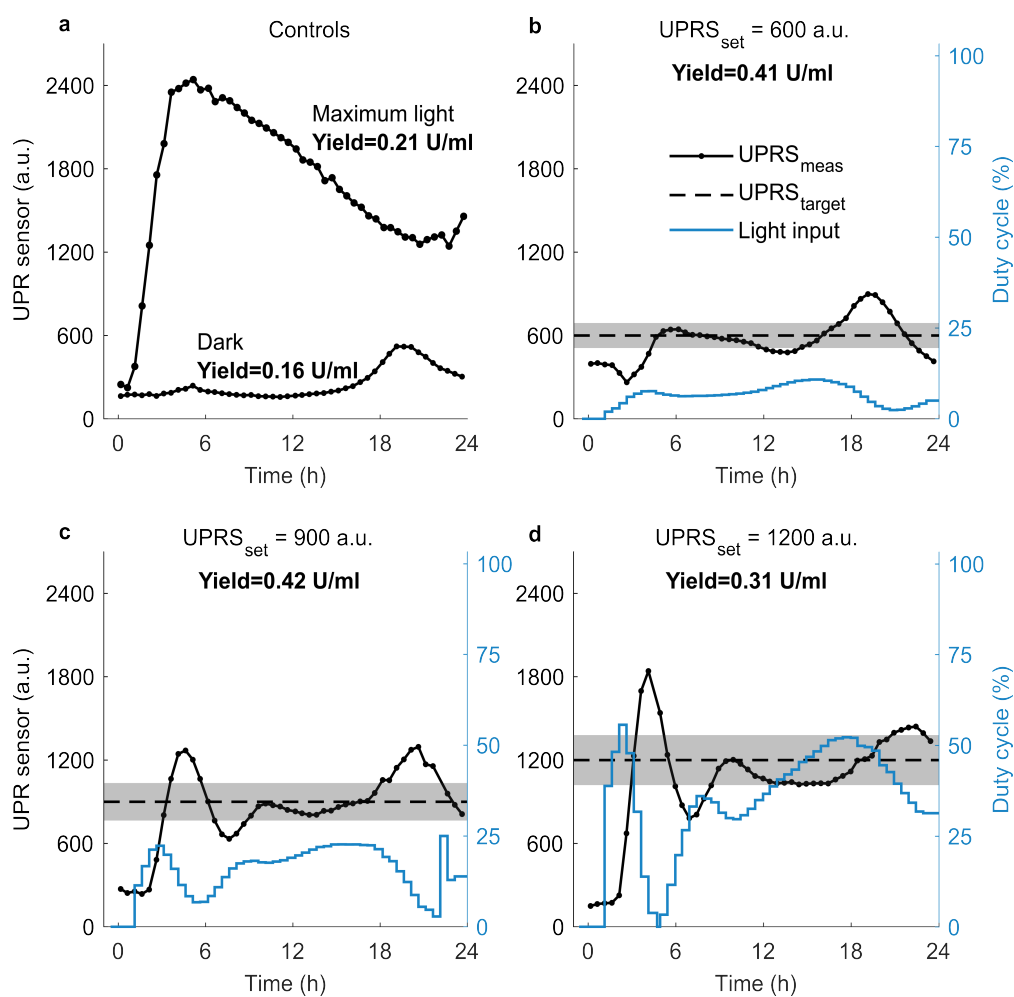


Figure 5.5: **a** Progression of the UPRS over time for dark and maximum light conditions with respective yield of α -amylase. **b-d** Progression of the UPRS over time in the bioreactor setup in the first 24h during which feedback control is performed (second run is shown in figure S5). Black solid line represents the measured UPRS, dashed the target UPRS level with 15% area around it. The blue line represents the blue light duty cycle in % resulting from closed-loop control. The closed-loop UPRS setpoints are 600 (**b**), 900 (**c**) and 1200 a.u. (**d**) resp.

5.4 Closed-loop control of the unfolded protein response optimizes the production of a model, hard-to-fold protein

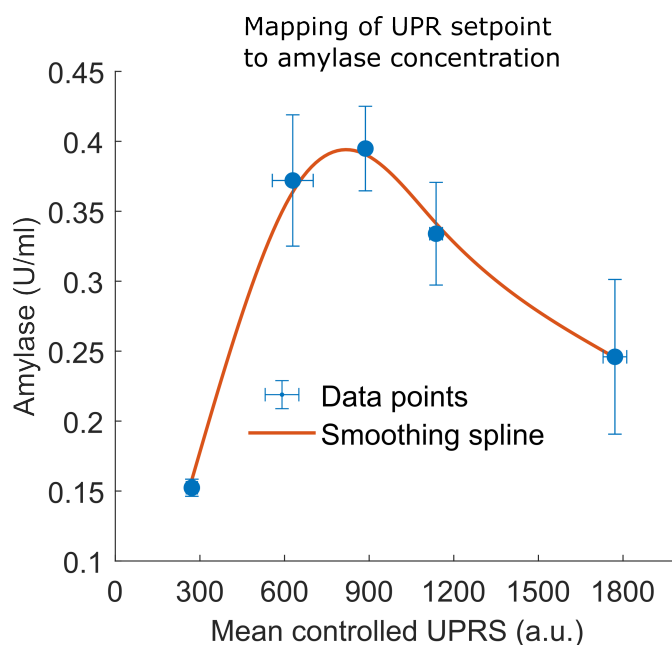


Figure 5.6: Mean and standard deviation of the final amylase concentration after 70 hours is plotted against the mean and standard deviation of the UPRS during the controlled time (first 24 hours) as shown for panels b-e (two independent experiments per condition, three for maximum light). We perform Welch's t-test and find that the α -amylase yield at the closed-loop setpoints 600 and 900 a.u. is higher than the yield with maximal induction at a significance level of 4.6% and 1.5%. The closed-loop setpoint at 1200 a.u. achieves a yield that can not be deemed statistically higher from maximal induction ($\alpha = 6.2\%$).

5.5 Insights into the time dynamics of α -amylase production

We had also implemented an automatic amylase sampler in section 2.3. This sampler allowed the collection of samples every six to twelve hours during the 72 hour fermentation run. Samples were either acquired automatically, or manually at the end of the run. The amylase concentration was quantified offline.

Figure 5.7 shows the progression of amylase concentration in the supernatant over time for the different experimental conditions. One can see that the production in the first 18 hours is very low, followed by a jump around 25 hours after inoculation. This suggests that the control of UPRS that was achieved in the first 24 hours has a long-lasting effect on production. It would also be interesting to see whether maintaining the cells in the state achieved around 25 hours (early stationary phase), can boost productivity for a longer period.

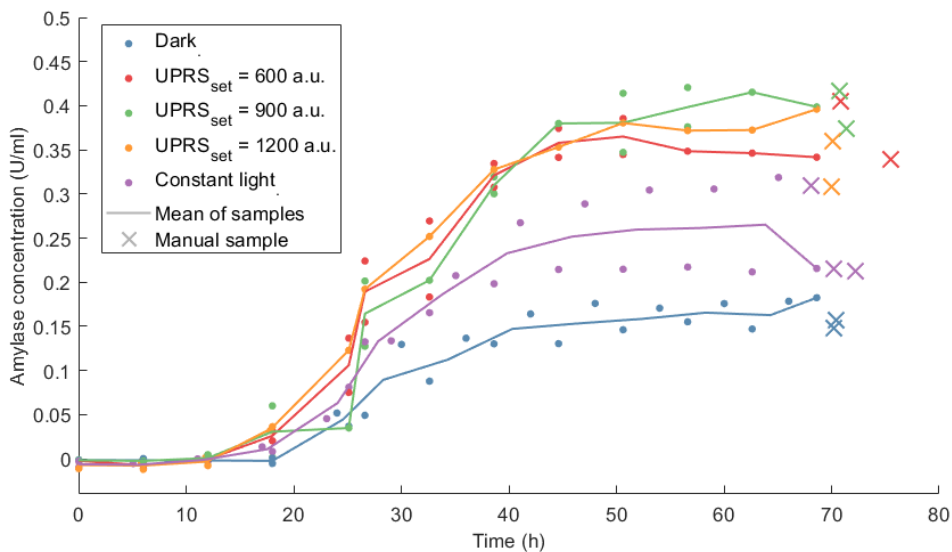


Figure 5.7: The progression of the concentration of α -amylase in the supernatant as a function of time for the five different experimental conditions. Shown are the individually acquired samples, the mean of the samples (if more than one sample was acquired) and the manually sampled measurement. Note that only one replicate for the condition of $\text{UPRS}_{set} = 1200$ a.u. is shown.

While allowing to gain some general insights, the valves of the amylase sampler proved to be prone to clog or fail. Thus, over the course of the experiments only 85% of possible samples were able to be collected. Additionally, to biologically inactivate the sample and kill any yeasts, the biocide sodium azide (NaAz) was added to each vial. The assay to quantify amylase is highly pH sensitive. Thus it is important to keep the ratio of

sodium azide to sample constant. With constant amount of NaAz, the sampled mass of sample was 0.889 ± 0.29 g. Thus, while allowing insights into the general dynamics of production, this setup does not allow to differentiate between different runs. We have thus reported only the manually acquired samples in figure 5.6.

5.6 Discussion

We leveraged the platform to control the level of our UPR sensor to different setpoints and study the effect on protein yield. To our knowledge, closed-loop control of a cellular internal parameter has not been performed in exponentially growing liquid cultures and at the high cell densities obtained in this fermentation. The control architecture for closed-loop cybergenetic control of the UPR is centered around simple proportional-integral control as a first proof-of-principle. More advanced control schemes such as model predictive control (MPC) or lead-lag controllers could generate even more robust transient behaviour. The existing mathematical model can be quickly adapted for such an implementation. With the rise of new *in situ* measurements of the cell physiological state at high cell densities (e.g. online raman spectroscopy probes), we could potentially further improve the process control.

The idea of using closed-loop optogenetic feedback control for optimal bioproduction has been extensively discussed in the community [17, 18, 30, 31]. However, a demonstration of improved bioproduction by closed-loop optogenetic feedback control at an industrially relevant volume scale is still missing. In this study, we close this gap. We demonstrate that controlling the level of UPR to intermediate levels maximizes production of proteins in comparison to maximal protein expression. The bioprocessing community has historically focused to control global process parameters such as temperature, dissolved oxygen and pH. This study highlights the importance of cell internal states for bioproduction. We hypothesize that with a given cell background, media and all other things unchanged, the UPR-yield map (figure 5.6) could be applicable to other proteins of interest. However, further investigation is required to show this. This new way of producing proteins using cybergenetics has the potential to revolutionize and massively simplify the upscaling of new proteins of interest. Additionally, simultaneous characterization of the production process in combination with closed-loop real-time feedback could allow a reduction of overall turnaround and upscaling times. This implementation of *in silico* closed-loop feedback at a high cell density cultivation (HCDC) level opens new possibilities about how to express and produce proteins efficiently.

Bibliography

- [1] Guokun Wang, Mingtao Huang, and Jens Nielsen. Exploring the potential of *Saccharomyces cerevisiae* for biopharmaceutical protein production. *Current Opinion in Biotechnology*, 48:77–84, 2017. ISSN 18790429. doi: 10.1016/j.copbio.2017.03.017. URL <http://dx.doi.org/10.1016/j.copbio.2017.03.017>.
- [2] A. Mellitzer, R. Weis, A. Glieder, and K. Flicker. Expression of lignocellulolytic enzymes in *Pichia pastoris*. *Microbial Cell Factories*, 11, 2012. ISSN 01681656. doi: 10.1186/1475-2859-11-61.
- [3] Elena Camara, Joan Albiol, and Pau Ferrer. Droplet Digital PCR-Aided Screening and Characterization of *Pichia pastoris* Multiple Gene Copy Strains. *Biotechnology and Bioengineering*, 113(7), 2016.
- [4] Carissa L. Young and Anne S. Robinson. Protein folding and secretion: Mechanistic insights advancing recombinant protein production in *S. cerevisiae*. *Current Opinion in Biotechnology*, 30:168–177, 2014. ISSN 18790429. doi: 10.1016/j.copbio.2014.06.018. URL <http://dx.doi.org/10.1016/j.copbio.2014.06.018>.
- [5] Claudia Rubio, David Pincus, Alexei Korennykh, Sebastian Schuck, Hana El-Samad, and Peter Walter. Homeostatic adaptation to endoplasmic reticulum stress depends on Ire1 kinase activity. *Journal of Cell Biology*, 193(1):171–184, 2011. ISSN 00219525. doi: 10.1083/jcb.201007077.
- [6] Sijie Yu, Liangtian Miao, He Huang, Yin Li, and Taicheng Zhu. High-level production of glucose oxidase in *Pichia pastoris*: Effects of Hac1p overexpression on cell physiology and enzyme expression. *Enzyme and Microbial Technology*, 141(1):109671, 2020. ISSN 18790909. doi: 10.1016/j.enzmictec.2020.109671. URL <https://doi.org/10.1016/j.enzmictec.2020.109671>.
- [7] Carine Beaupere and Vyacheslav M. Labunskyy. (Un)folding mechanisms of adaptation to ER stress: lessons from aneuploidy. *Current Genetics*, 65(2):467–471, 2019. ISSN 14320983. doi: 10.1007/s00294-018-0914-9. URL <http://dx.doi.org/10.1007/s00294-018-0914-9>.
- [8] José L. Martínez, Eugenio Meza, Dina Petranovic, and Jens Nielsen. The impact of respiration and oxidative stress response on recombinant α -amylase production by *Saccharomyces cerevisiae*. *Metabolic Engineering Communications*, 3:205–210, 2016. ISSN 22140301. doi: 10.1016/j.meteno.2016.06.003. URL <http://dx.doi.org/10.1016/j.meteno.2016.06.003>.

-
- [9] Peng Xu, Qin Gu, Wenya Wang, Lynn Wong, Adam G.W. Bower, Cynthia H. Collins, and Mattheos A.G. Koffas. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nature Communications*, 4:1408–1409, 2013. ISSN 20411723. doi: 10.1038/ncomms2425. URL <http://dx.doi.org/10.1038/ncomms2425>.
- [10] Francesca Ceroni, Alice Boo, Simone Furini, Thomas E. Goroehowski, Olivier Borkowski, Yaseen N. Ladak, Ali R. Awan, Charlie Gilbert, Guy Bart Stan, and Tom Ellis. Burden-driven feedback control of gene expression. *Nature Methods*, 15(5):387–393, 2018. ISSN 15487105. doi: 10.1038/nmeth.4635.
- [11] Andreas Miliadis-Argeitis, Sean Summers, Jacob Stewart-Ornstein, Ignacio Zuleta, David Pincus, Hana El-Samad, Mustafa Khammash, and John Lygeros. In silico feedback for in vivo regulation of a gene expression circuit. *Nature Biotechnology*, 29(12):1114–1116, 2011. ISSN 10870156. doi: 10.1038/nbt.2018.
- [12] Justin Melendez, Michael Patel, Benjamin L. Oakes, Ping Xu, Patrick Morton, and Megan N. McClean. Real-time optogenetic control of intracellular protein concentration in microbial cell cultures. *Integrative Biology (United Kingdom)*, 6(3):366–372, 2014. ISSN 17579708. doi: 10.1039/c3ib40102b.
- [13] Andreas Miliadis-Argeitis, Marc Rullan, Stephanie K. Aoki, Peter Buchmann, and Mustafa Khammash. Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth. *Nature Communications*, 7(May):1–11, 2016. ISSN 20411723. doi: 10.1038/ncomms12546. URL <http://dx.doi.org/10.1038/ncomms12546>.
- [14] Patrick Harrigan, Hiten D. Madhani, and Hana El-Samad. Real-Time Genetic Compensation Defines the Dynamic Demands of Feedback Control. *Cell*, 175(3):877–886.e10, 2018. ISSN 10974172. doi: 10.1016/j.cell.2018.09.044. URL <https://doi.org/10.1016/j.cell.2018.09.044>.
- [15] Harrison Steel, Robert Habgood, Ciarán Kelly, and Antonis Papachristodoulou. In situ characterisation and manipulation of biological systems with Chi.Bio. *PLoS Biology*, 18(7):1–12, 2020. ISSN 15457885. doi: 10.1371/journal.pbio.3000794.
- [16] Chetan Aditya, François Bertaux, Gregory Batt, and Jakob Rues. A light tunable differentiation system for the creation and control of consortia in yeast. *Nature Communications*, 12(1), 2021. ISSN 20411723. doi: 10.1038/s41467-021-26129-7. URL <http://dx.doi.org/10.1038/s41467-021-26129-7>.

- [17] Makoto A. Lalwani, Evan M. Zhao, and José L. Avalos. Current and future modalities of dynamic control in metabolic engineering. *Current Opinion in Biotechnology*, 52:56–65, 2018. ISSN 18790429. doi: 10.1016/j.copbio.2018.02.007.
- [18] Sylvain Pouzet, Alvaro Banderas, Matthias Le Bec, Thomas Lautier, Gilles Truan, and Pascal Hersen. The Promise of Optogenetics for Bioproduction : Dynamic Control Strategies and Scale-Up Instruments. *Bioengineering*, 2020.
- [19] Mari Valkonen, Merja Penttilä, and Markku Saloheimo. Effects of inactivation and constitutive expression of the unfolded-protein response pathway on protein production in the yeast *Saccharomyces cerevisiae*. *Applied and Environmental Microbiology*, 69(4):2065–2072, 2003. ISSN 00992240. doi: 10.1128/AEM.69.4.2065-2072.2003.
- [20] Mengmeng Huang, Yanyun Gao, Xiangshan Zhou, Yuanxing Zhang, and Menghao Cai. Regulating unfolded protein response activator HAC1p for production of thermostable raw-starch hydrolyzing alpha-amylase in *Pichia pastoris*. *Bio-process and Biosystems Engineering*, 40(3):341–350, 2017. ISSN 16157605. doi: 10.1007/s00449-016-1701-y.
- [21] Keith E.J. Tyo, Zihe Liu, Dina Petranovic, and Jens Nielsen. Imbalance of heterologous protein folding and disulfide bond formation rates yields runaway oxidative stress. *BMC Biology*, 10, 2012. ISSN 17417007. doi: 10.1186/1741-7007-10-16.
- [22] Zihe Liu, Keith E.J. Tyo, Jose L. Martinez, Dina Petranovic, and Jens Nielsen. Different Expression Systems for Production of Recombinant Proteins in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering*, 2012. doi: 10.1002/bit.24409.
- [23] K. D. Wittrup and V. Benig. Optimization of amino acid supplements for heterologous protein secretion in *Saccharomyces cerevisiae*. *Biotechnology Techniques*, 8(3):161–166, 1994. ISSN 0951208X. doi: 10.1007/BF00161582.
- [24] Karl Johan Åström and Tore Hägglund. PID controllers: theory, design, and tuning, 1995.
- [25] Dirk Benzinger and Mustafa Khammash. Pulsatile inputs achieve tunable attenuation of gene expression variability and graded multi-gene regulation. *Nature Communications*, 9(2018):1–38, 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05882-2. URL <http://dx.doi.org/10.1038/s41467-018-05882-2>.
- [26] Evan M. Zhao, Makoto A. Lalwani, Jhong Min Chen, Paulina Orillac, Jared E. Toettcher, and José L. Avalos. Optogenetic Amplification Circuits for Light-Induced

-
- Metabolic Control. *ACS Synthetic Biology*, 10(5):1143–1154, 2021. ISSN 21615063. doi: 10.1021/acssynbio.0c00642.
- [27] Huong Thi Phuong, Yuki Ishiwata-Kimata, Yuki Nishi, Norie Oguchi, Hiroshi Takagi, and Yukio Kimata. Aeration mitigates endoplasmic reticulum stress in *Saccharomyces cerevisiae* even without mitochondrial respiration. *Microbial Cell*, 8(4):77–86, 2021. ISSN 23112638. doi: 10.15698/MIC2021.04.746.
- [28] Nozomi Kawazoe, Yukio Kimata, and Shingo Izawa. Acetic acid causes endoplasmic reticulum stress and induces the unfolded protein response in *Saccharomyces cerevisiae*. *Frontiers in Microbiology*, 8(JUN):1–10, 2017. ISSN 1664302X. doi: 10.3389/fmicb.2017.01192.
- [29] Sebastián Bernales, Feroz R. Papa, and Peter Walter. Intracellular signaling by the unfolded protein response. *Annual Review of Cell and Developmental Biology*, 22:487–508, 2006. ISSN 10810706. doi: 10.1146/annurev.cellbio.21.122303.120200.
- [30] Evan M. Zhao, Yanfei Zhang, Justin Mehl, Helen Park, Makoto A. Lalwani, Jared E. Toettcher, and José L. Avalos. Optogenetic regulation of engineered cellular metabolism for microbial chemical production. *Nature*, 555(7698):683–687, 2018. ISSN 14764687. doi: 10.1038/nature26141.
- [31] César Carrasco-López, Sergio A. García-Echauri, Therese Kichuk, and José L. Avalos. Optogenetics and biosensors set the stage for metabolic cybergenetics. *Current Opinion in Biotechnology*, 65:296–309, 2020. ISSN 18790429. doi: 10.1016/j.copbio.2020.07.012.

6 Discussion and Outlook

In this PhD thesis I have described the development of an illumination system for a commercial bioreactor setup and the development for an automatic sampling and dilution system for automation. Moreover, I have genetically engineered *S. cerevisiae* to respond to blue light [1, 2], express transcriptional reporter and unfolded protein response sensor [3] and produce α -amylase. I have proposed a mathematical model on the basis of previous mathematical model [2, 4], which is able to capture the key dynamics of the UPR as a response of the optogenetic input. I use this model to tune a proportional-integral (PI) controller *in silico*. Lastly, I demonstrate that controlling the UPRS to intermediate levels improves α -amylase production by 60% in comparison to maximal UPR stress. This work has been published in Metabolic Engineering (Benisch et al. [5]).

6.1 Strain development

The goal of the strain development was to produce a *S. cerevisiae* strain that is expressing α -amylase only in the presence of light and for long production runs. In contrast to many publications in the literature using optogenetic promoters that run experiments only in exponential growth and for short time periods (<12h), protein production is highest at high cell densities that are reached in the late-exponential phase. Significant work was performed to ensure tight optogenetic regulation, by changing the promoter sequence and adapting the media composition. While we achieve tight optogenetic regulation during exponential growth in the first 18 hours (figure 3.6), upon the onset of stationary phase the expression of reporters is reduced, while amylase secretion is present in all conditions, whether exposed to light or not (figure 5.7). This is in line with previous work observing considerable activity of some promoters when cells transition from exponential to stationary growth [6]. Closed loop control was thus stopped after 24 hours and the

production system was run in open loop at maximum induction.

We had originally engineered our yeast with an EL222-responsive promoter from Benzinger and Khammash [2] and had subsequently moved to a synthetic CYC1 minimal promoter from Ottoz et al. [7]. While this new promoter contains less transcription factor binding sites, expression of amylase was still observed. Characterization of this observation is additionally hampered, as the fluorescent reporters for transcriptional and UPR readout indicate a leveling off of transcription and folding stress. This is likely due to reduced maturation rates of the fluorescent reporters. To understand the gene expression from the P_{EL222} promoter in the stationary phase, one could run a qPCR or Western Blot analysis to quantify the mRNA transcripts and protein levels respectively of the reporters and amylase inside the cell. Transcription in the stationary phase is not well understood [8]. Korber and Barbaric [9] e.g. found that chromatin state has an effect on transcription factor site occlusion massively impacting gene transcription. Similarly, the different control mechanisms to regulate intracellular pH in stationary cells [10] could lower the functionality of the optogenetic system. Two possible strategies can be applied to overcome these issues.

One could try to engineer the promoter and fluorescent reporters to be fully operational in the stationary phase. In *E. coli* at the onset of stationary phase, many targets are controlled by the sigma factor σ^{S} [11]. This could be a potential starting site for engineering a synthetic promoter that is tight but also responsive during stationary phase. Yu et al. [12] engineered a promoter in *B. subtilis* for high-level expression during stationary phase. They used publicly available genome-wide microarray data to identify potential promoters that were active during late exponential and early stationary phase. This approach could be used to find promoter regions also in yeast that allow transcriptional control during stationary phase. With such an improved promoter-reporter system, closed-loop control could be run for longer times reaping potentially greater benefits for production.

An alternative strategy would be to run the process differently, moving away from batch production with glucose limitation to a fedbatch or continuous culture. In both these systems, cells are initially grown in a batch phase, after which a nutrient feed is switched on to allow for further growth. Because the cells are not in nutrient starvation mode, it is expected that optogenetical control and fluorescent protein maturation works better. The difference between fedbatch and continuous culture is that in continuous cultures, part of the cell-containing media is removed, resulting in a steady-state. Both processes would ensure nutrient availability after the initial batch phase, thus possibly allowing further control over the production system. Continuous culture would be addi-

tionally advantageous for process control, as steady-states are inherently reached with a constant input. This has been shown in Carrillo et al. [13] who run their optogenetic production of amylase in turbidostat mode at very low optical densities and achieve constant UPR with constant input even without process control. While disadvantages of turbidostat mode of operation are apparent as massively reduced cell density and thus production efficiency, this may still be a possible direction for future work.

With the current version of the strain, we stopped closed-loop control at the onset of stationary phase after 24 hours. Following this, we set the light intensity to 100%, to allow comparability of the closed-loop conditions with maximum constant light. At the same time, the initial 24 hours yielded much lower duty cycles of 20-30%. We expect that setting the light duty cycle after 24 hours to these values, could result in improved product formation.

In [5], we demonstrated single-input single-output control. Already now, we are able to track the dynamics of multiple nodes of the circuit allowing multi-objective control. Our photobioreactor platform is equipped with three different LEDs (blue, red and green), allowing the illumination of the culture with three different wavelengths. The first publication of an optogenetic system in *S. cerevisiae* was from Shimizu-Sato et al. [14] and demonstrated the implementation of the red-light optogenetic tool PhyB-PIF3. Implementing such a second optogenetic tool with different actuation wavelength would double the degrees of freedom. Similarly, the multiplexing approach from Benzinger et al. allows multi input with a single wavelengths. These strategies could target a second actuating knob in protein production, such as the availability of chaperone or the upregulation of the heat shock response transcription factor Hsf1p [16].

So far, we also have not investigated the effect of transcriptional reporter expression on the unfolded protein response. There is clear evidence however, that the expression of an easy-to-fold protein such as fluorescent proteins will likely not significantly upregulate the UPR [13]. There may also be an effect of the high-level expression of amylase and thus overburdening of the ER on the expression and folding rate of the fluorescent reporters. A worry is, that a significant downregulation of the fluorescent proteins will hinder control of the UPR itself. To study this, one could introduce a constitutive expressing fluorescent color and observe how the levels are varying as a function of amylase induction. This would allow the quantification of burden, whether stemming from the transcriptional or folding side. Additionally, one could replace a fluorescent protein with fluorescent RNA aptamers, which would allow quantification of transcriptional and folding stress, as RNA aptamers in contrast to fluorescent proteins do not require folding in the ER.

We also propose using the circuit for some in-depths studying of the unfolded protein response in yeast. Currently, this part has not been explored. For example, the strain would lend itself ideally for culturing in a microfluidic device with single-cell measurement and tracking. With this, we could be able to understand different subpopulations of UPR response (as suggested in [13]). Further, a more frequent measurement of the cellular response with more light inputs, could allow a rigorous identification of parameters of the unfolded protein response. Lastly, Leber et al. [17] report the presence of a super UPR (SUPR) in the presence of two stressors (inositol deficiency and unfolding). We suspect, that this effect is also present at sustained high UPR stress, but were unable to rigorously show this.

Overall, the strain is a nice tool to study expression-induced unfolded protein response *in vivo* especially during exponential growth. Usage of it for characterization of the UPR in different setups (e.g. microscope and mother machine), even without the described changes, could reap quick wins.

6.2 UPR-yield map

The project demonstrates the advantage of closed-loop control on the production of a single model protein. We show that intermediate levels of UPR achieve higher yield than maximum gene expression. In our publication [5], we do not propose an exact mechanism that explains the shape of the curve. However, we believe that the presence of excessive amounts of unfolded proteins in the endoplasmic reticulum has a long lasting negative impact on the cell as necessary maintenance can not be carried out to sufficient levels. Possibly, there is also an upregulation of the ER-associated degradation pathway, resulting in many unfolded proteins being degraded rather than secreted. PDI1 is responsible for the formation of disulfide bonds in yeasts [18]. In the case of unsuccessful formation of the cysteine bonds, either a glucosyltransferase binds and initiates another cycle of PDI1 aided folding or YOS9 binds and initiates ER associated degradation (ERAD) [19]. One can imagine that high levels of unfolded proteins in the ER increase the chance of unsuccessful disulfide bond formation and thus elevated ERAD levels.

We also believe that the obtained UPR-Yield map (figure 5.6) could be extended to other proteins of different size, and complexity, as well as other metabolites. Carrillo et al. [13] have already moved in this direction. In their recent preprint, they expressed six different proteins of varying complexity in *S. cerevisiae* to different induction levels and followed the stress and secretion levels over time. These experiments were performed in turbidostat cultures at optical densities of $OD_{600} = 0.5$. For the two most complex

proteins (single chain antibody variable fragment 4M5.3, scFv) and α -amylase, they find a similar optimum between production and UPR stress. Finally, they perform feedback on their UPR sensor for cells producing scFv and find an optimum with 60% increased production over maximum induction. These are very exciting findings, as this preprint confirms our findings, even in different experimental contexts (turbidostat vs. batch), strain backgrounds (BY4741 vs. Cen.PK) and proteins of interest (scFv vs. α -amylase). It should also be noted, that for simple proteins of interest, with the same expression strength range, the trade-off may not be present. I predict that for very simple proteins such as common well-folding fluorescent proteins, the expression strengths that are accessible by our current optogenetic toolkit, all result in UPR stress levels in the left, upwards branch of figure 5.6.

Generally, this approach could also be extended to metabolite production, where either the product itself or intermediates are toxic. Here again, we propose that there exists an optimum expression level, which does not overburden the cell while achieving maximal production. Xu et al. [20] for example tuned the expression levels and found optimal production for non-maximal gene expression. One common bottleneck that is observed in biofuel production is the excessive concentration of biofuel inside the cell resulting in cell death. Swidah et al. [21] report end-product toxicity of butanol at 1%, which severely hampers efficient conversion of substrate into final product. Here, our approach could work nicely, by using closed-loop control to achieve product titers below toxicity levels and continuously separate product from cells, thus allowing more production of biofuels.

We also have to put the yield improvement of 60% in context with the existing literature. Generally, the strain that we had used for amylase production has seen significant production improvements in the last years. Both using random mutagenesis, as well as rational design of the secretory pathway allowed the authors to improve product titer of amylase 6-fold [22, 23]. Other publications presented in the introduction often achieve similar fold improvements. Potentially, these genetic variations have removed any remaining bottlenecks in the expression of proteins. Thus, it is yet to be investigated whether our proposed approach is able to always improve significantly over these genetically modified variants.

6.3 Platform

The platform developed in this PhD has shown reliable performance and has allowed us to achieve closed-loop control of a cell internal state in batch cultures, so far not

achieved in the literature. Full automation of all processes involved have allowed me to prepare and start a fermentation and then walk away for the next 65 hours. The data of cell growth and fluorescent reporters that we were able to generate as a side product are rarely found in the literature. In comparison to the standard bioproduction literature, our data on expression, stress, cell density and product concentration is much more finely resolved. Thus, this dataset has allowed significant insights into the production dynamics of proteins.

Figure 5.7 shows the evolution of amylase concentration during the run. Interestingly, the instantaneous production rate of product is highest in the late exponential phase. This could be because of genetic upregulation of chaperones and secretion factors, thus increasing production in the short term. Fedbatch systems reproduce these growth conditions, by intermittently adding fresh nutrients to the growth chamber. Thus, it could be possible to improve production efficiency by even varying the feeding timing, rather than just concentration and nutrients.

The throughput of testable conditions was limited to one condition every week, as the fermentation run time plus preparation, autoclaving and finally disassembling and cleaning take a majority of a regular work week. For future work in this direction, parallelization of the setup is required to test more conditions and gain more knowledge in a shorter amount of time. Since the beginning of my PhD alternative platforms have come to market and could be useful for this goal. The development of the evotron in the CTSB [24] as well as commercial solutions for mini-bioreactors (e.g. Eppendorf DASbox[®] Mini Bioreactor or Beckman Coulter BioLector), could fulfill the main requirements for this proposed parallelization.

Additionally, we are currently operating our closed-loop optogenetic control system with a flow cytometer attached to the photobioreactor platform. While feasible for this work, cheaper solutions [25] would be possible. One idea is to develop a custom-built flow cell with attached microscope or fluorimeter to measure the mean fluorescence. Such an approach is performed in Melendez et al. [26]. The disadvantages have been discussed extensively in chapter 2, but further development could broaden the applicability in a real industrial setting.

Not presented in this thesis is the distributions of cells that are measured using the flow cytometer. Both transcriptional reporter and UPR sensor show a bimodal behaviour. We had tested a strain that only contained the transcriptional reporter and where the TR was on the production plasmid. We sorted for the two populations and found that only the high-expressing population was able to grow. This hints towards

the second population experiencing plasmid loss and thus the inability to metabolize glucose. The UPR reporter exhibits similar bimodal behaviour. It could be interesting to sort for these two populations as well and see, whether highly producing cells are enriched in the highly expressing UPR cell line, or whether the low UPRS cells are able to better cope with UPR stress. Secondly, as the two populations arise spontaneously, one could think about using the shape of the two distributions as variable for process control. After identifying the high-amylase expressing strain (=producers), one could instead of targetting constant UPRS, try the maximize the portion of producers as a variable for process control.

6.4 Modelling and Control

The mathematical modelling presented in chapter 4 has been sufficient for this project. However, we have discussed several possibilities to further improve the model. Currently, we are also assuming instant maturation of the fluorescent reporters, while in reality fluorescent reporters mature in the order of 10s of minutes to a couple of hours. More specifically, the fluorescent reporters used in this study mature in 17.6 minutes (Venus) and 36 minutes (mScarlet-I) [27]. Some delays that were described in section 4.4.1 could thus be easily captured by including these mechanisms.

At the moment, the data used for characterization stem from a different platform (waterbath) and are limited to the exponential and early stationary phase. At the same time, since the initial model development, at least 11 experimental runs in the bioreactor have been performed, each with finely resolved data about cell size, density and the two fluorescent reporters. This data (chapter 5) could be used to further improve the model fit. Firstly, because the data stems from the bioreactor experiments, better predictions for bioreactor experiments can be made. Secondly, because this data also includes the stage of stationary phase, growth parameters can be fit more accurately.

A missing link in the current model is the understanding of amylase secretion. The next steps in modelling should focus on identifying mechanisms that can clearly explain the transition of unfolded amylase proteins to secreted proteins in the supernatant. Feizi et al. [28] developed a genome-scale model of the secretory pathway in *S. cerevisiae* and allows a comprehensive insights into the underlying processes. The last experimental runs on the bioreactor also contain granular amylase concentration data sampled every 6-12 hours. So far this data has not been used for any model structure identification. Currently, characterization is limited by the lack of sensitivity of the amylase assay for low amounts of product. Thus experiments only yield discernable concentrations

after >18 hours. Thus, for further characterization, the implementation of a secretion sensor, similar as in [13] would be helpful, as its sensitivity is superior to traditional ELISA-based assays.

Industrial heterologous production is run in markedly bigger bioreactor volumes (m^3 vs. L) and at higher cell densities. As the surface to volume ratio decreases with increasing volumes, sufficient illumination of the entire cellular population could be a problem. Pouzet et al. [29] discussed several strategies in their review. Besides the classical illumination of the bioreactor from the sides, they discuss the possibility of adding illumination sources inside the bioreactor, thus increasing irradiation. Lastly, using photobioreactors for algae and repurposing them for microbial fermentation could be investigated. Generally, all these strategies try to overcome the decrease of surface to volume ratio by more lights. Zhao et al. [30] propose the usage of amplification circuits which achieve activation at light duty cycles as low as 1%. Lastly, using slow reverting optogenetic variants [15] could overcome the limitations of insufficient illumination. All these strategies should be investigated first theoretically, before implementing them. Initial efforts in this direction (data not shown) showed, that internal illumination does not increase area sufficient to overcome the needs. With our automated photobioreactor platform, experimental data could be generate that mimic industrial conditions and serve as a basis for future theoretical work on this topic.

As discussed at various points in this thesis the implementation of PI control is only the first step for performing closed-loop control. The platform is easily extendable to run control with PID, model-predictive (MPC), extremum seeking (ESC) or machine learning control systems [31]. Miliás-Argeitis et al. [32] nicely demonstrated the advantages of MPC over PID in a turbidostat setup of *E. coli* controlling the expression of a fluorescent protein with the optogenetic CcaS/CcaR. Given the complex nature of the control task, with limited time horizon, the implementation of MPC could reap quick benefits. Additionally, as the dynamics of amylase production are more rigorously understood, one could derive more optimal operating conditions, that could further improve productivity. In a recent preprint Espinel-Ríos et al. [33] propose the usage of dynamic constraint-based models to formulate a model-based optimal control problem and identify optimal optogenetic inputs in a fedbatch production process. In their theoretical work they can only improve production of the product lactate by $\approx 14\%$ raising the questions, how much improvements can be gained in real-life scenarios.

6.5 Final Remark

Overall with this work, we have demonstrated the successful implementation of optogenetic feedback control at unmatched size and volume. Additionally, we finally close this gap and demonstrated the benefits that closed loop control can have on bioproduction of proteins. We are excited to see, in which direction the field will be moving and looking forward to real industrial implementations of optogenetic closed loop control

Bibliography

- [1] Laura B. Motta-Mena, Anna Reade, Michael J. Mallory, Spencer Glantz, Orion D. Weiner, Kristen W. Lynch, and Kevin H. Gardner. An optogenetic gene expression system with rapid activation and deactivation kinetics. *Nature Chemical Biology*, 10(3):196–202, 2014. ISSN 15524469. doi: 10.1038/nchembio.1430.
- [2] Dirk Benzinger and Mustafa Khammash. Pulsatile inputs achieve tunable attenuation of gene expression variability and graded multi-gene regulation. *Nature Communications*, 9(2018):1–38, 2018. ISSN 2041-1723. doi: 10.1038/s41467-018-05882-2. URL <http://dx.doi.org/10.1038/s41467-018-05882-2>.
- [3] Philip I. Merksamer, Ala Trusina, and Feroz R. Papa. Real-Time Redox Measurements during Endoplasmic Reticulum Stress Reveal Interlinked Protein Folding Functions. *Cell*, 135(5):933–947, 2008. ISSN 00928674. doi: 10.1016/j.cell.2008.10.011. URL <http://dx.doi.org/10.1016/j.cell.2008.10.011>.
- [4] Ala Trusina, Feroz R. Papa, and Chao Tang. Rationalizing translation attenuation in the network architecture of the unfolded protein response. *PNAS*, 105(51):20280–20285, 2008. ISSN 00278424. doi: 10.1073/pnas.0803476105.
- [5] Moritz Benisch, Dirk Benzinger, Sant Kumar, Hanrong Hu, and Mustafa Khammash. Optogenetic closed-loop feedback control of the unfolded protein response optimizes protein production. *Metabolic Engineering*, 77(March):32–40, 2023. ISSN 1096-7176. doi: 10.1016/j.ymben.2023.03.001. URL <https://doi.org/10.1016/j.ymben.2023.03.001>.
- [6] Alon Zaslaver, Anat Bren, Michal Ronen, Shalev Itzkovitz, Ilya Kikoin, Seagull Shavit, Wolfram Liebermeister, Michael G. Surette, and Uri Alon. A comprehensive library of fluorescent transcriptional reporters for *Escherichia coli*. *Nature Methods*, 3:623–628, 2006.
- [7] Diana S.M. Ottoz, Fabian Rudolf, and Joerg Stelling. Inducible, tightly regulated and growth condition-independent transcription factor in *Saccharomyces cerevisiae*. *Nucleic Acids Research*, 42(17), 2014. ISSN 13624962. doi: 10.1093/nar/gku616.
- [8] Paul K. Herman. Stationary phase in yeast. *Current Opinion in Microbiology*, 5: 602–607, 2002. doi: [https://doi.org/10.1016/S1369-5274\(02\)00377-6](https://doi.org/10.1016/S1369-5274(02)00377-6).
- [9] Philipp Korber and Slobodan Barbaric. The yeast PHO5 promoter: from single locus to systems biology of a paradigm for gene regulation through chromatin . *Nucleic Acids Research*, 42:10888–10902, 2014. doi: <https://doi.org/10.1093/nar/gku784>.

-
- [10] Minoska Valli, Michael Sauer, Paola Branduardi, Nicole Borth, Danilo Porro, and Diethard Mattanovich. Intracellular pH distribution in *Saccharomyces cerevisiae* cell populations, analyzed by flow cytometry. *Applied and Environmental Microbiology*, 71(3):1515–1521, 2005. ISSN 00992240. doi: 10.1128/AEM.71.3.1515-1521.2005.
- [11] Stephan Lacour and Paolo Landini. σ S-Dependent Gene Expression at the Onset of Stationary Phase in *Escherichia coli*: Function of σ S-Dependent Genes and Identification of Their Promoter Sequences. *Journal of Bacteriology*, 186:7186–7195, 2004. doi: <https://doi.org/10.1128/JB.186.21.7186-7195.2004>.
- [12] Xiaoxia Yu, Jingtao Xu, Xiaoqing Liu, Xiaoyu Chu, Ping Wang, Jian Tian, Ningfeng Wu, and Yunliu Fan. Identification of a highly efficient stationary phase promoter in *Bacillus subtilis*. *Scientific Reports*, 5, 2016.
- [13] Sebastián Sosa Carrillo, Henri Galez, Sara Napolitano, François Bertaux, and Gregory Batt. Maximizing protein production by keeping cells at optimal secretory stress levels using real - time control approaches. *bioRxiv*, 2022.
- [14] Sae Shimizu-Sato, Enamul Huq, James M. Tepperman, and Peter H. Quail. A light-switchable gene promoter system. *Nature Biotechnology*, 20(10):1041–1044, 2002. ISSN 10870156. doi: 10.1038/nbt734.
- [15] Dirk Benzinger, Serguei Ovinnikov, and Mustafa Khammash. Synthetic gene networks recapitulate dynamic signal decoding and differential gene expression. *Cell Systems*, 13(5):353–364.e6, 2022. ISSN 24054720. doi: 10.1016/j.cels.2022.02.004. URL <https://doi.org/10.1016/j.cels.2022.02.004>.
- [16] Jin Hou, Tobias Österlund, Zihe Liu, Dina Petranovic, and Jens Nielsen. Heat shock response improves heterologous protein secretion in *Saccharomyces cerevisiae*. *Applied Microbiology and Biotechnology*, 97(8):3559–3568, 2013. ISSN 01757598. doi: 10.1007/s00253-012-4596-9.
- [17] Jess H. Leber, Sebastián Bernales, and Peter Walter. IRE1-independent gain control of the unfolded protein response. *PLoS Biology*, 2(8), 2004. ISSN 15449173. doi: 10.1371/journal.pbio.0020235.
- [18] Ronnie Farquhar, Neville Honey, Susan J. Murant, Peter Bossier, Loren Schultz, Donna Montgomery, Ronald W. Ellis, Robert B. Freedman, and Mick F. Tuite. Protein disulfide isomerase is essential for viability in *Saccharomyces cerevisiae*. *Gene*, 108:81–89, 1991. doi: [https://doi.org/10.1016/0378-1119\(91\)90490-3](https://doi.org/10.1016/0378-1119(91)90490-3).

- [19] Reka Szathmary, Regula Biemann, Mihai Nita-Lazar, Patricie Burda, and Claude A. Jakob. Yos9 Protein Is Essential for Degradation of Misfolded Glycoproteins and May Function as Lectin in ERADE. *Molecular Cell*, 19:765–775, 2005. doi: <https://doi.org/10.1016/j.molcel.2005.08.015>.
- [20] Peng Xu, Qin Gu, Wenya Wang, Lynn Wong, Adam G.W. Bower, Cynthia H. Collins, and Mattheos A.G. Koffas. Modular optimization of multi-gene pathways for fatty acids production in *E. coli*. *Nature Communications*, 4:1408–1409, 2013. ISSN 20411723. doi: 10.1038/ncomms2425. URL <http://dx.doi.org/10.1038/ncomms2425>.
- [21] R. Swidah, H. Wang, P.J. Reid, H.Z. Ahmed, A.M. Pisanelli, K.C. Persaud, C.M. Grant, and M.P. Ashe. Butanol production in *S. cerevisiae* via a synthetic ABE pathway is enhanced by specific metabolic engineering and butanol resistance. *Biotechnology for Biofuels*, 8, 2015.
- [22] Mingtao Huang, Yunpeng Bai, Staffan L. Sjöström, Björn M. Hallström, Zihe Liu, Dina Petranovic, Mathias Uhlén, Haakan N. Joensson, Helene Andersson-Svahn, and Jens Nielsen. Microfluidic screening and whole-genome sequencing identifies mutations associated with improved protein secretion by yeast. *Proceedings of the National Academy of Sciences of the United States of America*, 112(34):E4689–E4696, 2015. ISSN 10916490. doi: 10.1073/pnas.1506460112.
- [23] Mingtao Huang, Guokun Wang, Jiufu Qin, Dina Petranovic, and Jens Nielsen. Engineering the protein secretory pathway of *Saccharomyces cerevisiae* enables improved protein production. *PNAS*, 115(47):E11025–E11032, 2018. doi: 10.1073/pnas.1809921115.
- [24] Joaquín Gutiérrez Mena, Sant Kumar, and Mustafa Khammash. Dynamic cybergenetic control of bacterial co-culture composition via optogenetic feedback. *Nature Communications*, 13(1):1–16, 2022. ISSN 20411723. doi: 10.1038/s41467-022-32392-z.
- [25] Harrison Steel, Robert Habgood, Ciarán Kelly, and Antonis Papachristodoulou. In situ characterisation and manipulation of biological systems with Chi.Bio. *PLoS Biology*, 18(7):1–12, 2020. ISSN 15457885. doi: 10.1371/journal.pbio.3000794.
- [26] Justin Melendez, Michael Patel, Benjamin L. Oakes, Ping Xu, Patrick Morton, and Megan N. McClean. Real-time optogenetic control of intracellular protein concentration in microbial cell cultures. *Integrative Biology (United Kingdom)*, 6(3):366–372, 2014. ISSN 17579708. doi: 10.1039/c3ib40102b.

-
- [27] Enrique Balleza, J. Mark Kim, and Philippe Cluzel. Systematic characterization of maturation time of fluorescent proteins in living cells. *Nature Methods*, 15:47–51, 2018.
- [28] A. Feizi, T. Osterlund, Dina Petranovic, S. Bordel, and Jense Nielsen. Genome-scale modeling of the protein secretory machinery in yeast. *PLoS One*, 8, 2013.
- [29] Sylvain Pouzet, Alvaro Banderas, Matthias Le Bec, Thomas Lautier, Gilles Truan, and Pascal Hersen. The Promise of Optogenetics for Bioproduction : Dynamic Control Strategies and Scale-Up Instruments. *Bioengineering*, 2020.
- [30] Evan M. Zhao, Makoto A. Lalwani, Jhong Min Chen, Paulina Orillac, Jared E. Toettcher, and José L. Avalos. Optogenetic Amplification Circuits for Light-Induced Metabolic Control. *ACS Synthetic Biology*, 10(5):1143–1154, 2021. ISSN 21615063. doi: 10.1021/acssynbio.0c00642.
- [31] Robert J. Lovelett, José L. Avalos, and Ioannis G. Kevrekidis. Partial Observations and Conservation Laws: Gray-Box Modeling in Biotechnology and Optogenetics. *Industrial and Engineering Chemistry Research*, 59(6):2611–2620, 2020. ISSN 15205045. doi: 10.1021/acs.iecr.9b04507.
- [32] Andreas Miliadis-Argeitis, Marc Rullan, Stephanie K. Aoki, Peter Buchmann, and Mustafa Khammash. Automated optogenetic feedback control for precise and robust regulation of gene expression and cell growth. *Nature Communications*, 7(May):1–11, 2016. ISSN 20411723. doi: 10.1038/ncomms12546. URL <http://dx.doi.org/10.1038/ncomms12546>.
- [33] Sebastián Espinel-Ríos, Bruno Morabito, Johannes Pohlodek, Katja Bettenbrock, Steffen Klamt, and Rolf Findeisen. Towards a modeling, optimization and predictive control framework for fed-batch metabolic cybergenetics. *arXiv*, pages 1–17, 2023. URL <http://arxiv.org/abs/2302.02177>.

Supplement

S1 Supplementary figures

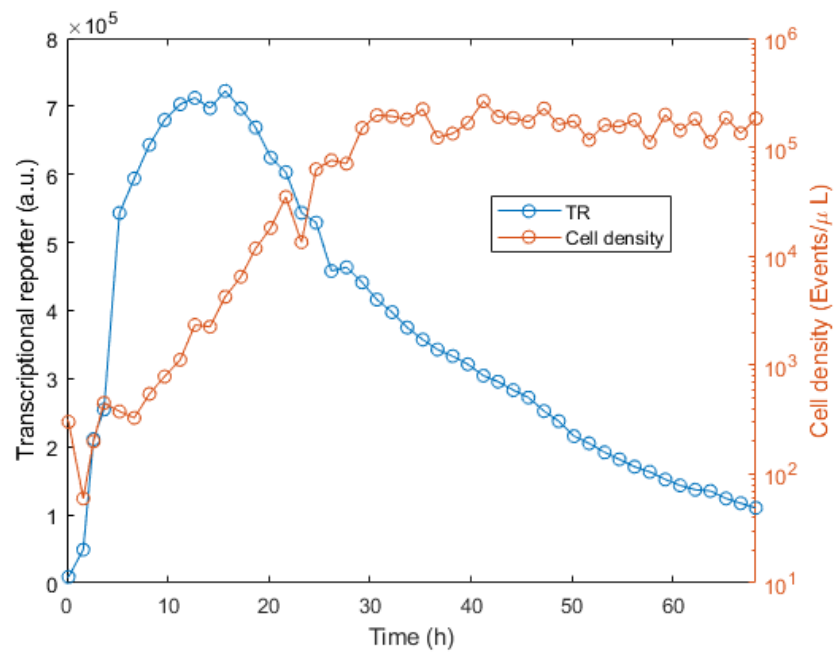


Figure S1: yMB44 is grown in the bioreactor with constant light and the cell density (right y-axis) and transcriptional reporter (left y-axis) are measured with the automatic sampling and dilution device.

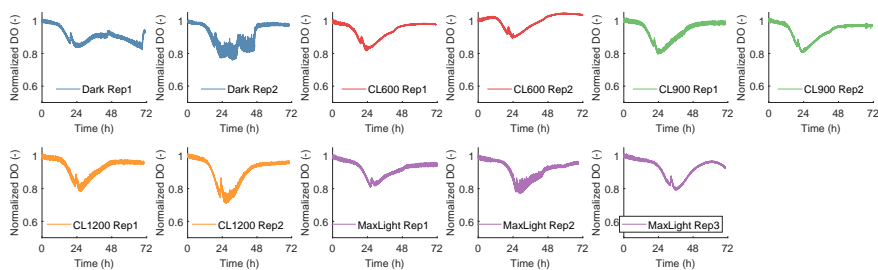


Figure S2: Progression of the dissolved oxygen (DO) normalized by the initial DO value at $t=0$ h.

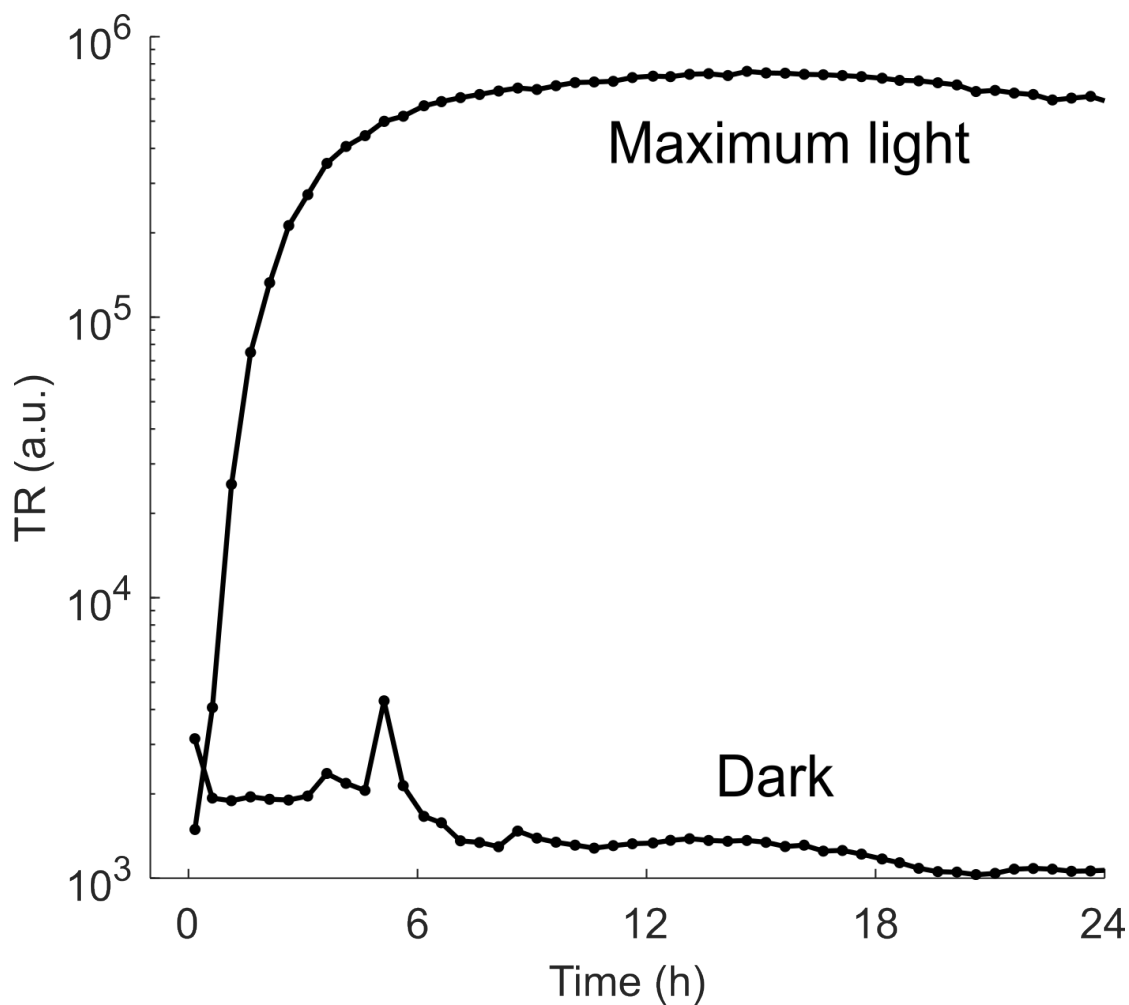


Figure S3: Progression of the TR for constant maximum light and in the dark for repetition 1.

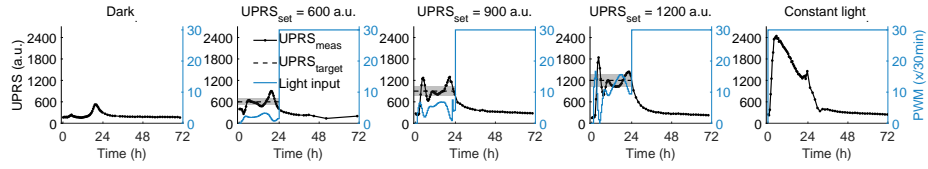


Figure S4: Progression of UPRS over the entire 72 hours of experiment for repetition 1.

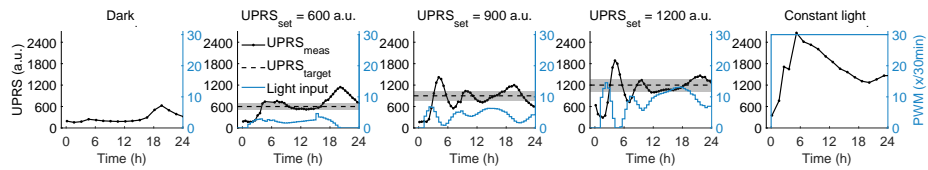


Figure S5: Progression of UPRS for the 2nd replicate and 3rd replicate for maximum light over the first 24 hours of experiment.

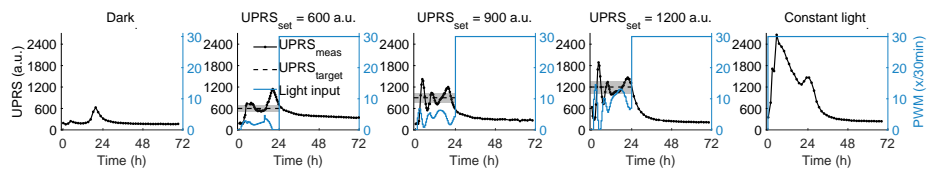


Figure S6: Progression of UPRS for the 2nd replicate over the full 72 hours of experiment.

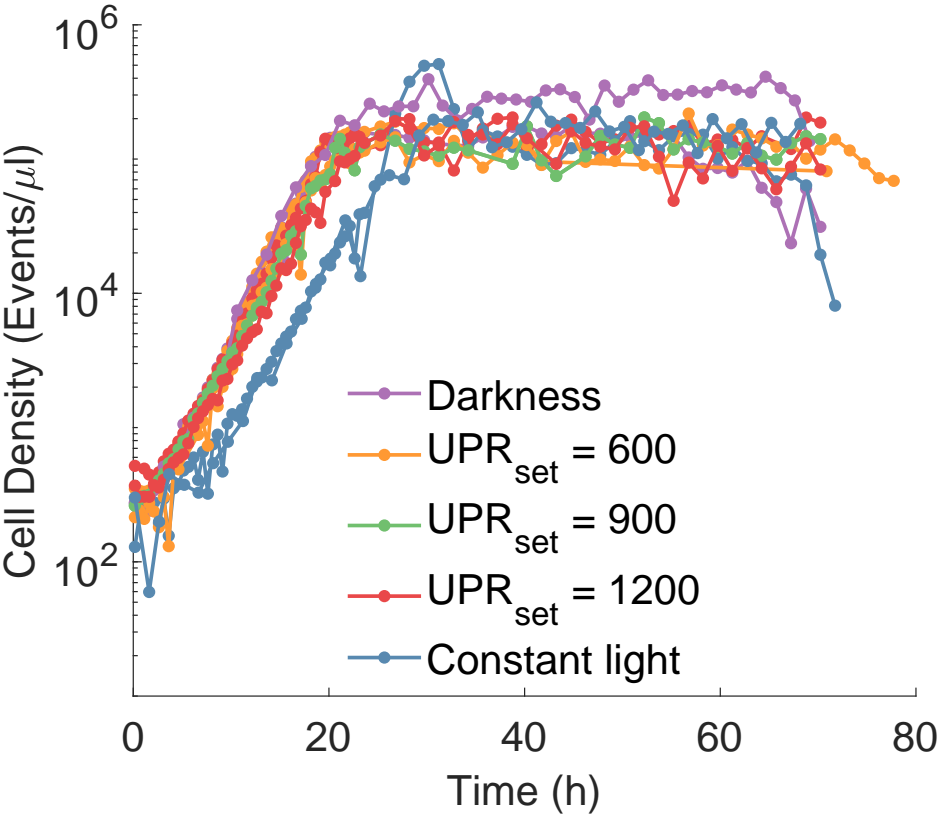


Figure S7: Progression of the cell density in cells per μl over the full 72 hours of experiment.

S2 Supplementary tables

Table S1: Plasmids used for strain construction. Promoters are represented by “pr”, terminators are represented by “t”.

Plasmid	Description	Marker	Source
pAlphaAmyCPOT	TPIpr-Amylase	POT1	[1]
pYMB4	4xUPREpr-UbiY-Venus-ScPGK1t	Hygromycin B	This study
pYMB10	EL222pr-Amylase-TPI1t	POT1	This study
pYMB12v	EL222pr-UbiY-Venus-ScPGK1t	Hygromycin B	This study
pYMB15m	4xUPREpr-UbiY-mScarletI-ScPGK1t	Zeocin	This study
pYTKmk48	ScRPL18Bpr-Msn2AD-EL222-ScENO2t	URA	[2]
pYTK042	3a Part, Ubi-Y (medium)	Chloramphenicol	[3]
pYTK045	3b Part, Venus	Chloramphenicol	[3]
pYTK054	4 Part, tPGK1	Chloramphenicol	[3]
pYTKmk110	Backbone: Hygromycin B resistance, HO-locus homology arms	Kanamycin	[2]
pYTKmk123	3b Part, mScarlet-I	Chloramphenicol	This study
pYTKmk190	2 Part, UPRE-pCYCmin	Chloramphenicol	This study
pYTKmk216	new EL222 promoter	Chloramphenicol	This study
pYTKmk217	MoClo compatible pYTKmk216	Chloramphenicol	This study
pYTKmk227	Backbone: Zeocin resistance, YIRCΔ6-locus homology arms	Kanamycin	This study

Table S2: Oligos used in this study.

Oligo name	Sequence
oMB020	attaagcggccgcGcagctgaAGCTTGCGTT
oMB021	attaagcggccgcGcagctgaAGCTTGCGTT

Table S3: Strains used in this study.

Name	Genotype	Source
CEN.PK 530-1D	MAT α ura3-52 HIS3 LEU2 TRP1 SUC2 MAL2-8 ^c tpi1(41-707)::loxP-KanMX4-loxP	[1]
CEN.PK 113-7D	MAT α URA3 HIS3 LEU2 TRP1 SUC2 MAL2-8 ^c	Euroscarf
yMB9	CEN.PK 113-7D, HO::pYMB4	This study
yMB31	CEN.PK 530-1D, pYMB10	This study
yMB32	yMB31, URA3::pYTKmk48	This study
yMB39	yMB32, HO::pYMB12v	This study
yMB44	yMB39, YIRC Δ 6IX::pYMB15m	This study

Table S4: Amino acid concentration for SD-2xSCAA.

Amino acid	Concentration (mg/l)
Arginine	190
Aspartic Acid	400
Glutamic acid	1260
Glycine	130
Histidine	140
Isoleucin	290
Leucine	400
Lysine	440
Methionine	108
Phenylalanine	200
Threonine	220
Tryptophan	40
Tyrosine	52
Valine	380

Table S5: Estimated parameters for the optogenetic-UPR model.

Parameter	Fitted or fixed	Value	Unit
k_{off}	Fitted	4.77	min^{-1}
TF_{tot}	Fitted	$2.23 \cdot 10^3$	-
k_{on}	Fitted	$1.86 \cdot 10^{-3}$	-
$k_{m,\text{basal}}$	Fixed	0	min^{-1}
k_m	Fitted	$2.51 \cdot 10^4$	min^{-1}
n	Fixed	1	-
K_m	Fitted	$1.50 \cdot 10^3$	-
$k_{m,\text{deg},b}$	Fitted	$5.24 \cdot 10^{-2}$	min^{-1}
$k_{u,\text{trans}}$	Fitted	$1.81 \cdot 10^1$	min^{-1}
δ	Fitted	$2.55 \cdot 10^{-2}$	min^{-1}
K_{CU}	Fitted	$2.51 \cdot 10^4$	-
β	Fitted	$4.08 \cdot 10^2$	min^{-1}
γ	Fitted	$9.11 \cdot 10^{-3}$	min^{-1}
α	Fitted	$3.75 \cdot 10^1$	min^{-1}
K_{IU}	Fitted	$9.93 \cdot 10^5$	-
K_{CI}	Fitted	$2.55 \cdot 10^3$	-
$k_{UPRS,\text{trans}}$	Fitted	2.36	min^{-1}
$k_{UPRS,0}$	Fitted	$2.14 \cdot 10^1$	min^{-1}
S_0	Fitted	5.00	g L^{-1}
n_S	Fitted	4.20	-
$K_{d,UPRS}$	Fitted	2.60	$(\text{g}^{n_S} \text{L}^{-n_S})$
$k_{UPRS,\text{deg},b}$	Fitted	$1.73 \cdot 10^{-3}$	min^{-1}
$Y_{X/S}$	Fitted	$3.89 \cdot 10^4$	$\text{L g}^{-1} \text{Events } \mu\text{L}^{-1} = 10^6 \text{Events g}^{-1}$
μ_{max}	Fitted	$1.47 \cdot 10^{-2}$	min^{-1}
U_{max}	Fitted	$4.42 \cdot 10^5$	-
K_S	Fitted	8.31	g L^{-1}
$k_{rib,max}$	Fitted	1.04	-
k_{sec}	Fixed	$1.00 \cdot 10^{-4}$	U (activity of amylase)

Table S6: Proportional-Integral control parameters as plotted in figure 4.12.

	UPRS _{target} (a.u.)	$K_p(-)$	$K_i(min^{-1})$
Operating point	600	$1.6 \cdot 10^{-3}$	$4.2 \cdot 10^{-5}$
	900	$3.0 \cdot 10^{-3}$	$7.7 \cdot 10^{-5}$
	1200	$8.5 \cdot 10^{-3}$	$9.2 \cdot 10^{-5}$
Ziegler-Nichols	600	$9.0 \cdot 10^{-3}$	$4.4 \cdot 10^{-5}$
	900	$10.8 \cdot 10^{-3}$	$5.3 \cdot 10^{-5}$
	1200	$11.7 \cdot 10^{-3}$	$5.7 \cdot 10^{-5}$
Predicted optimal operating point	600	$2.7 \cdot 10^{-3}$	$2.3 \cdot 10^{-5}$
	900	$3.6 \cdot 10^{-3}$	$4.1 \cdot 10^{-5}$
	1200	$3.6 \cdot 10^{-3}$	$4.9 \cdot 10^{-5}$

Bibliography

- [1] Zihe Liu, Keith E.J. Tyo, Jose L. Martinez, Dina Petranovic, and Jens Nielsen. Different Expression Systems for Production of Recombinant Proteins in *Saccharomyces cerevisiae*. *Biotechnology and Bioengineering*, 2012. doi: 10.1002/bit.24409.
- [2] Dirk Benzinger, Serguei Ovinnikov, and Mustafa Khammash. Synthetic gene networks recapitulate dynamic signal decoding and differential gene expression. *Cell Systems*, 13(5):353–364.e6, 2022. ISSN 24054720. doi: 10.1016/j.cels.2022.02.004. URL <https://doi.org/10.1016/j.cels.2022.02.004>.
- [3] Michael E. Lee, William C. DeLoache, Bernardo Cervantes, and John E. Dueber. A Highly Characterized Yeast Toolkit for Modular, Multipart Assembly. *ACS Synthetic Biology*, 4(9):975–986, 2015. ISSN 21615063. doi: 10.1021/sb500366v.

