LUND UNIVERSITY

LTH

FACULTY OF ENGINEERING

Master's Thesis

Novel tools for high-throughput genetic engineering, selection, and screening for improved Biotin production in *E. coli* cell factories

February 2023 – September 2023

Wibke Maria Vellguth

Examiner: Rajni Hatti-Kaul, Professor of Biotechnology, Lund University

Supervisor at Lund University: Nelida Leiva Eriksson

Supervisor at Biosyntia ApS: David Lennox-Hvenekilde

Degree Project in Biotechnology (30 ECTS)

Preface

This master's thesis was performed at the R&D department of Biosyntia ApS in Copenhagen, Denmark. The project spanned seven months, from February to August, and serves as a part of the Master of Science in Biotechnology program at Lund University. This thesis aims to develop an *E. coli* strain with enhanced biotin production capabilities to create an alternative to current unsustainable industrial production methods.

First and foremost, I would like to thank my supervisor at Biosyntia, David Lennox-Hvenekilde, for his continuous support, mentoring, and patience. I would also like to thank my supervisor, Nelida Leiva-Eriksson, for her helpful guidance and availability throughout this project. I also want to express my gratitude to every one of the employees at Biosyntia for their advice, guidance, and friendship. Special thanks go out to Silvia Peri for answering all my questions and guiding me through most of the experiments, Nils Myling-Petersen for advising me at the beginning of my project, and Jonathan Thompson Wood for being a true friend and great company.

Finally, I would like to dedicate this project to my late grandfather, Gerhard Kittel, who ignited my passion for science and is dearly missed.

"The pain of missing you is a beautiful reminder of the joy of loving you." – Dean Jackson

Abstract

Biotin is an essential co-factor in various metabolic processes and is crucial for the growth and maintenance of living organisms. The market need for biotin is multifaceted and driven by its importance in human health, beauty and cosmetics applications, pharmaceuticals, animal nutrition, and other industries. Current methods of industrial biotin production are unsustainable and generate environmentally hazardous waste, fueling the demand for new, greener strategies. The production of biotin through microbial fermentation has thus become an attractive alternative. Biotin biosynthesis is a complex and energy-intensive process that requires several enzymatic reactions, and biotin titers in microbial production have not reached economically viable levels. Biotin Synthase (BioB), the enzyme that catalyzes the last step of this process in *E. coli,* the conversion of dethiobiotin to biotin, has been identified as the bottleneck for this pathway. Overexpression of the enzyme generates oxidative stress and inhibits growth, though the exact mechanism has not yet been elucidated. This thesis employed a multiplexed engineering approach via high-throughput genetic engineering, selection, and screening methods to explore whether creating *E. coli* strains more resistant to oxidative stress would lead to higher biotin production capabilities. Using the expression of a DNA methylase to introduce controlled genetic mutations in Biosyntia's proprietary biotin-producing *E. coli* strains and selecting strains with improved resistance to oxidative stress yielded a selection of strains with increased resistance, which were screened for their biotin production. The improved resistance to oxidative stress and BioB induction in the strains did not lead to higher biotin production levels.

Keywords: *biotin, BioB, biotin synthase, oxidative stress, E. coli*

Popular Science Summary

The global need for biotin in the food, feed, and cosmetics industries is increasing steadily. Because current production methods use petrochemicals as precursors for synthesis, they are unsustainable and generate environmental waste. This has fueled the demand for alternative, more sustainable strategies. Biotin production via microbial fermentation could be a greener alternative but has not reached economically viable levels and cannot compete with current production methods. *E. coli* can naturally produce biotin but does so in minute amounts. This thesis aims to improve biotin production levels in biotin-producing *E. coli strains* to contribute towards a greener, more sustainable process.

A bottleneck in the biosynthetic pathway, which has limited production levels, has been identified in previous research: the enzyme BioB, which catalyzes the last step of biotin synthesis. This enzyme creates oxidative stress in the cells and thereby inhibits cell growth. It was hypothesized that creating strains more resistant to oxidative stress and higher levels of intracellular BioB could lead to higher biotin production capabilities. By generating a controlled level of genetic diversity or introducing new genes into the strains and then selecting those with higher oxidative stress resistance, this thesis aims to find a strain that can produce higher levels of biotin. By nature, the randomness of the mutagenesis process generates a high number of strains that need to be tested for their biotin production. In this thesis, the number of strains that need to be tested is systematically reduced by selecting strains with a higher oxidative stress resistance by exposing them to various stressor concentrations. Only the fittest, most resistant strains survive this and are subsequently tested for their biotin-production capabilities.

The work revealed that strains with higher resistance to oxidative stress and BioB production did not produce higher levels of biotin, indicating other limiting factors in the biosynthetic pathway. However, further investigations are needed to draw definitive conclusions, as not all the data generated during this project is reliable due to measurement issues.

Table of Contents

Abbreviations

Abbreviation	Definition
ACP	acyl carrier protein
Amp, A	ampicillin
BioB	biotin synthase
CO	cobalt
cPCR	colony polymerase chain reaction
DTB	Dethiobiotin
EDTA	Ethylenediaminetetraacetic acid
HABA	4'-hydroxyazobenzene-2-carboxylic acid
IPTG	Isopropyl β- d-1-thiogalactopyranoside
isc	iron-sulfur cluster systems
IscR	iron-sulfur cluster regulator
Kan, K	kanamycin
LB	lysogeny broth
NEB	New England Biolabs
OD	optical density
PQ	paraquat
Rif, R	rifampicin
ROS	reactive oxygen species
SAM	S-Adenosyl methionine
SOC	Super Optimal broth with Catabolite repression
suf	sulfur mobilization

1. Introduction

1.1 Industrial Biotechnology and Metabolic Engineering

Disregarding microbial food and beverage production, which dates to the seventh millennium BC (McGovern, 2004), the use of microorganisms in the industrial production of value-added compounds began more recently, at the start of the 20th century (Soetaert & Vandamme, 2010). At the time, solvents like ethanol, acetone, and butanol were first produced using large-scale yeast and bacterial fermentations, and citric acid and penicillin were first produced using filamentous fungi.

Alexander Flemming's discovery of penicillin marked a significant moment in the history of modern biotechnology. Despite Flemming discovering penicillin as early as 1929, it remained a "laboratory curiosity" until the arrival of the Second World War, when the study of penicillin became essential to the war efforts as wounded soldiers were dying on the battlefield from bacterial infections (Soetaert & Vandamme, 2010). Although the inhibitory effects of penicillin on bacterial colonies were clearly described, and early clinical trials were successful, the large-scale production of penicillin proved much more difficult. The fungal strain initially identified by Flemming produced only trace amounts of penicillin, but genetic manipulation made significant strides in production capabilities and led to the emergence of a new technology known as "strain improvement." Since early cell factories such as *Penicillium chrysogenum* (Penicillin), *Aspergillus niger* (citric acid), and *Clostridium acetobutylicum* (Acetone, Butanol, Ethanol) were native producers of the compounds of interest, the focus of these early genetic studies was the enhancement of production titers through random mutagenesis using X-ray or UV treatments. However, after Cohen et al. (1973) demonstrated that *in vitro*-created DNA constructs were biologically functional after transformation, the use of recombinant constructs became popular and led to the establishment of a new biopharmaceutical industry.

In the years following this discovery, recombinant DNA constructs were used for the heterologous production of proteins like somatostatin, human insulin, and erythropoietin. While it was discovered that the production of a complex molecule like insulin could be achieved through the overexpression of a single gene, the overproduction of a compound like ethanol required the modulation of a collection of genes in the ethanol production pathway (Woolston et al., 2013). This led to the establishment of the field of metabolic engineering, which focuses on the engineering of complete metabolic networks and pathways. A central task in this field is the directed improvement of product formation through quantifying and modulating fluxes within the cell. Its express purpose is optimizing bio-fermentative processes for cost-effective commercial production.

1.2 Biotin and its Applications

Biotin (vitamin B7) is an essential vitamin that serves as a co-factor of five mammalian carboxylases (Dasgupta, 2019). These enzymes are crucial in biological processes like fatty acid synthesis, gluconeogenesis, and amino acid metabolism. Vitamins, like Biotin, are essential for all living organisms (Wang et al., 2021). Because the human body cannot synthesize most vitamins independently, it must obtain all necessary vitamins from food (Gironés-Vilaplana et al., 2017). Often, nutritional requirements are not met, which requires supplementation. This need has created a market for vitamins as supplements in the food, feed, and cosmetics industries. However, current industrial production methods are unsustainable, employing non-renewable sources like petrochemicals as

precursors and generating environmentally hazardous waste (Acevedo-Rocha et al., 2019). The industrial production of biotin, for example, requires a multi-step chemical synthesis (Eggersdorfer et al., 2012). Nevertheless, despite the rapid development of organic synthesis methods, environmentally friendly protocols are still lacking. Thus, the desire for a greener, more sustainable production process has fueled the demand for new strategies. Since many microorganisms can synthesize vitamins naturally (Burkholder & Mcveigh, 1942), vitamin production through microbial fermentation has become an attractive alternative to chemical synthesis.

Like other microbes, *E. coli* can synthesize biotin but does so in minute amounts (Brown & Kamogawa, 1991). Consequently, its productivity must be improved to enable a cost-competitive process. However, forward metabolic engineering has yet to yield satisfactory results. This is due to the complex metabolic pathway (Figure 1) facilitating biotin synthesis *in vivo.*

1.3 The Biotin Biosynthetic Pathway and its Challenges for Bioproduction

1.3.1 The Biotin Biosynthetic Pathway in *E. coli*

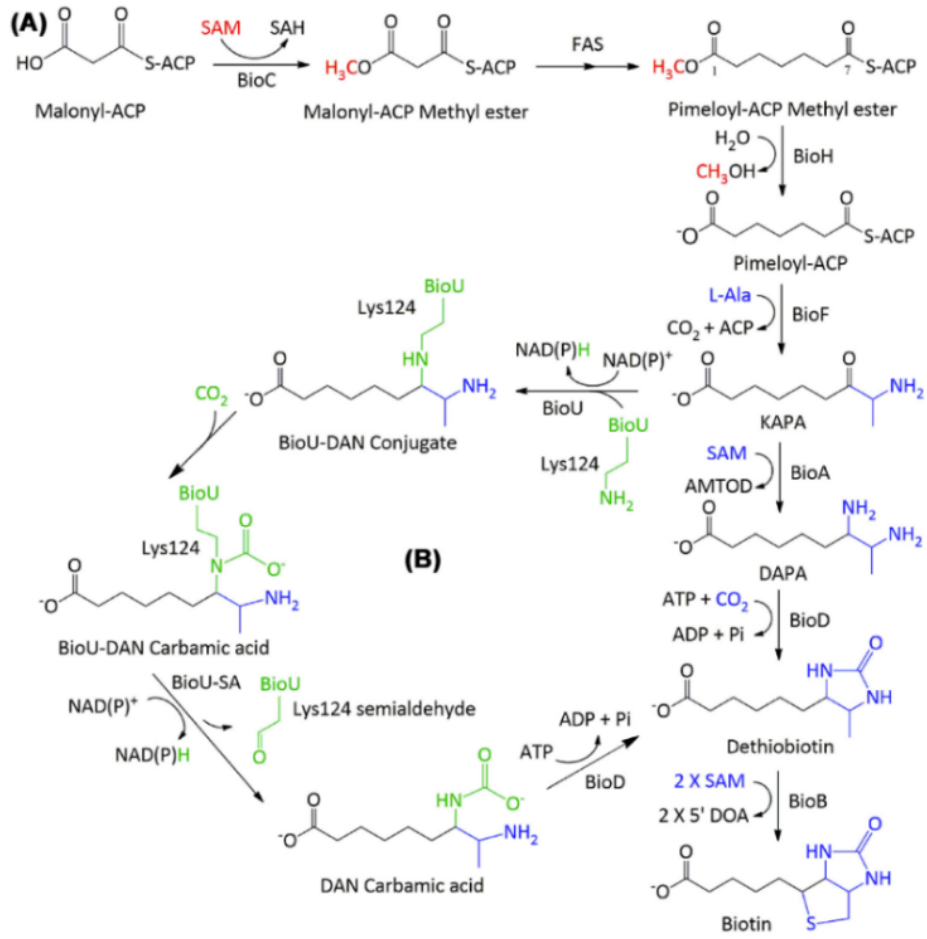


Figure 1-(A) The E. coli biotin biosynthetic pathway and (B) the proposed reaction mechanism of the new biotin biosynthetic enzyme BioU that bypasses BioA. (Sirithanakorn & Cronan, 2021)

There are two ways *E. coli* can acquire biotin. It either scavenges from the environment or produces it *de novo* (Sirithanakorn & Cronan, 2021). Because biotin synthesis requires S-adenosyl methionine (SAM) at several points of the pathway, the process is metabolically expensive; hence, biotin synthesis

is shut down when biotin is exogenously available. The *bio* genes that regulate the biosynthetic pathway are organized into operons and thus co-transcribed, allowing the enzymes to be produced in stoichiometric quantities. Bacteria generally require biotin in only modest quantities for growth, but the enzymes, especially the ones catalyzing the last part of the pathway, are poor catalysts and must be synthesized in sufficient amounts.

The biosynthetic pathway (Figure 1) for biotin in *E. coli,* as described in Sirithanakorn & Cronan (2021), consists of two stages: The generation of the pimelate moiety and the assembly of the biotin rings. In the first stage, the free carboxyl group of the malonyl-acyl carrier protein (ACP) is methylated by BioC, an S-adenosyl methionine (SAM)-dependent methyltransferase. This methylation enables the fatty acid biosynthetic enzymes to use the malonyl-ACP methyl ester as a substrate. Following methylation, the malonyl-ACP methyl ester undergoes two cycles of fatty acid synthesis. The first cycle of fatty acid chain elongation produces C5 carboxylate, and the second cycle elongates this to the C7 dicarboxylate to yield pimeloyl-ACP methyl ester. BioH then removes the methyl group from the methyl ester, yielding pimeolyl-ACP. BioF subsequently condenses pimeloyl-ACP with L-alanine, yielding 8-Amino-7-oxononanoate (KAPA) and initiating stage two of the pathway. BioA then catalyzes the transamination of the two NH_2 groups to form 7,8-diamino pelargonic acid (DAPA). This is followed by the ATP-dependent insertion of CO_2 between the N7 and N8 nitrogen atoms of DAPA, forming the ureido ring of biotin (Sirithanakorn & Cronan, 2021). The final step is the catalysis of the resulting dethiobiotin (DTB) to biotin by BioB, an iron-sulfur enzyme, which inserts sulfur obtained from its own $[2Fe\text{-}2S]^{2+}$ (iron-sulfur) cluster between the C6 methylene and the C9 methyl groups of DTB, forming the tetrathiopane ring of biotin (Dunn, 2019).

1.3.2 Challenges in the Biotin Pathway

The conversion of DTB to biotin, catalyzed by the biotin synthase (BioB), is considered the main bottleneck in this pathway, which is evidenced by the accumulation of DTB in biotin cell factories (Bali et al., 2020). The catalysis of DTB to biotin requires two SAM equivalents and sulfur donated from the $[2Fe\text{-}2S]^{2+}$ cluster, one of two iron-sulfur clusters present in the active enzyme (Cramer & Jarrett, 2018). The mechanism of sulfur donation from the iron-sulfur cluster requires cluster regeneration after each turnover to regenerate an active enzyme form, resulting in a slow catalysis rate. BioB is degraded by a proteolysis mechanism that sequentially cleaves small fragments from the C-terminus but is seemingly resistant to degradation and capable of multiple turnovers in a high-iron environment, which favors iron-sulfur cluster regeneration (Reyda et al., 2008).

The Fe-S cluster's destruction during the catalysis may lead to the release of free iron, which can cause oxidative stress by producing highly reactive oxygen species *in vivo* (Py & Barras, 2010). Usually, this effect is negligible in the cell, as the expression of BioB is extremely low due to the minute amounts of biotin needed for cell growth and maintenance. Due to the toxic potential of free iron in the cells, pathways for producing iron-sulfur clusters are tightly regulated to minimize toxic effects. In *E. coli*, iron-sulfur clusters are primarily produced through sulfur mobilization (*suf*) and iron-sulfur cluster systems (*isc*). The *isc* pathway is part of the housekeeping system and supplies iron-sulfur clusters to cluster proteins. The *suf* pathway is likely induced under stress conditions, such as during oxidative stress or iron limitation (Py & Barras, 2010).

Adding to the already complex issue of iron-sulfur cluster regeneration, BioB has been found to inhibit growth independent of its biotin-forming activity (Ifuku et al., 1995). Multiple hypotheses have been made regarding the cause of the observed toxicity, including the generation of reactive oxygen species (via iron-sulfur clusters), insufficient iron-sulfur cluster supply, or limitations in iron availability (Reyda et al., 2008). Ifuku et al. (1995) also postulated that growth inhibition has something to do with a specific conformation of the enzyme.

Another factor is the high demand for SAM as a co-factor for BioB, coupled with the inhibiting effect of SAM byproducts on several enzymes in the pathway, including BioB (Parveen & Cornell, 2011).

1.3.3 Limitations in Traditional Approaches

Several factors contribute to the complexity of the DTB to biotin conversion step. Most importantly, the unresolved question of the BioB toxicity mechanism and the iron-sulfur cluster regeneration issue. Previous efforts to improve biotin production involved the use of random mutagenesis and antimetabolites (Streit & Entcheva, 2003). Three strategies have been applied: cloning or overproduction of *bio* genes, selection for improved biotin production through chemical mutagenesis methods, and a combination of the two.

Attempts to improve iron-sulfur cluster generation have included overexpression of the *isc* operon and the plasmid-based overexpression of *isc* and *suf*. This led to increased biotin production and indicated that iron-sulfur cluster supply is a limiting factor (Bali et al., 2020). However, plasmid overexpression of the *isc* and *suf* operon was hypothesized to burden the cells, constituting a significant drawback for cell factory scale-up. Using BioB overexpression toxicity as the foundation for growth selection assays, Bali et al. (2020) were able to identify single mutations in the global IscR regulator (involved in iron-sulfur cluster biogenesis) that improve cellular tolerance towards BioB overexpression increasing DTB to biotin catalysis more than 2.2-fold.

Despite these efforts, biotin production titers have yet to reach economically viable levels. With rational and forward engineering strategies reaching their limits, other avenues should be explored. Rational, forward metabolic engineering is, by and large, limited by the current understanding of the metabolism one aims to engineer. It allows for the systematic building of pathways, ensuring predictability and robustness. Fine-tuning at each step yields well-characterized strains, but with the understanding of cellular metabolisms not yet at a point where they can be modeled reliably, it remains a challenge. Multiplexed engineering, on the other hand, offers distinct opportunities in cell factory development, as it involves the simultaneous manipulation of multiple genetic elements within an organism, enabling a more rapid optimization, accelerating strain development, and enhancing productivity. However, multiplexed engineering also has its limitations. Simultaneously manipulating several genetic elements can lead to unintended interactions, and its outcomes are hard to predict. This results in a lack of understanding of the contributions of each genetic change to the resulting phenotype. The concomitant changes in several genetic elements also increase the likelihood of off-target mutations, genetic instability, and reduced cell viability. Balancing these limitations with the benefits of accelerated strain development is crucial when deciding on the optimal strategy. Integrating both engineering strategies can optimize cell factories by leveraging their respective strengths.

1.4 Objectives of the Thesis

This project aims to address the problems that arise in forward engineering by using a multiplexed approach to engineering the biotin production strain. By introducing controlled levels of genetic diversity in Biosyntia's proprietary biotin-producing *E. coli* strains and using high-throughput selection and screening methods, this thesis aims to improve biotin production titers by developing a strain more resistant to the BioB-induced cellular stress. It is hypothesized that a strain more resistant to higher levels of intracellular BioB would produce higher levels of biotin.

There are three building blocks to this thesis. The first is the generation of genomic diversity, the second is the selection for improved phenotypes, and the third and final building block is the screening of potential hits. The generation of genetic diversity is split into two distinct approaches. The first is the development of genomic diversity, and the second is the development of plasmid diversity. Genomic diversity is created by amplifying the natural occurrence of mutations brought forth by the DNA polymerase during replication by expressing a DNA methylase, which masks the newly-synthesized DNA from the endogenous proofreading mechanisms. This increases the rate of point mutations in the genome, enhancing genetic diversity, while also controlling the mutagenesis rate, to minimize the number of off-target mutations. This system is then used to create libraries of controlled genetic diversity for the biotin-producing strains. Plasmid diversity is created by generating genomic DNA libraries using *E. coli* and *S. cerevisiae* strains. For this purpose, gDNA from these strains is fractured into a length between 1000 and 5000bp, large enough to cover the average ORF, and cloned into a plasmid backbone compatible with *E. coli* strains. This feasibly introduces possibly beneficial genes into the strain, which might lead to increased resistance to BioB-induced oxidative stress and, possibly, higher biotin production capabilities. Using DNA fragments from *E. coli* also enables testing of overexpression of all native genes without selection bias.

The next step is the selection for improved phenotypes. Previous work at Biosyntia has indicated that the expression of BioB, the bottleneck enzyme, causes oxidative stress to the cells, inhibiting growth. Favorable phenotypes are selected by imposing oxidative stress on the mutant populations created in the previous steps. Oxidative stress is induced by exposing the populations to media additives such as copper, hydrogen peroxide, and paraquat or inducing BioB expression. This selects for mutants with genotypes that confer the desired resistance to oxidative stress.

The last step is screening the selected strains for improved biotin-production titers using a robot-assisted microtiter-plate screening pipeline to characterize growth and production capabilities. Biotin production is quantified using an auxotrophic bioassay strain, that grows only in the presence of exogenously available biotin.

2. Material and Methods

2.1 E. coli diversity tool development

2.1.1 Creating Strain for MMR Inactivation Experiment

2.1.1.1 Plasmid-prep from E. Coli BS04704

5 mL of LB+A (lysogeny broth and ampicillin, ref. Abbreviations) medium were inoculated with a strain carrying the pBS2215 plasmid (BS04704) from a cryostock and grown overnight (ref. Strain List and Plasmid List in Appendix). This plasmid carries the *dam* gene which encodes the DNA methylase used to create genetic diversity. The plasmid was purified from the strain using the Nucleo Spin Plasmid EasyPure DNA Purification kit (Macherey-Nagel).

2.1.1.2 Making Electrocompetent Cells

An overnight culture of the strain (BS08072) was grown at 37°C in 3 mL 2xYT medium in culture tubes. 20 mL of 2xYT medium were inoculated with 100 µL from the overnight culture the next day and incubated at 37°C. When the culture reached an OD of ~ 0.6 (exponential phase), the cells were centrifuged (4000g, 5 min), the supernatant was discarded, and the cells were resuspended in 20 mL of ice-cold sterile water. The cells were centrifuged again and resuspended in 20 mL of ice-cold sterile water a second time. The cells were centrifuged a third time and resuspended in 1 mL of ice-cold water.

2.1.1.3 Transformation

100 µL of electrocompetent cells and 10 µL of plasmid DNA (from pBS2215) were transferred into a sterile electroporation cuvette (Fisherbrand, #FB101, 1 mm gap, Fisher Scientific). The cuvette was placed in an electroporation system (ECM 399, BTX Harvard apparatus), and 1800 V was applied to the cells. 1 mL of pre-warmed (37°C) SOC medium was added to the cuvette immediately after. The cuvette contents were transferred to an Eppendorf tube and incubated for 1 h at 37°C and 275 rpm. After the incubation, the cells were streaked out on an LB agar plate with ampicillin (to select positive transformants) and incubated for 48h; the resulting strain was named BS07701.

2.1.1.4 Creation of Cryo-Stock

A colony from strain BS07701 was picked from the agar plate, and 2 mL of LB medium was inoculated. The culture was grown overnight. The culture was centrifuged (4000g, 5 minutes) and washed twice in mMOPS medium without a carbon source. The culture was then resuspended in half of the original culture volume. 600 µL of culture volume and 400 µL of 50% glycerol solution were mixed. The mix was transferred to a cryo-tube and stored in a -80°C freezer. This was done in duplicates.

2.1.2 Generating Strain Libraries and Data for MMR Inactivation Model

Day 1

Three cultures from the BS06757 (or BS07701) culture were inoculated in 3 mL 2xYT+A medium and incubated overnight. These cultures represent the S0 cultures. 40 LB+K+A+R (lysogeny broth with kanamycin, ampicillin, and rifampicin) plates were prepared.

Day 2

The ODs of the overnight culture were measured, and production was verified by spinning down 1 mL of culture, mixing 100 µL HABA solution with 100 µL of supernatant, and observing a color change from orange to yellow if biotin is present. 100 µL of each of the three overnight cultures were inoculated into 5 mL 2xYT+A+K+5 mM cAMP medium to a start OD_{620} of 0.01, and production was induced with arabinose to a final concentration of 20 g/L. A replicate of the first of the three cultures was induced into the same medium but not induced to act as a control. These cultures represent the S1 cultures. The cultures were incubated for an hour at 37°C and 280 rpm, after which the start OD was measured. The cultures were subsequently incubated in the same conditions overnight. The S0 cultures were also left to incubate overnight.

Day 3

The ODs of all cultures were measured, and 200 µL of the S1 culture were plated onto LB+A+R agar. 100 µL of overnight culture were transferred into fresh 2xYT+A+K+5mM cAMP medium and incubated at 37°C and 280 rpm for an hour. These represent the S2 cultures. Their OD was measured before adding arabinose. Production of the S0 cultures was verified using HABA solution as described above. S0 and S1 cultures were washed twice using mMOPS medium without glucose and resuspended in half of the original culture volume. The same volume of 50% glycerol solution was added to the cultures. Five 1 mL cryo-stocks were prepared for each culture and stored at -80°C. The day 3 protocol was repeated until an S7 culture was reached. After two days of incubation at 37°C, the colonies on the LB+K+A+R plates were counted. Colonies S1, S3, S5, and S7 were sent to an external facility for NGS sequencing.

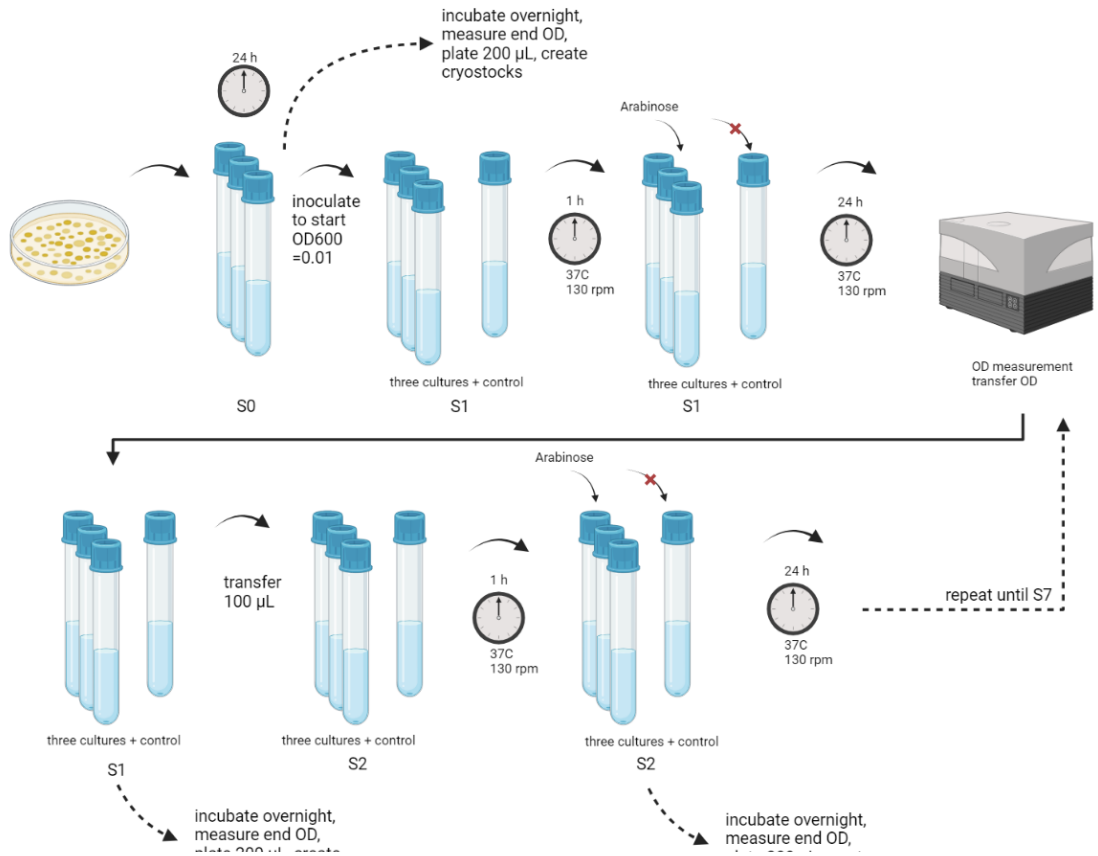


Figure 2- Generating strain libraries and data for MMR inactivation model – workflow created in BioRender.

2.1.3 Creating gDNA Libraries

2.1.3.1 Extraction of gDNA

a) Saccharomyces cerevisiae gDNA

A *S. cerevisiae* strain (BS07084) was inoculated into 5 mL YPD medium from a cryo-stock and grown overnight at 30°C and 300 rpm. The gDNA was subsequently purified using the YeaStar Genomic DNA Kit (Zymo Research).

b) Escherichia coli gDNA

An *E. coli* strain (BS04755) was inoculated into 5 mL 2xYT medium from a cryo-stock and grown overnight at 37°C and 280 rpm. The genomic DNA was purified using the DNeasy Kit (Qiagen).

2.1.3.2 Plasmid backbone digestion

Table 1- Master mix for plasmid backbone digestion and phosphorylation

Component	**Volume**
Plasmid pBS1897 (backbone for library, ref. Plasmid List)	1 µg
Anza 10X Buffer	5 µL
Eco32I	5 µL
Anza Phosphatase	2.5 µL
H2O	Up to 50µL

The plasmid backbone includes a promoter in front of a restriction site so inserts can be expressed. The pBS1897 plasmid has an inducible pTet promotor followed by a ribosomal binding site. The components listed in Table 1 were mixed in a PCR tube and incubated at 37°C for 3 hours. 5 µL of the reaction mix were subsequently run on a 1% agarose gel to verify the size of the backbone. The rest of the reaction mix was purified using the E.Z.N.A Cycle Pure Kit from OMEGA bio-tek and eluted in 30 µL water.

2.1.3.3 Fractionation protocol (adapted from New England Biolabs) for gDNA

1. NEBNext dsDNA Fragmentase was vortexed for 3 seconds and placed on ice.
2. The components in Table 2 were mixed in a sterile PCR tube and vortexed. The genomic DNA used in this fractionation was purified in protocols 2.1.3.1 a) and b).

Table 2 - Fractionation protocol reaction components.

Component	**Volume**
Genomic DNA (*E. coli* or *S. cerevisiae*) (5ng-3µg)	1-16 µL
10X Fragmentase Reaction Buffer v2	2 µL
Sterile Water	variable
Final Volume	18 µL

3. 2.0 µL dsDNA Fragmentase was added to the reaction and vortexed for 3 seconds.
4. The mixture was incubated at 37°C for the recommended times to generate the desired fragment size. A time course study determined the exact incubation time for a given sample type. Times between 2.5 and 5 mins were best for *E. coli* (BS04755) for S. cerevisiae (BS07084).
5. 5 µL of 0.5 M EDTA was added to stop the reaction.
6. The reaction mix was run on a 1% agarose gel. Sections between 1 and 5kbp were cut from the gel and purified using the Monarch DNA Gel Extraction Kit (New England Biolabs).

2.1.3.4 gDNA End Repair Procedure

To enhance ligation efficiency, the overhangs from the fragmented DNA are extended to produce blunt and phosphorylated ends. The components listed in Table 3 were mixed in a sterile microcentrifuge tube and placed on ice. The mixture was incubated in a thermal cycler for 20 minutes at 20°C and purified using the E.Z.N.A Cycle Pure Kit (OMEGA bio-tek).

Table 3 – gDNA End Repair Reaction Mix

Component	**Volume**
gDNA fragments	1.5 µg
10X End Repair Reaction Mix	5 µL
End Repair Enzyme Mix	2.5 µL
Water (sterile)	to 50µL
Final volume	50 µL

2.1.3.5 Ligation of Backbone and DNA Fragments

An NEB highly concentrated T4 Ligase was used for blunt-end ligation of the inserts (1000bp to 5000bp gDNA fragments from *E. coli* (BS04755) and *S. cerevisiae* (BS07084)) and the backbone (pBS1897). The mixture described in Table 4 was incubated at 16°C for 12 h and then heat-inactivated at 70°C for 5 minutes. The reaction mix was then purified using the E.Z.N.A Cycle Pure Kit (OMEGA bio-tek) and eluted in 15 µL.

Table 4 - Ligation Mixture

Component	**Volume**	**Comment**
T4 ligase buffer	2 µL	Thaw at 37°C and smell sulfur to ensure ATP is breaking down.
PEG4000	2 µL	
Backbone (pBS1897)	100 ng	
gDNA fragments	200 ng	
T4 ligase	1µL	Vortex briefly before use.
dH20	to 20µL	

2.1.3.6 Transformation of Construct

The *E. coli* strain chosen for transformation was BS08072. Electrocompetent cells were produced following the protocol *2.1.1.2 Making Electrocompetent Cells*. The purified ligation mix was transformed into the electrocompetent cells following the *2.1.1.3 Transformation* protocol. The empty pBS1897 plasmid was used as a positive control for the transformation. After incubation of the transformation mixture, the mixture was diluted up to 10^{-4}, dilutions 10^{-3} and 10^{-4} were plated on LB agar plates with kanamycin and the success of the transformation was verified with a colony PCR (protocol in Appendix) of resulting colonies.

2.2 Functional Selection Development

2.2.1 Determination of Stressor Levels

Gradients for paraquat (PQ), cobalt (CO), hydrogen peroxide (H_2O_2), and Isopropyl β-D-1-thiogalactopyranoside (IPTG) (Table 5) were created to determine the range for inhibitory concentrations. mMOPS+DTB+A plates with varying concentrations of stressors were prepared, and 10^8 cells were streaked out on the plates, with a subsequent incubation at 37°C for 48 hours.

Table 5 – Concentrations of selection stressors

Stressor	Plate 1	Plate 2	Plate 3	Plate 4	Plate 5	Plate 6	Unit
PQ	100	150	200	250	300	350	µM
Cobalt	1	2	3	4	5	-	mM
H2O2	0.1	0.5	1	1.5	2	-	mM
IPTG	50	100	500	1000	1500		µM

2.2.2 Generation of 2D gradient

A 2D gradient was created using the determined paraquat and hydrogen peroxide stressor levels (Figure 3 and Table 6). The other stressors were excluded from further experiments due to time constraints. Three sets of nine agar plates with mMOPS+DTB+A and varying stressor concentrations were made.

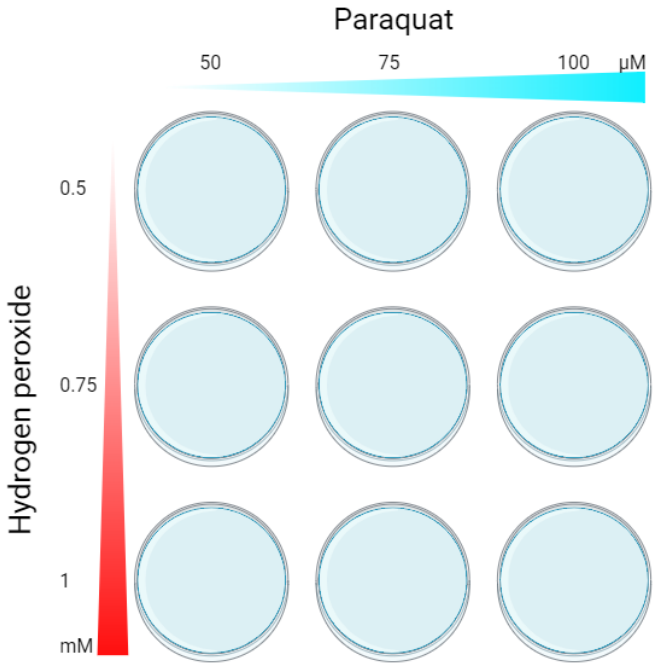


Figure 3 - Gradient illustration made with BioRender

Table 6 – Stressor concentrations of hydrogen peroxide and paraquat

	50 µM PQ	75 µM PQ	100 µM PQ
0.5 mM H2O2	0.5 mM H2O2 50 µM PQ	0.5 mM H2O2 75 µM PQ	0.5 mM H2O2 100 µM PQ
0.75 mM H2O2	0.75 mM H2O2 50 µM PQ	0.75 mM H2O2 75 µM PQ	0.75 mM H2O2 100 µM PQ
1.0 mM H2O2	1.0 mM H2O2 50 µM PQ	1.0 mM H2O2 75 µM PQ	1.0 mM H2O2 100 µM PQ

2.2.3 Selection for Improved Phenotypes

Day 1

A pre-culture of BS07701 was inoculated into 3 mL of 2xYT+A medium and incubated at 37°C and 280 rpm overnight. Twenty-seven selection gradient agar plates were made (Table 6) with mMOPS+DTB+A medium and different combinations of stressor concentrations, as well as six control plates without stressors.

Day 2

The cells were washed twice with mMOPS medium without a carbon source, then 10^8 cells were plated on mMOPS+DTB+A plates with a stressor gradient and control plates without stressors. These plates represent the uninduced (Round 1) plates. The plates were placed in an incubator at 37°C for 48 h. 100 µL of the overnight culture were transferred into 5 mL 2xYT+A+5 mM cAMP medium and induced with 20% arabinose. The culture was placed in the shaker and incubated overnight at 37°C and 280 rpm.

Day 3

The cells from the overnight culture were washed twice with mMOPS medium without a carbon source, and 10^8 cells were plated on the stressor gradient and control plates. These represent the once-induced (Round 2) plates. The plates were placed in an incubator at 37°C for 48 h. 100 µL of the overnight culture was transferred to 5 mL 2xYT+A+5 mM cAMP medium and induced with 20% arabinose. The culture was placed in the shaker and incubated overnight at 37°C and 280 rpm.

Day 4

The cells from the overnight culture were washed twice with mMOPS medium without a carbon source, and 10^8 cells were plated on the stressor gradient and control plates. These represent the twice-induced (Round 3) plates. The plates were placed in an incubator at 37°C for 48 h.

All agar plates were removed from the incubator after 48 h of incubation and photographed. The largest colonies growing in the highest concentrations were picked and streaked out on LB+A agar plates and incubated at 37°C.

2.3 High-throughput Screening

2.3.1 Growth Assay with Stressors

Day 1

Colonies were picked from the LB+A agar plates, which were streaked out from the largest colonies from Round 1 and Round 3 of the selection plates, inoculated into 3 mL mMOPS+DTB+A medium and grown overnight at 37°C and 280 rpm. An additional overnight culture of the background strain BS07701 was inoculated into 3 mL mMOPS+DTB+A medium from cryostock.

Day 2

2 µL of overnight culture were inoculated into 200 µL of mMOPS+DTB+A medium with nine different stressor concentrations on a microtiter plate in duplicates for the colony strains and in triplicates for the control strain. The microtiter plates were covered with a seal and incubated in a spectrophotometer for 24 h and OD_{620} measurements were taken every 20 minutes.

Table 7 -Stressor concentrations for microtiter plates

Label for concentration	**11**	**12**	**13**	**21**	**22**	**23**	**31**	**32**	**33**
PQ[µM]	0	0	0	50	50	50	100	100	100
H2O2[mM]	0	0.5	1	0	0.5	1	0	0.5	1

2.3.2 Biotin Production Assay with IPTG

2.3.2.1 Sample Preparation

Three single colonies (biological triplicates) were picked from each of the LB+A agar plates and inoculated into 200 µL mMOPS+DTB+A medium on a microtiter plate. The microtiter plate was placed in an incubator for 24 h at 37°C and 275 rpm. See layout in Figure 4.

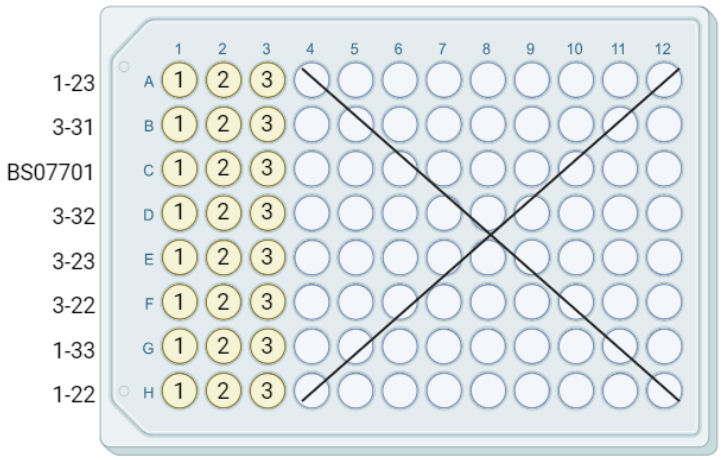


Figure 4 - Pre-culture layout for cultures with biological triplicates. All cultures except BS07701, which acts as a control, stem from the selection with stressors. The crossed-out wells were left empty. Layout created in BioRender.

2.3.2.2 Inoculation of Production Plates

5 µL of the precultures were inoculated into 500 µL of mMOPS+DTB+A medium in 96-well production plates. The plates had a selection of different concentrations of IPTG in the medium, ranging from 0 µM to 100 µM (Figure 5). IPTG induces bioB production. The *bioB* gene is under a T5-lacO promoter (ref. Strain List). Each concentration was prepared in triplicates. The contents of the wells were mixed thoroughly via aspiration. 200 µL were taken from each well and transferred into 96-well microtiter

plates for use in the growth assay. The production plates were subsequently covered with an aluminum seal and incubated for 48 h at 37°C and 275 rpm.

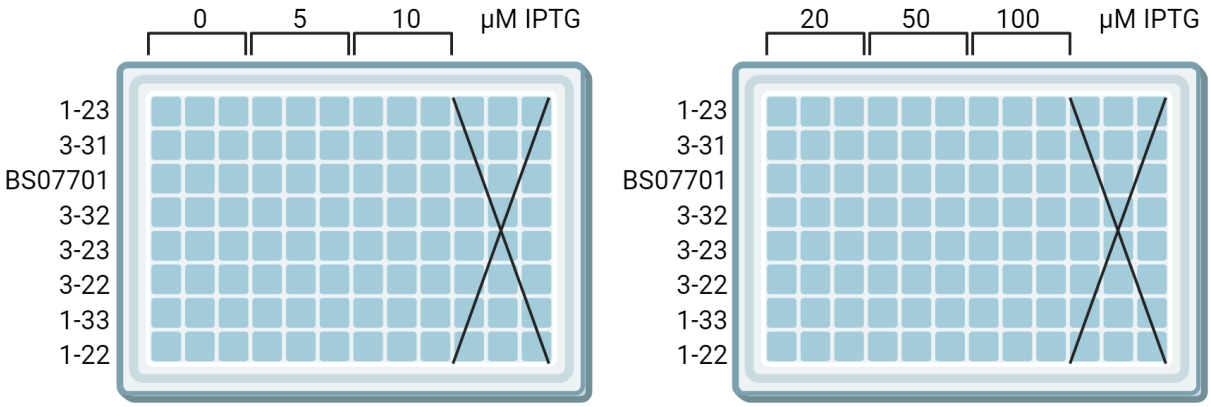


Figure 5 - Layout for production plates. Crossed-out wells are empty. Layout created in BioRender.

2.3.2.3 Harvest

A new 96-well plate was prepared to measure the end OD of the production plates. 20 µL of culture was mixed with 180 µL of sterile water to obtain a 1:10 dilution, and the OD_{620} was measured with the plate reader. The remaining cultures in the production plates were centrifuged at 4000g for 15 minutes. After centrifugation, 200 µL of supernatant was carefully harvested and moved to a 96-well plate.

2.3.2.4 Biotin Bioassay

Day 1

Ten 96-well microtiter plates were labeled for the dilutions and the bioassay. Six plates for the dilution were labeled according to Figure 6, Column 3, and four plates for the bioassay were labeled according to Figure 6, Column 6. Sterilized dH20 was added to the dilution plates using an automated multichannel pipettor according to the volumes specified in column 4. 52 mL of mMOPS+zeocin (1:1000 dilution to mMOPS volume) for the bioassay plates. 135 µL of volume were added in each well. A bioassay strain (BS07210) bullet was taken from the freezer and thawed at room temperature. The bioassay strain is an auxotrophic strain that grows only when biotin is present. Growth of the bioassay strain is indicative of biotin production by the assayed strains. Biotin in exported from the cells of the producing strain, enabling growth of the auxotrophic strain. To prepare the bioassay solution a dilution of mMOPS+zeocin was inoculated to an OD of 0.005-0.01 from the bioassay strain bullet. 135 µL of

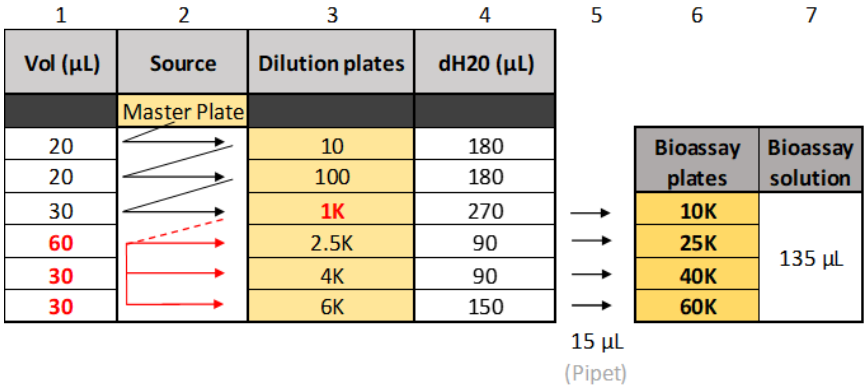


Figure 6 – Bioassay preparation guide

bioassay solution were inoculated into each well using the automated multichannel pipettor. To prepare the bioassay master plate, the supernatant samples taken from the production plates were added to a 96-well plate.

The microtiter plates containing biotin standards (preparation refer to Appendix) were taken from the -20°C freezer and thawed at room temperature. The contents of the plates (200 µL) were transferred to columns 9-12 of the master plate.

The contents of the bioassay master plate and the standards were mixed thoroughly using the multichannel pipettor. The multichannel pipettor was subsequently used to make dilutions of the supernatants in the pre-filled dilution plates according to the volumes listed in columns 1 and 2, and each plate was mixed thoroughly. The multichannel pipettor was then used to transfer 15 µL from the supernatant dilutions to the bioassay plates. The bioassay plates were then covered with breathable seals and loaded into the shaker to incubate at 37°C and 275 rpm overnight.

Day 2

The seals were removed from the plates, and the OD_{620} was measured. The OD measurement can be converted into cell growth which is directly proportional to the biotin production and can be used to quantify the biotin production when considering the dilution level and growth of 21 included biotin standard concentrations. The cell growth of the auxotrophic strain in the standards with known biotin concentrations is used to generate a standard curve for each plate, enabling the conversion of OD to biotin levels.

2.3.3 Growth Assay with IPTG

To determine the resistance of the strains to BioB induction, the 96-well microtiter plates prepared during the *2.3.2.2 Inoculation of Production Plates* were sealed and placed in a spectrophotometer for 48 h at 37°C and 275 rpm. OD_{620} measurements were taken every 20 minutes to create growth curves for each strain.

3. Results and Discussion

3.1 *E. coli* Genomic Diversity Tool Development

3.1.1 Generation of Strain Libraries and Data for MMR Inactivation Model

The goal of the first part of the project was to generate genetic diversity in the strain while simultaneously controlling the mutagenesis rate to enable testing of all resulting distinct mutations for an improved phenotype and to limit the number of off-target mutations. The first approach was the inactivation of the endogenous mismatch repair system. The methyl-directed mismatch repair system (MMR) in *E. coli* is activated after replication, and its primary function is the removal of base-base mismatches or insertion-deletion mismatches that have escaped the replicative polymerases' proofreading function (Hsieh, 2001). The system can be deactivated by expressing a DNA methylase (Dam methylase EC 2.1.1.72), which hyper-methylates DNA during replication, shielding the newly replicated DNA from the mismatch repair system and increasing the rate of spontaneous mutations within the organism, resulting in a higher genetic diversity. The plasmid pBS2215, carrying the *dam* gene which encodes the DNA methylase was successfully transformed into the wild-type strain BS08072, creating strain BS07701 used for the experiment.

To generate genetic diversity, three colonies were taken from the newly created strain BS07701, and expression of the DNA methylase was induced over several generations alongside a control strain that was not induced. The rate of mutagenesis was monitored using the occurrence of rifampicin resistance in the colonies. Rifampicin resistance is developed through mutations in the rpoB subunit of the DNA-dependent RNA polymerase (Weinstein & Zaman, 2018). Rifampicin resistance should develop more often in the colonies where the DNA methylase was expressed than in the uninduced control culture, indicating a higher mutation rate due to the DNA methylase's expression. The number of rifampicin-resistant colonies should, therefore, increase steadily in the induced cultures, thus serving as a tool for determining the mutation rate. The obtained data should be compared to the NGS data to verify the accuracy of the rifampicin readout and its relation to actual mutagenesis rates in the strain.

Due to unexpected results in the full-length experiment, it was re-done later using a shortened protocol version, which was performed until S3 (Figure 7). The shortened protocol yielded much improved results and showed a steady increase in rifampicin resistance in the cultures where the DNA methylase was induced. In contrast, the number of resistant colonies in the control culture remained unchanged, as previously postulated. The graphs obtained in the full-length experiment, using strain BS06757, showed a substantial fluctuation in rifampicin resistance. Especially in the S3 generation, the number of rifampicin-resistant colonies in the induced cultures was lower than in the control strain. However, the plates prepared in this generation had excessive amounts of contaminating (likely fungal) colonies growing on the plates and are not representative of the actual number of rifampicin-resistant colonies due to possible competition for resources leading to growth inhibition. If the S3 data set is excluded, the occurrence of resistance still varies much more than expected and decreases at several points (Figure 8, full graph in the Appendix). Due to these unexpected results, the NGS samples were excluded from the analysis as the generated cryo-stock libraries were not used in further experiments. After re-examining the full-length experiment, the error was found to be improper lab practice. More specifically, an insufficient vortexing time of the cell suspension directly before plating, allowing the cells to settle and smaller numbers of cells to be plated, skewing the number of resistant colonies growing on the plates.

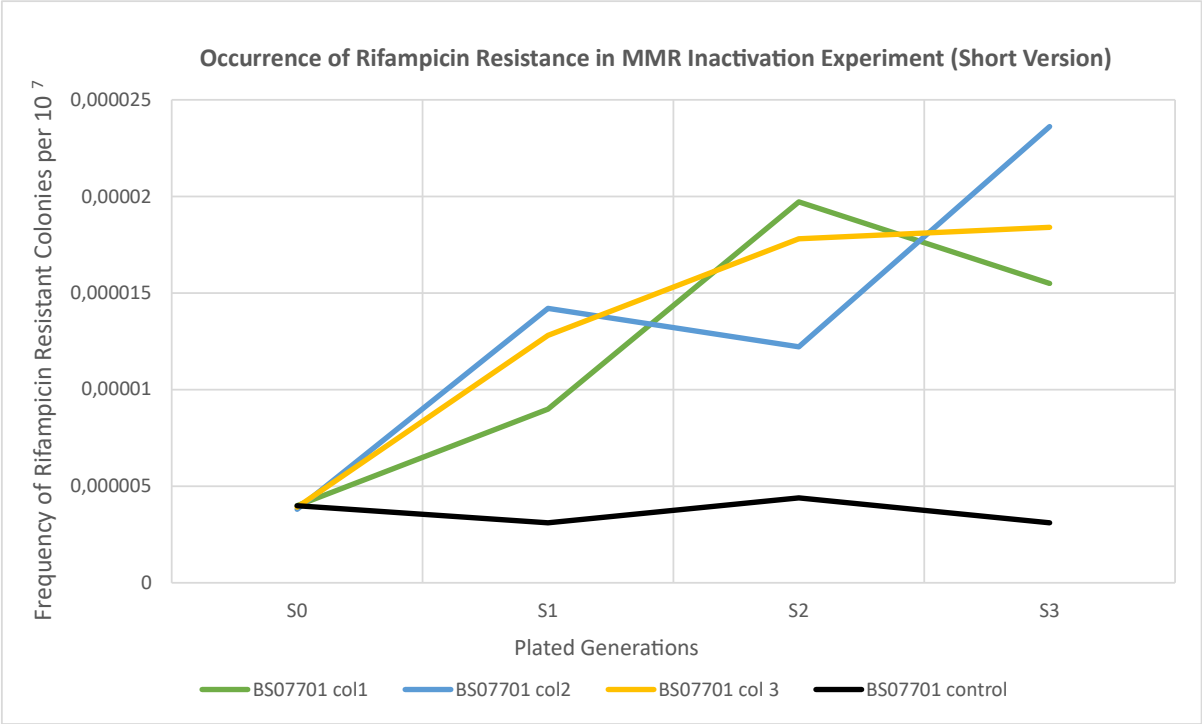


Figure 7 - Shortened MMR inactivation experiment showing steady increase of resistant colonies in the induced (col1, col2, and col3) cultures and no significant increase in the number of resistant colonies in the control.

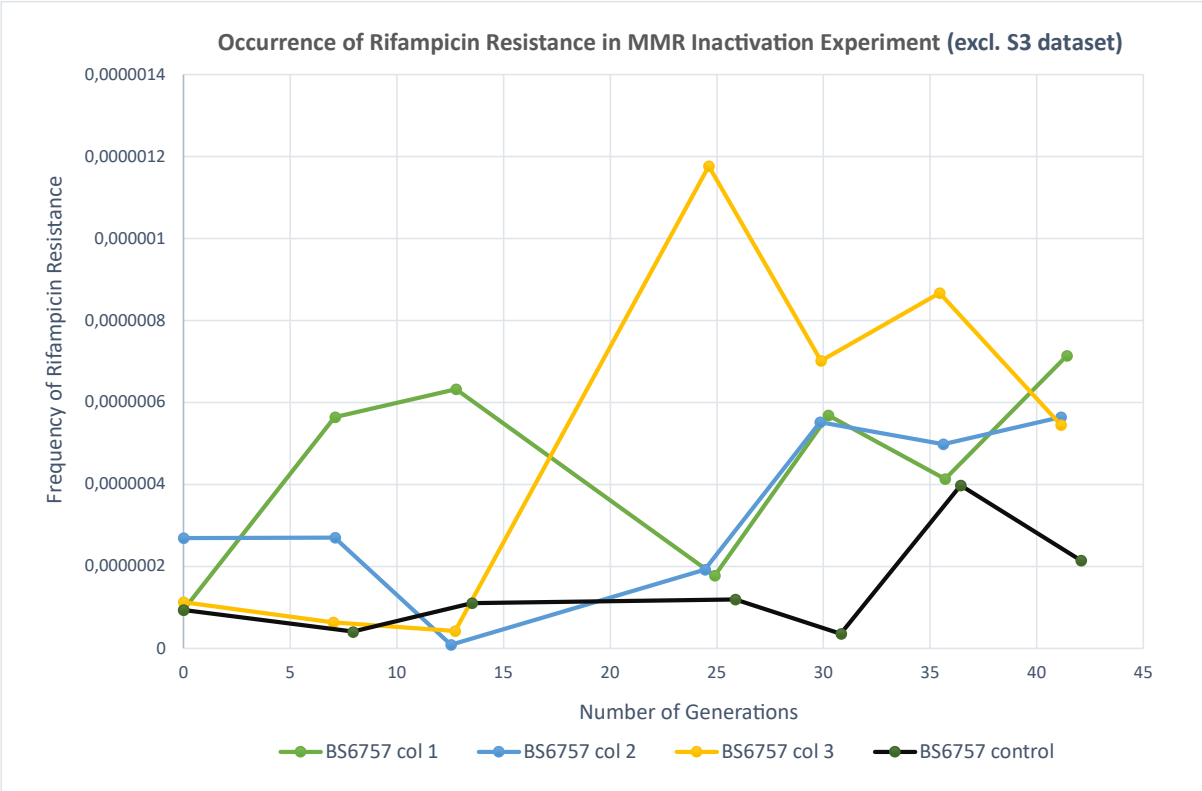


Figure 8 - Occurrence of rifampicin resistance in MMR inactivation experiment performed until S7. S3 dataset is excluded from the graph.

3.1.2 Creation of gDNA Libraries

To explore further diversity-creating measures, gDNA plasmid libraries were created using the DNA of *E. coli* (BS04755) and *S. cerevisiae* (BS07084) strains. This would enable the introduction of possibly beneficial genes into the strain, conceivably leading to the creation of fitter phenotypes, more resistant to the stress of BioB overexpression. The first step entailed purifying gDNA from both organisms and fracturing the genomic DNA into pieces between 1000 and 5000bp long, large enough to cover the average ORF (>1kB) and cloning them into a plasmid behind a constitutive promotor. The gDNA purification yielded samples with 373,3 ng/µL and 120 ng/µL of gDNA for *E. coli* (BS04755) and *S. cerevisiae* (BS07084), respectively. The fractionation protocol was performed with the full amount of available gDNA to ensure that there would be sufficient gDNA for the following steps in the protocol. The results of the fractionation protocol can be seen in Figure 9. The bands labeled BS07084 contain the *S. cerevisiae samples* and are less visible than those under BS04755, which contain the *E. coli* samples, indicating a lower amount of gDNA for the *S. cerevisiae* samples. This was expected, as less S. cerevisiae gDNA was available for the fractionation protocol.

Table 8 - Colony counts and diversity calculations for gDNA libraries.

Colony count			**Diversity**		**cPCR**	
Sample	1/1.000	1/10.000	1/1.000	1/10.000	%of true clones	Plasmids with inserts
BS07084	11	1	11000	10000	10	1050
BS04755	33	0	33000	0	15	2475
pos. control pBS1897	41	7	41000	70000		0

Following ligation of the DNA fragments with the vectors and transformation into *E. coli* (BS08072) a cPCR of the resulting colonies was performed to determine the diversity of gDNA lengths in the libraries. The number of colonies for the transformation of the gDNA libraries (Table 8) was very low overall. The *E. coli* strain BS04755 showed growth of 33 colonies at a 1:1000 dilution and none at 1:10000, and the *S. cerevisiae* strain BS07084 had 11 colonies growing on the 1:1000 dilution plate and 1 colony at the 1:10000 dilution. The positive control with the pBS1897 plasmid had 41 colonies and 7 colonies growing at a dilution of 1:1000 and 1:10000, respectively.

The results of the cPCR for the *E. coli* library can be seen in Figure 11. The length of the plasmid region augmented during the cPCR process with an empty vector should be around 241bp long. The results in Figure 11 indicate that most colonies were transformed with a vector carrying only a small insert or none. The desired insert length was achieved in lanes 13, 18, and 20 in three colonies. The band in lane 13 is at the height of 1500bp, indicating an insert length of around 1300bp, and the bands in lanes 18 and 20 are in between the 1000 and 1500bp markings, at around 1200bp, indicating insert lengths of just under 1000bp. The results of the cPCR for the *S. cerevisiae* library can be seen in Figure 10. The results show a single vector with an insert within the desired length (1000-5000bp) at a height of 2000bp in lane 7. All other lanes contain samples with a low insert length or no insert.

The percentage of true clones for both libraries was calculated by taking the number of samples in the cPCR with the correct insert length and dividing it by the total number of colonies tested in the cPCR. The diversities are calculated by multiplying the number of colonies with the dilution factors. The percentage of true clones for BS04755 is 15%, and for BS07084, it is 10%. This percentage, multiplied by the average diversity, yields the number of plasmids in the library that carry inserts of the desired length. For BS07084, that number is 1050, and for BS04755, it is 2475, meaning that this is the number of plasmids that might carry a gene that could confer an advantage to the strain.

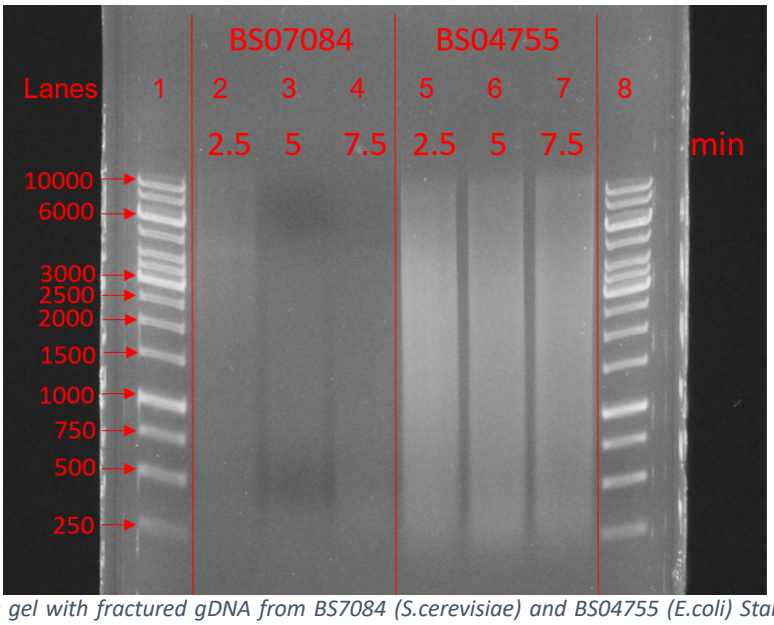


Figure 9 - 1% agarose gel with fractured gDNA from BS7084 (S.cerevisiae) and BS04755 (E.coli) Stained with gel red and photographed under UV light. Times indicated in image refer to fractionation times in the 2.1.3.3 Fractionation protocol (adapted from New England Biolabs)

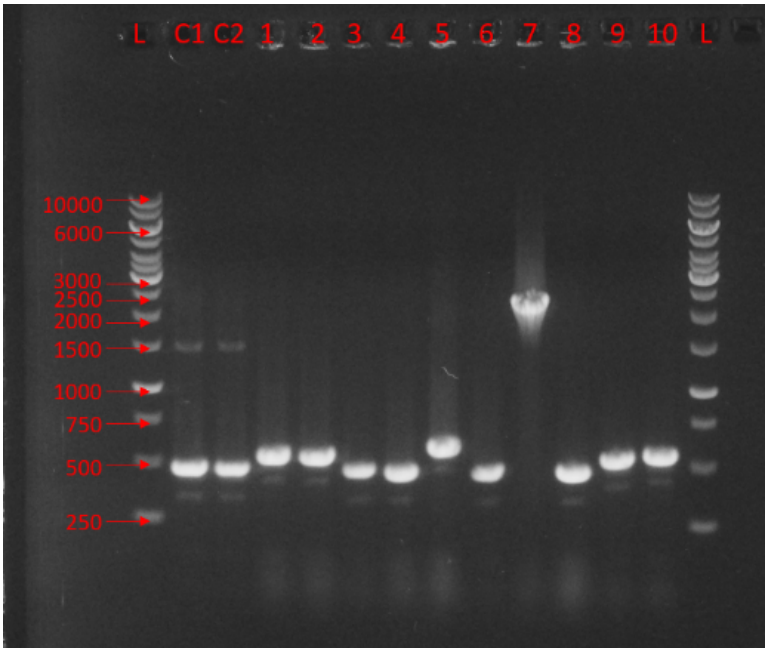


Figure 10 - 1% agarose gel with cPCR samples from gDNA library using 1kbp-5kpb BS078084 (S. cerevisiae) fragments. C1 and C2 are controls using the empty plasmid vector pBS1897.

The difficulty in producing libraries with sufficient insert length lies in the complications that blunt-end cloning brings with it. Blunt-end cloning is a less efficient technique than cloning with overhangs, and smaller DNA fragments usually ligate easier than longer fragments. The fractionated DNA possibly contained too many smaller DNA fragments, which interfered with the ligation of the longer DNA fragments, leading to a higher number of smaller inserts in the vectors. When comparing the colony counts of the library transformations to the colony counts of the positive control using the empty plasmid vector, the overall number of colonies in the positive control (Table 8) was equally low, indicating a low transformation efficiency. This could be due to the samples possibly containing ethanol residue from the DNA purification steps, which could have interfered with the transformation. It could

have also been caused by the electrocompetent cells not being in the exponential phase when they were prepared for electroporation. Another problem with this protocol is the loss of DNA going through all the protocol steps. After several purification steps, the amount of DNA left for transformation was significantly reduced. A way to optimize this protocol would be to increase the concentration of initial gDNA, concentrate the samples between steps, and pool samples to obtain enough DNA. Overall, the diversities of the libraries were not high enough to be used in further experiments as most inserts are not long enough to cover a standard ORF length and thus are unlikely to confer any beneficial genes to the strain. The libraries were, therefore, excluded from further experiments.

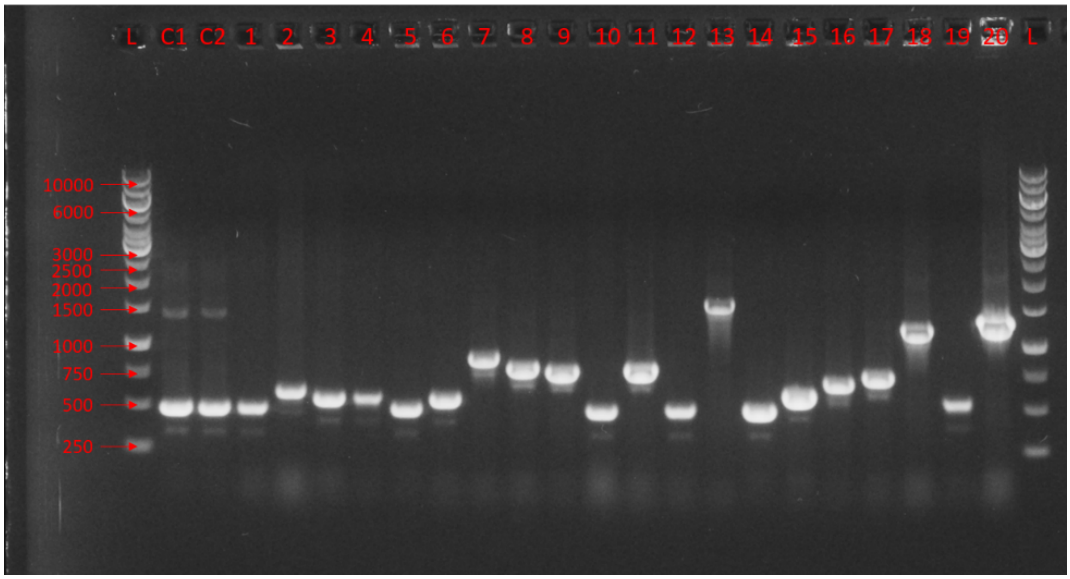


Figure 11 - 1% agarose gel with cPCR samples from the gDNA library using 1kbp-5kpb BS04755 (E. coli) fragments. C1 and C2 are controls using the empty plasmid vector pBS1897.

3.2 Functional Selection Development

3.2.1 Determination of Stressor Levels

This part of the project focused on the creation of selection gradients to select for phenotypes with higher resistance to oxidative stress. Cobalt, paraquat, and hydrogen peroxide are compounds known to induce oxidative stress in *E. coli* cultures (Fantino et al., 2010; Hassan & Fridovich, 1978; Nur et al., 2014). Oxidative stress is a cellular imbalance in oxidizing versus reducing species (Fleming & Burrows, 2020). It affects every sub-step of the central dogma of molecular biology (Fasnacht & Polacek, 2021). In bacterial cells, oxidative stress leads to damage to the backbone and bases of nucleic acids and damage to co-factors of proteins. The damage is mitigated by different stress responses, with different stress response regulons activated in bacteria, depending on the type of stressor. Oxidative stress is caused *inter alia* by the production of reactive oxygen species (ROS) like superoxide, hydrogen peroxide, and hydroxyl radicals.

Hydrogen peroxide can pass through the semipermeable membrane but cannot damage DNA directly. Its mutagenic effect can be explained via the Fenton reaction, which is mediated by DNA-associated Fe^{2+} ions. These react with hydrogen peroxide and form a highly reactive hydroxyl radical, which can react with most biomolecules. When produced near DNA, they can damage both the nucleobase and the deoxyribose moieties, which leads to mutations and strand breaks. Paraquat is also able to penetrate bacterial cells. Once inside, it oxidizes redox enzymes and produces superoxide by transferring electrons to oxygen. Superoxide promotes hydroxyl radical formation, resulting in DNA damage. It has also been found that it accelerates DNA damage by leaching iron from storage proteins or iron-sulfur clusters. Cobalt causes oxidative stress in cells by competing with iron in various metabolic processes, including iron-sulfur cluster assembly (Majtan et al., 2011). It also leads to the generation of additional ROS through cobalt-mediated free radical reactions (Leonard et al., 1998).

Paraquat and cobalt were chosen for selection development because their respective toxicity mechanisms are said to impact iron-sulfur cluster assembly, while hydrogen peroxide is a more classically mutagenic compound that causes random DNA damage. The first step included literature research for inhibitory concentrations, after which gradients were created to determine more strain-specific concentrations. The expression of BioB production using IPTG was also used as a selection factor, harnessing the toxicity of BioB.

For hydrogen peroxide (Figure 12), strong cell growth was seen in plates containing 0.1 and 0.5 mM of hydrogen peroxide, a few colonies growing at 1 mM, and no growth was visible at 1.5 and 2 mM. The plates with cobalt as a stressor (Figure 15) had strong growth at 1 mM and significantly reduced growth at 2 mM. No growth was visible in the higher concentrations. The plates where BioB production was induced (Figure 14) showed no reduction in growth across all concentrations. This was likely caused by a miscalculation during the preparation of the IPTG stock solution, leading to much lower amounts of IPTG used in the plates. In the plates containing paraquat as a stressor (Figure 13), small amounts of growth were visible at 100 and 200 µM concentration and the control plate was fully covered in growth. The plates with 300, 400, 500, and 600 µM of paraquat showed no visible growth. The compounds paraquat and hydrogen peroxide were selected for the creation of a 2D gradient, the other compounds were excluded due to time constraints.

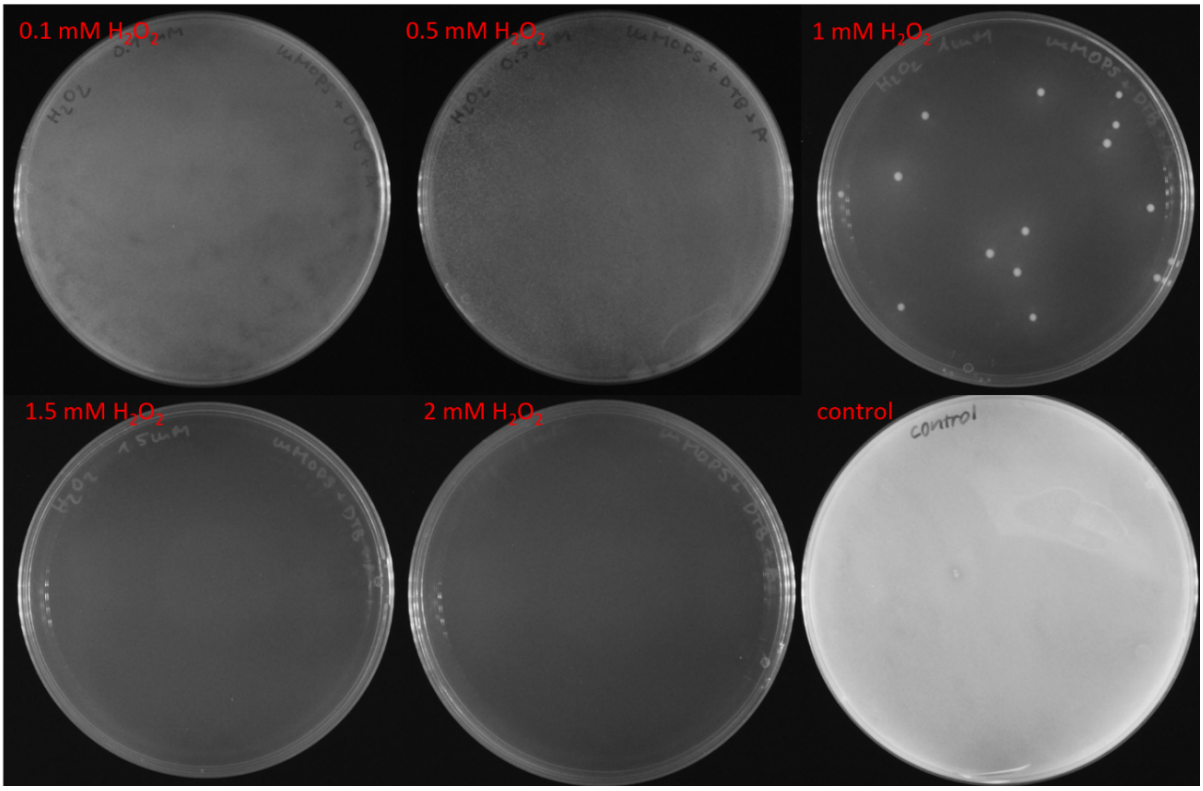


Figure 12 - mMOPS+DTB+A selection plates with hydrogen peroxide as stressor. Concentrations increase from left to right and top to bottom: 0.1 mM, 0.5 mM, 1 mM in the top row, 1.5 mM, and 2 mM and control plate without stressors in bottom row. Strong cell growth is visible at 0.1 and 0.5 mM, a few single colonies is visible at 1 mM, and no growth visible at 1.5 and 2 mM.

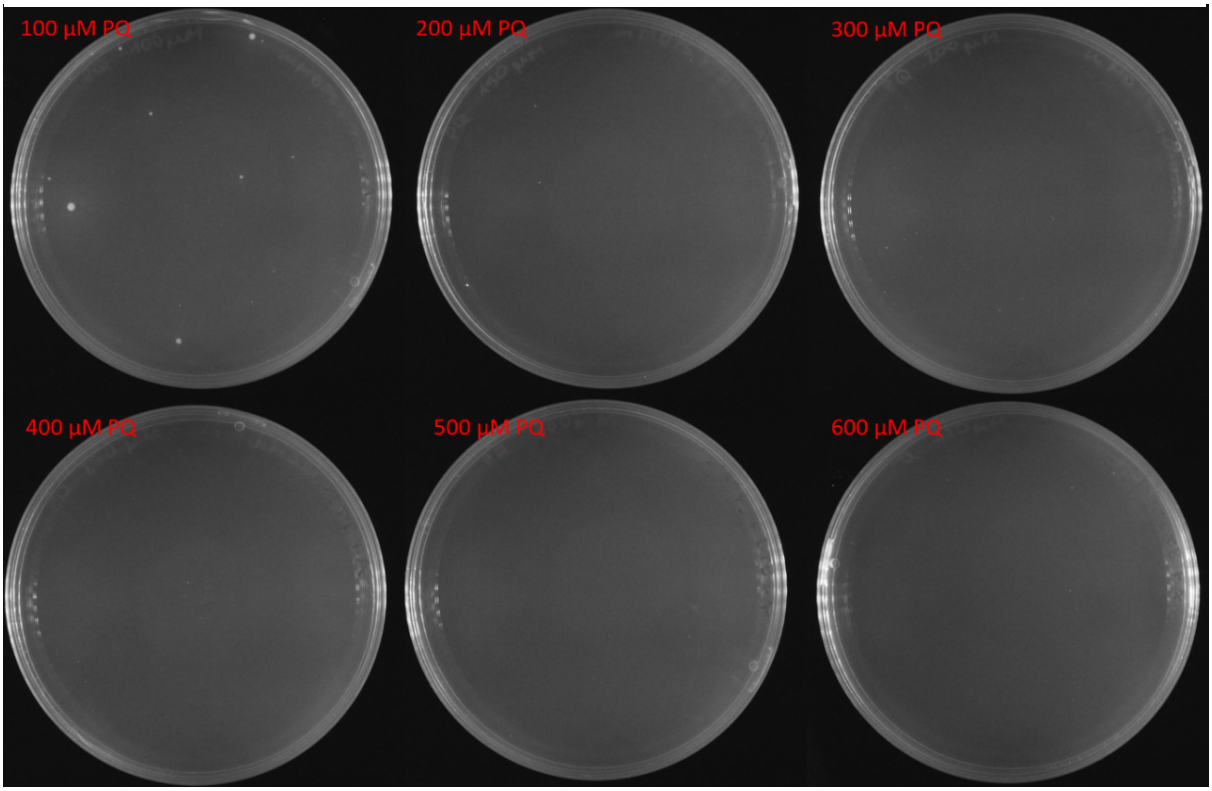


Figure 13 - mMOPs+DTB+A selection plates with paraquat as stressor. Concentrations in top row from left to right: 100 µM, 200 µM, 300 µM. Stressor concentrations in bottom row from left to right: 400 µM, 500 µM, and 600 µM. A few single colonies are visible at 100 µM and 200 µM, no growth in 300, 400, 500, and 600 µM concentration plates. The control plate for PQ is the same as for the hydrogen peroxide plates.

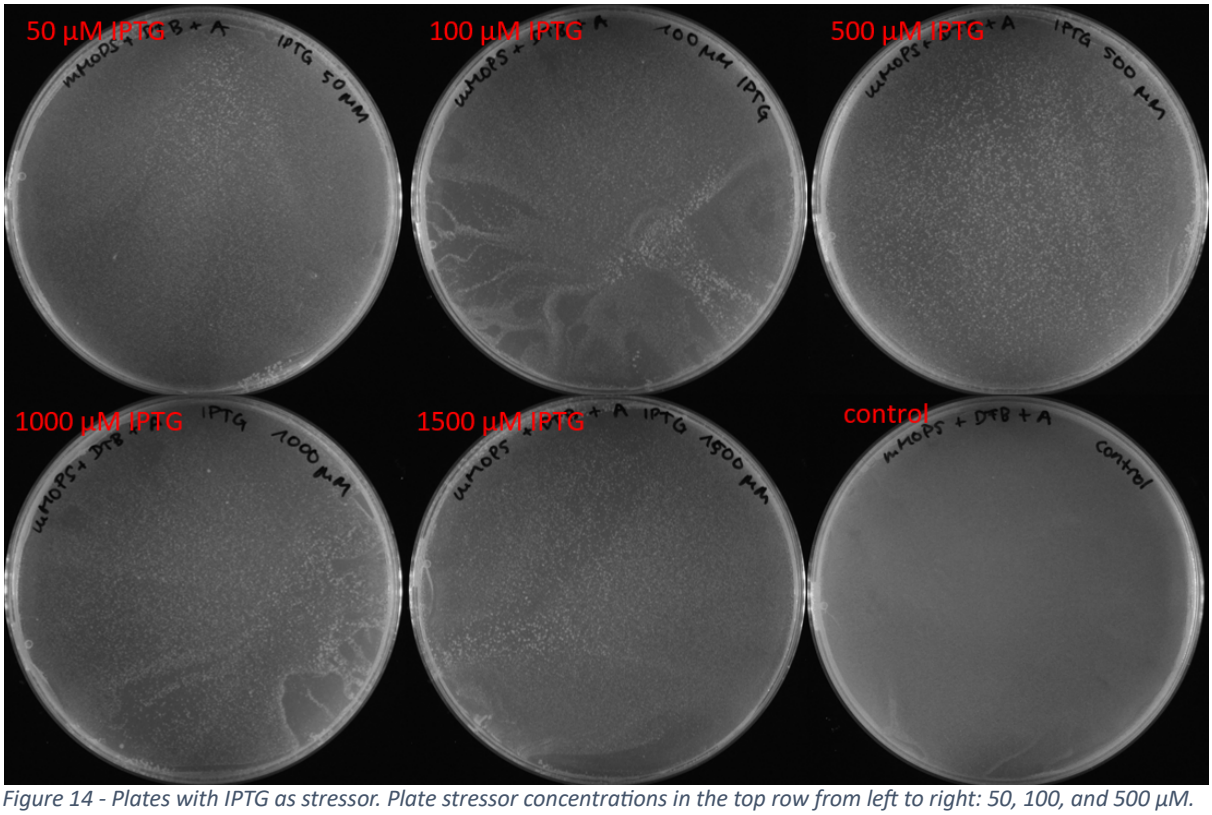


Figure 14 - Plates with IPTG as stressor. Plate stressor concentrations in the top row from left to right: 50, 100, and 500 µM. Stressor concentrations in bottom row from left to right: 1000, and 1500 µM. Control plate without stressor in bottom right corner. Strong growth is visible across all concentrations.

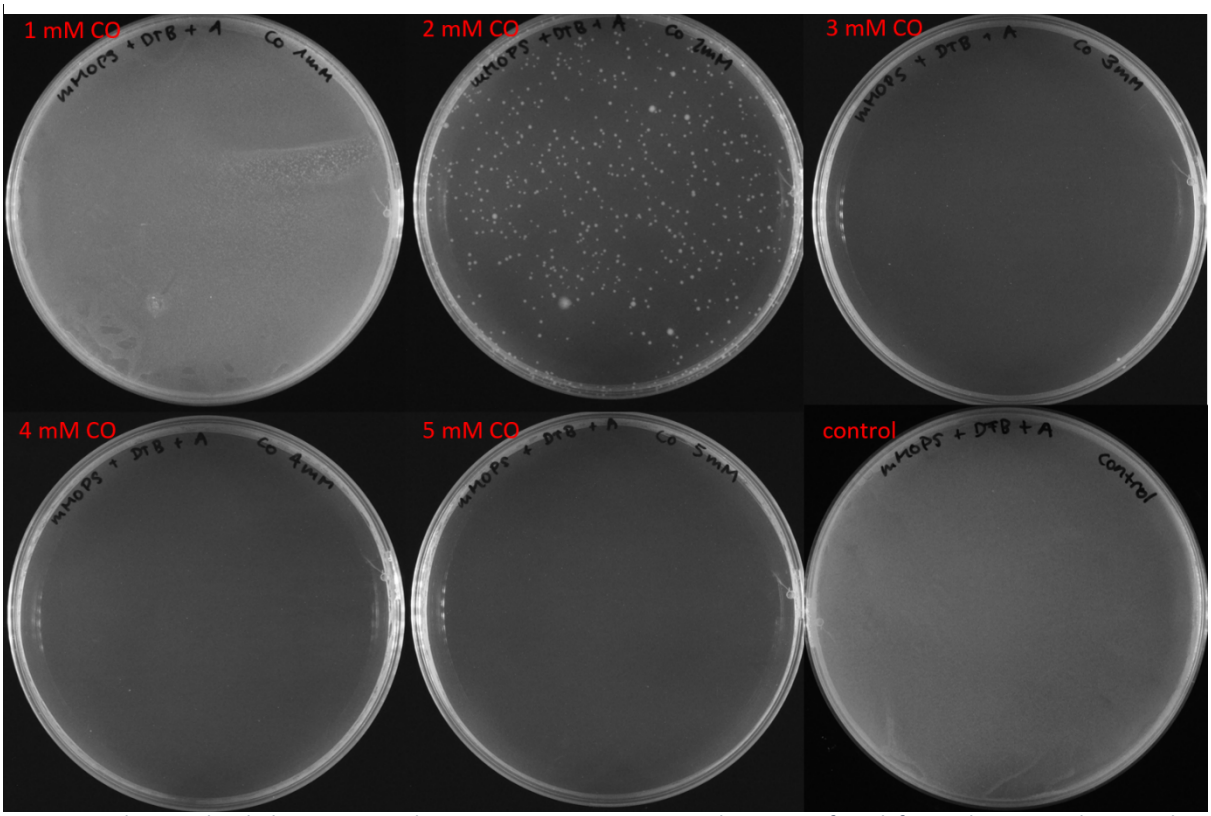


Figure 15 - Plates with cobalt as stressor. Plate stressor concentrations in the top row from left to right: 1, 2, and 3 mM. Plate stressor concentrations in the bottom row from left to right: 4 and 5 mM. Control plate without stressor in bottom right corner. Strong growth is visible at 1 mM and on the control plate. Single colonies are visible at 2 mM and no growth is visible at 3, 4, and 5 mM.

3.2.2 Generation of a 2D Gradient

The stressor concentrations of the plates that showed significantly reduced growth were chosen as upper limits for the 2D gradients, as a combination of two different stressors would conceivably increase cellular stress and kill the cells entirely if the concentrations were too high. Several gradients were tested using 16 different concentration combinations of hydrogen peroxide, cobalt and paraquat, (in Appendix) but due to time constraints and to reduce material costs, the number of selection plates was reduced to 9. A gradient was created and tested using 100 µM of PQ and 1 mM of H2O2 as upper stressor limits for 2.2.3 Selection for Improved Phenotypes.

3.2.3 Selection for an Improved Phenotype

The goal of the selection system was to act as a sieve, reducing the large population of plated cells into a manageable pool of potential hits to be screened more in-depth based on production. Cells growing at high concentrations of both stressors would be preferred candidates, as they show a higher resistance to oxidative stress, which could translate to resistance towards BioB-induced stress and thus higher biotin producing capabilities.

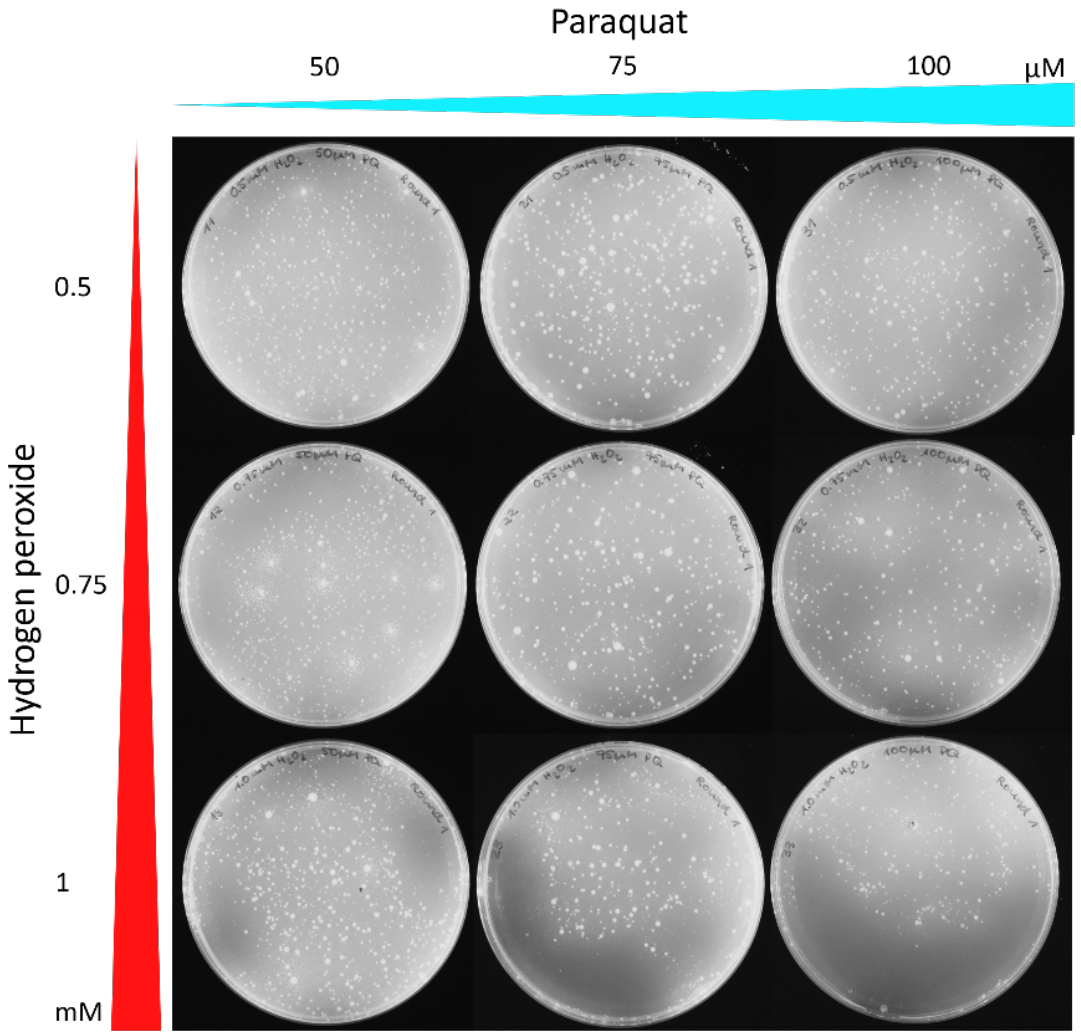


Figure 16 -Round 1 (not induced) 2D stressor gradient selection plates. Pictures taken after 5 days of incubation.

The cells growing on the plates in Figure 16 (round 1 of plating for selection protocol, uninduced) showed growth across all stressor concentrations, with only a minimal decrease in cell growth at the highest stressor concentrations. The cells growing after the first induction cycle (Figure 17), labeled "Round 2", showed decreased growth in nearly all the plates compared to the Round 1 plates, but especially in the higher concentrations, towards the lower right corner of the gradient. The cells from Round 3 (Figure 18), which were induced twice, also show decreased growth compared to the first round of plating, but higher growth than visible in the second round. The low cell growth in Round 2 is likely attributable to an extended drying time under the clean bench, leading to dryer plates with decreased water availability. The decreased number of cells growing on Round 2 and 3 plates compared to Round 1 plates is likely a consequence of disadvantageous off-target mutations, leading to a reduction in cell fitness and robustness and, therefore, a decrease in colonies.

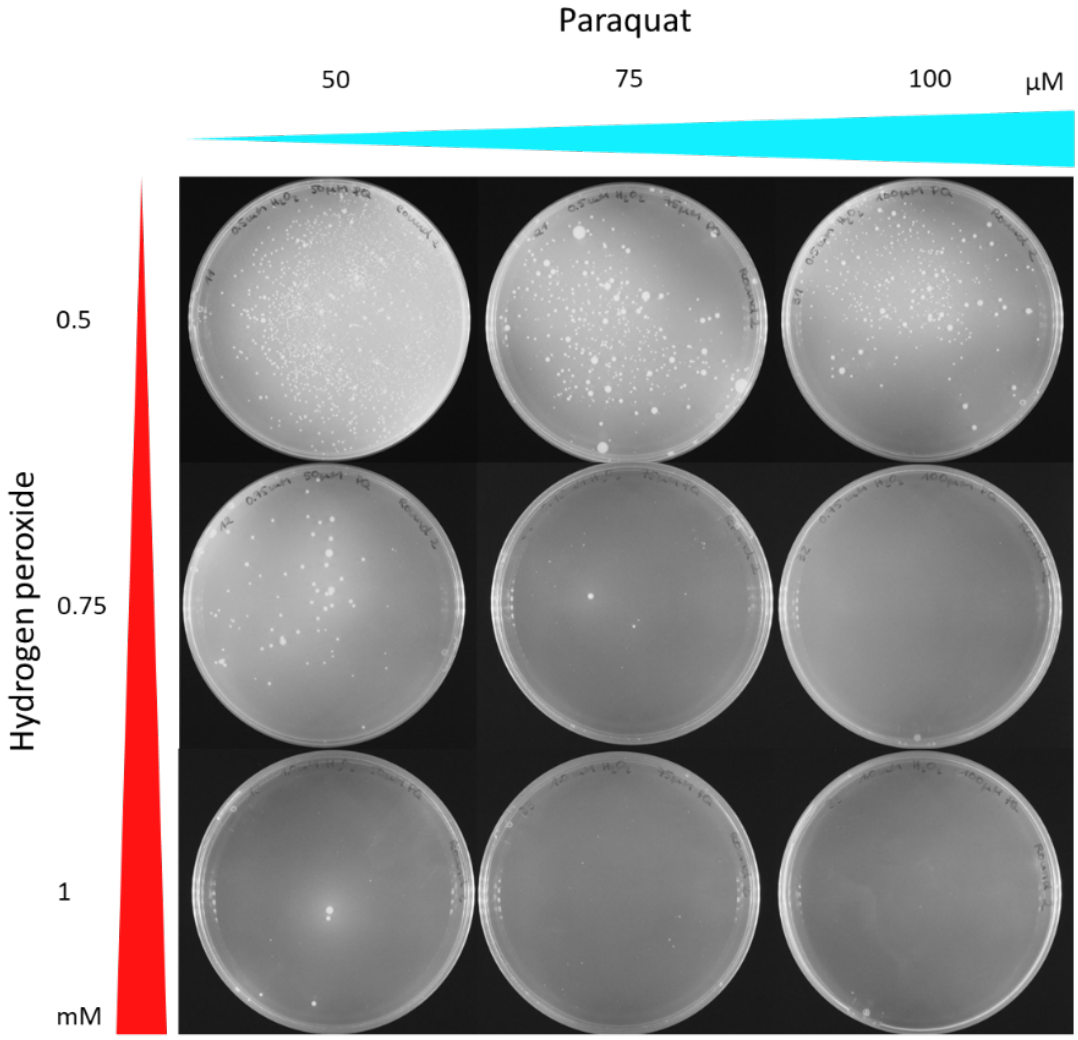


Figure 17 - Round 2 (induced once) 2D stressor gradient selection plates. Pictures taken after 5 days of incubation.

Due to the low number of colonies in the high stressor concentrations of Round 2 plates, only colonies from Rounds 1 and 3 were selected for screening experiments. Four of the largest colonies were picked from each batch of plates to be used in the growth experiment with stressor concentrations.

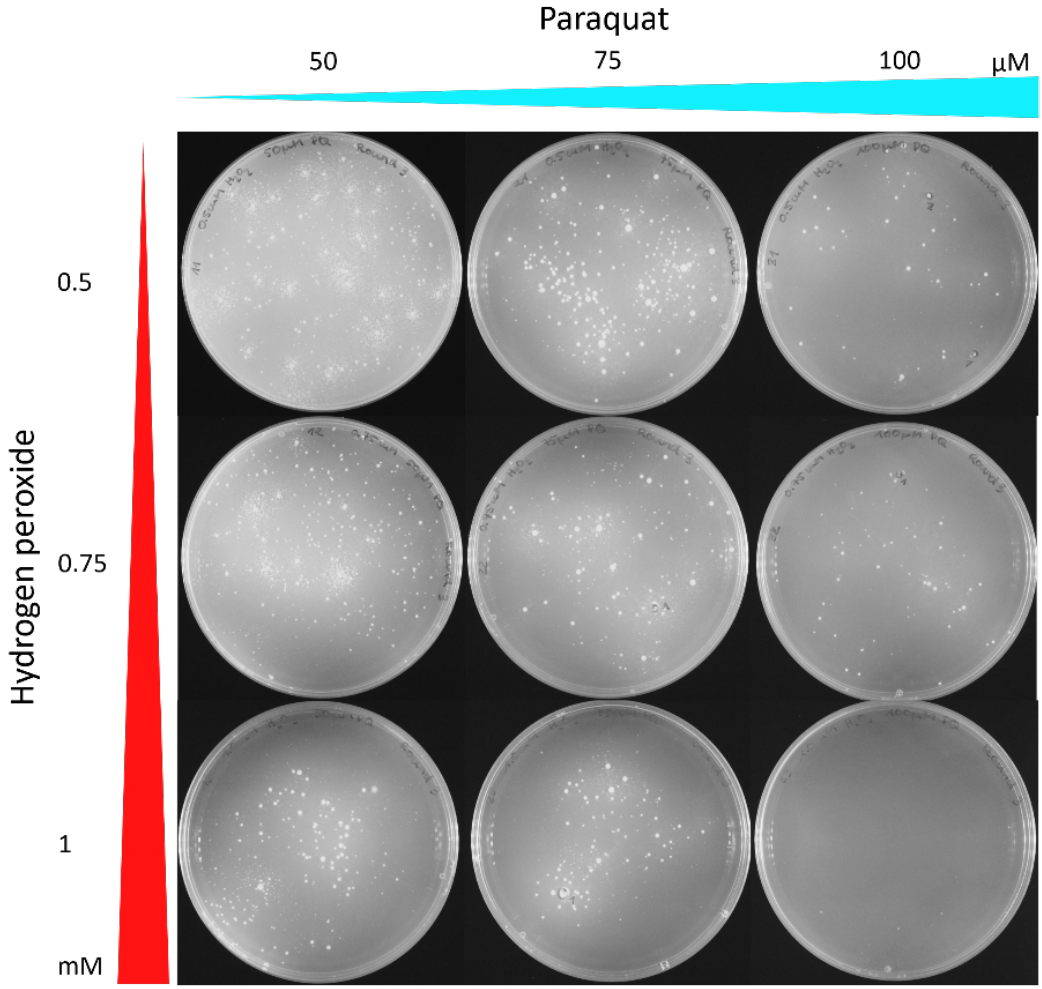


Figure 18 - Round 3 (twice induced) 2D stressor gradient selection plates. Pictures taken after 5 days of incubation.

3.3. High-throughput Screening

3.3.1 Growth Assay with Stressors

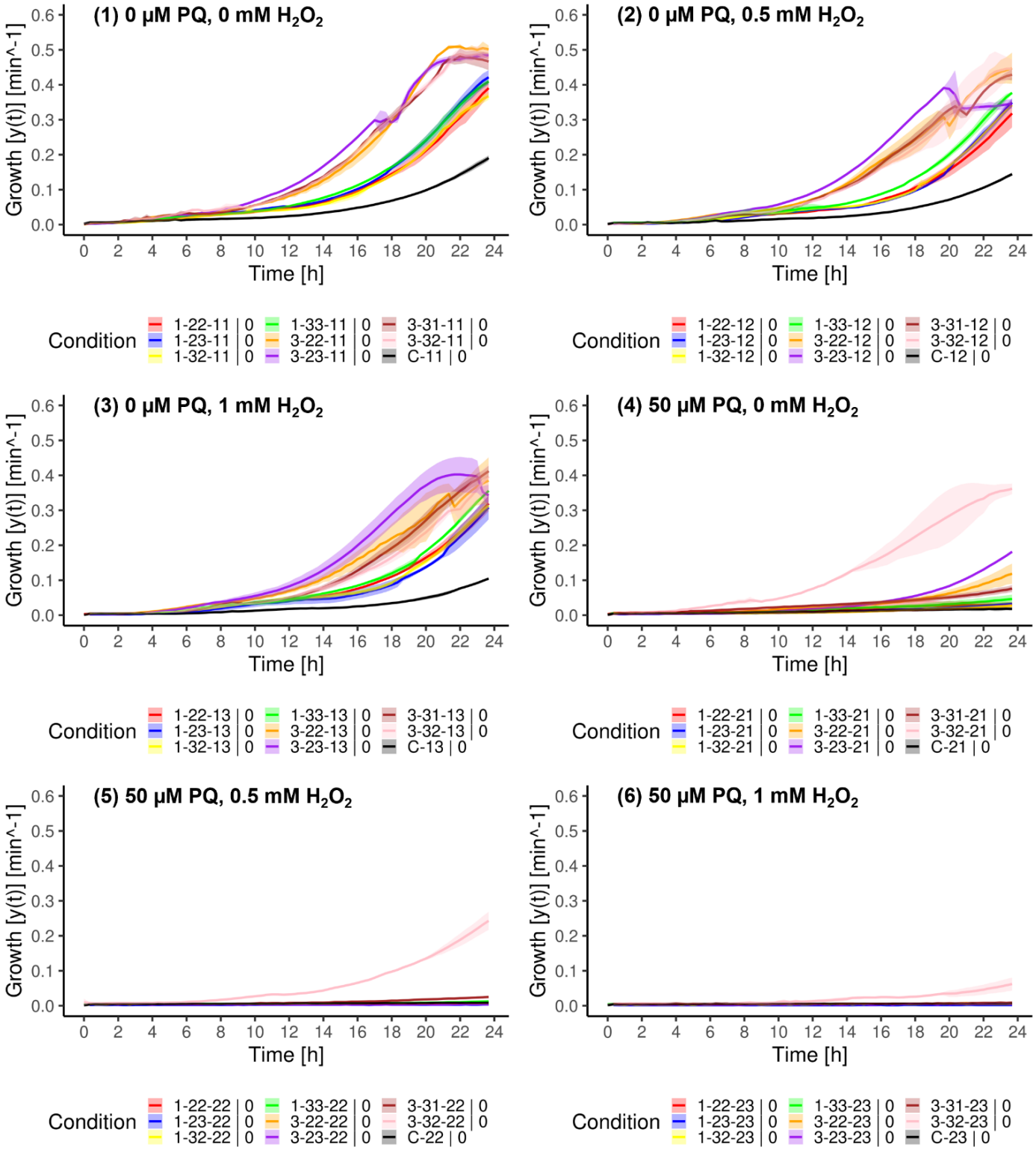


Figure 19 -Numbered graphs with the growth curves of six of the nine tested stressor concentrations, the graphs with the growth curves for the highest three concentrations (31,32,33) can be found in the appendix. Naming conventions for each strain: Round(1,3)-Plate(22,23,31,32,33)-stressor concentration(11, 12, 13, 21, 22, 23, 31, 32, 33). Example 1-22-11 = strain taken from round 1, plate 22, growing in stressor 11. Stressor concentration labels are explained in Table 7. A mutant strain key can be found in the appendix. Samples labeled C-[number] contain the background strain BS07701 and act as a control growing in the respective concentration.

A growth experiment using nine different stressor concentrations (Table 7) was performed to narrow down the strains which would be used for the biotin production assays. The strains which would have the highest growth rate in the different stressor concentrations would be used for further investigations as they would have the highest resistance to oxidative stress, which could translate into a higher resistance to BioB induction and thus higher biotin production capabilities. In graph (1) (Figure 19), which shows the individual growth curves with no stressors, the growth curves for Round 1 are grouped together, and the growth curves for Round 3 strains are grouped together. The control strain has the lowest growth curve. The strains from Round 1 show higher growth rates than the control strain, and the curves from Round 3 show the highest growth. In graph (2), which shows the growth rates for the strains growing in 0.5 mM of H_2O_2, the growth curves are still grouped as in graph (1), with the strains from Round 3 having the highest, strains from Round 1 the second highest, and the control strain having the lowest growth rates. In graph (3) (stressor concentration: 1 mM H_2O_2) Round 3 strains still dominate, showing the highest growth curves. In graph (4) (stressor concentration: 50 µM PQ) strain 3-32 shows the highest growth curve by far, indicating that this particular strain is most resistant to paraquat as a stressor. This is also evidenced by the fact that in graphs (5) and (6) this strain still shows the highest growth rate. Throughout all stressor concentrations the control strain remains the one with the lowest growth rate. The strain that was excluded from further experiments is strain 1-32, which has the lowest growth rate at the highest stressor concentration where growth of all strains was still observed: 50 µM PQ.

The most likely explanation for faster growth rates of the selection strains when compared to the control strains is that only the largest colonies on the selection plates were picked and used for the growth experiment, leading to a selection of strains with higher growth rates. The high growth rates of Round 3 strains also indicate that they are indeed more resistant to the various stressor concentrations, and that the induction of the DNA methylase during the selection experiment likely led to beneficial genetic changes with improved oxidative stress resistance phenotypes. Round 1 strains were not induced, they were selected simply for growing faster on the stressor concentrations, which also led to higher growth curves for these strains, though not as high as the Round 3 strains. This could be an indication that the acceleration of the mutation rate in the Round 3 strains led to more beneficial genetic changes faster.

3.3.2 Growth Assay with IPTG

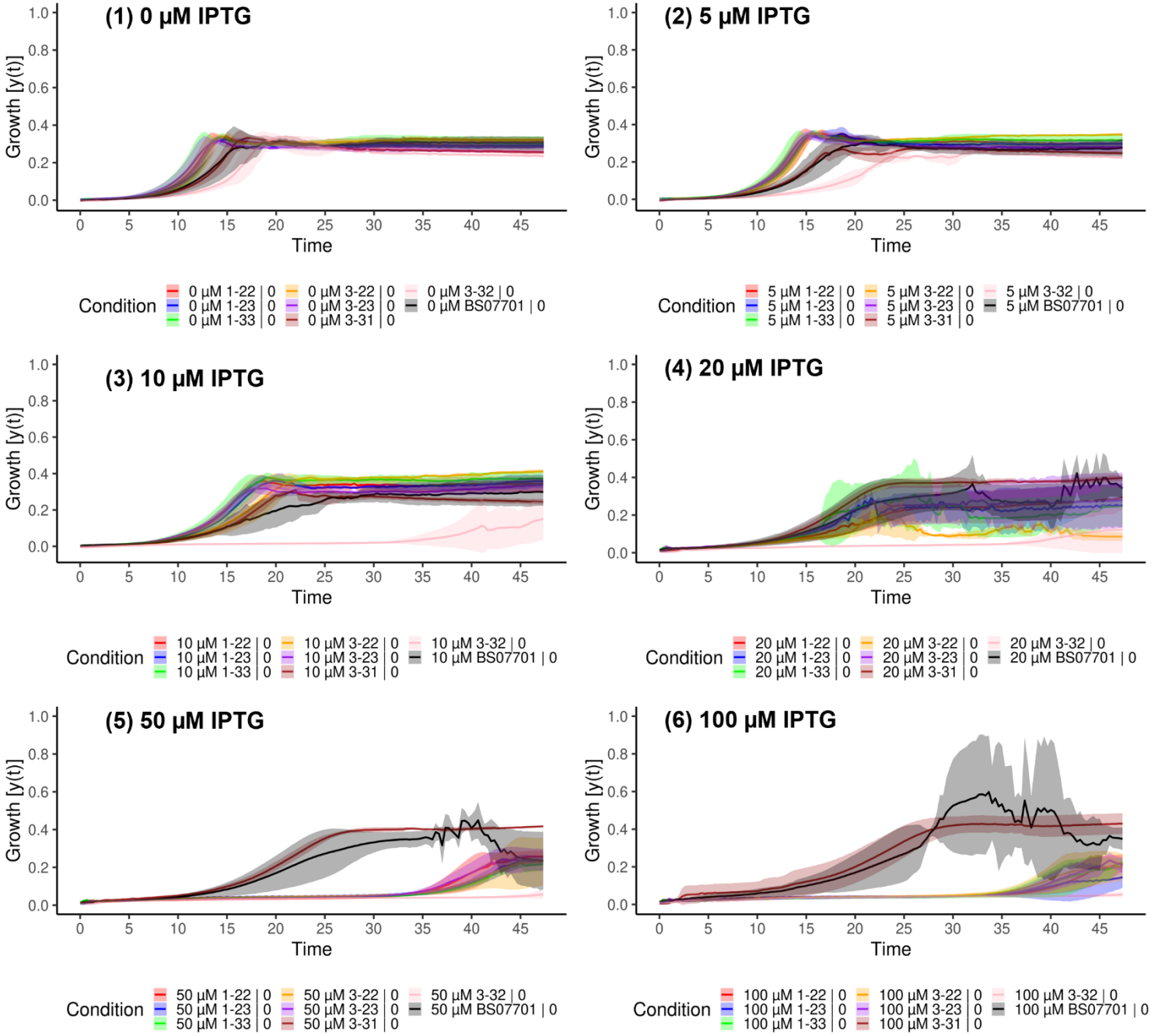


Figure 20 - Graphs showing growth curves of selection strains at 0, 5, 10, 20, 50, and 100 µM of IPTG

To observe the effects of BioB induction on the strain, a growth experiment using an IPTG gradient was conducted to evaluate the growth rates of the selection strains. At 0 µM IPTG the growth rate for strain 3-32 is the lowest, lower even than the growth rate of the control strain. Which does not match with the growth rates this strain showed in the growth experiment in Figure 19, graph (1), when no stressors were used. In the subsequent IPTG concentrations, 3-32 remains the strain with the lowest growth curve. Cross checking with the end ODs of the production plate (Table 9), the low growth curve of this strain in this experiment was likely the result of a low inoculum volume, caused by an issue with the pipettor, rather than a high susceptibility of the strain to BioB induction. This is further evidenced by the fact that ODs of this strain in the first plate (0, 5, and 10 µM IPTG concentrations) do not differ

significantly from the ODs of the other strains at these concentrations but are massively lower than the ODs of all the other strains in the second plate. The second plate was prepared last, so there was likely an issue with the pipettes in this row, which led to a lower inoculum volume. Strain 3-31 shows a faster growth rate than most strains throughout all IPTG levels. Especially at high IPTG concentrations, at 50 and 100 µM, this strain, along with the background strain, shows a much higher growth curve than the other strains, indicating higher resistance to IPTG induction.

3.3.3 Biotin Production Assay

A biotin production assay was carried out to determine whether the strains that show a higher resistance to oxidative stress, induced via hydrogen peroxide and paraquat, also show a higher resistance to BioB-induced oxidative stress and thus higher biotin production capabilities. The production titers using different IPTG induction levels are plotted in Figure 21. Strain 1-23-1 produces its highest amount of biotin at around 3.05 mg/L using 10 µM IPTG for induction. The optimum for strain 3-31-1 is at around 2.8 mg/L using 5 µM IPTG. The control strain reaches its maximum production at 3.1 mg/L and 20 µM IPTG induction. In strain 3-32-1 the levels measured have wide error bars from the 10 µM concentration onwards, and the optimum is at around 2.7 mg/L (50 µM IPTG), though the value is not reliable. The results for this strain indicate that there was a measurement issue during the biotin assay because the bars usually follow a curve that has an optimal IPTG induction level, with biotin titers tapering off at both sides of the optimum. The low levels of the bars at 10 and 20 µM interrupt this curve, and due to the high error bars indicating a large degree of uncertainty, they are not necessarily true representations of the true biotin production titers for this strain. Strains 3-32-1, 1-33-1, and 1-22-1 follow this optimum curve nicely, though there are wide error bars in any of

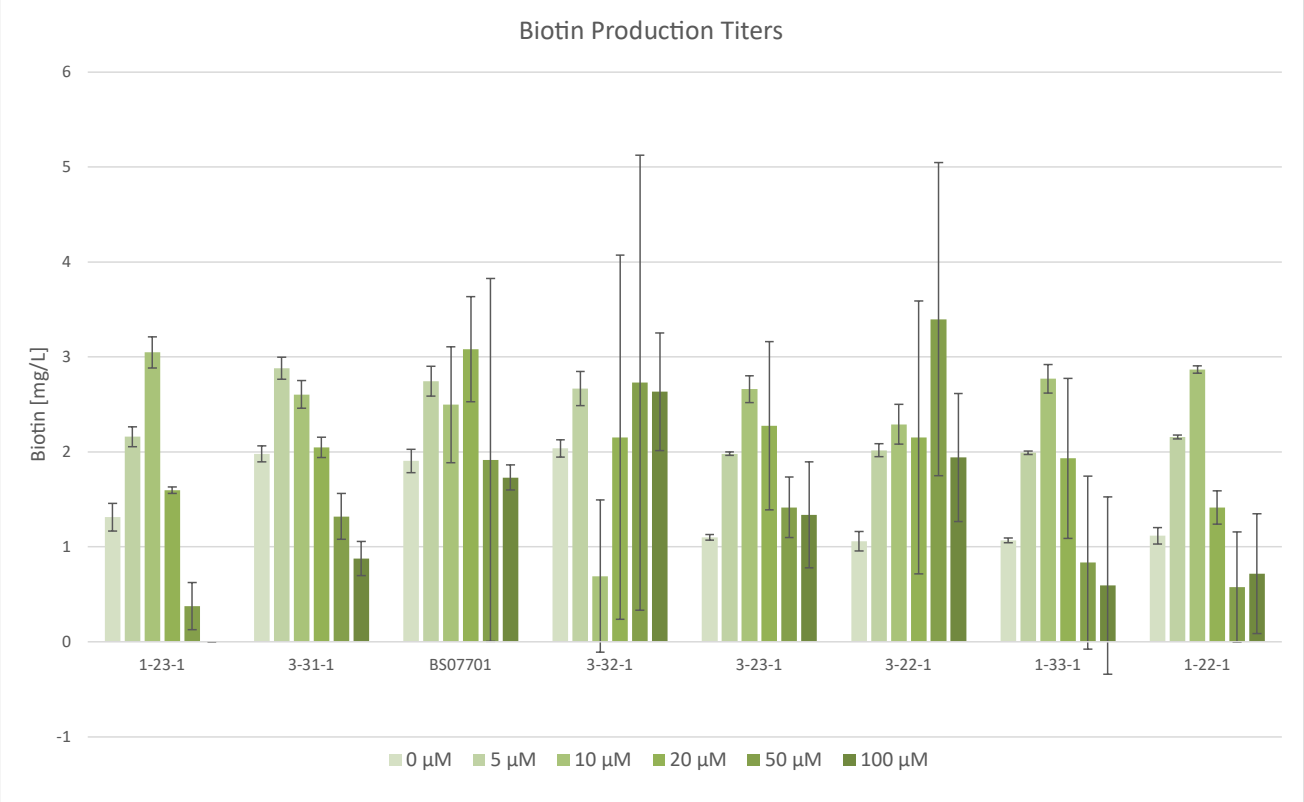


Figure 21 - Biotin production titers using the selected strains and BS07701 as a control strain.

the measurements from 20 µM upwards, and all their highest titers remain lower than the highest of the control strain. The high error in the measurements from 20 µM onwards could also be a result of issues with the pipettor when preparing the second plate. Strain 3-22-1 shows the highest production titer overall, but the error bar is wide, so the result is not reliable and likely not representative of this strain's true production capabilities.

The biotin production of strain 3-31-1, which had the highest growth rates at high IPTG concentrations, did not differ considerably from the biotin titers of the background strain. Though this strain was able to maintain a better growth curve at high IPTG than all the other strains, this did not translate into higher biotin production levels. Since it can be assumed that BioB expression is proportional to IPTG concentration, the strain likely produced more BioB, but this did not result in a higher rate of DTB to biotin conversion. This could be caused by iron-sulfur cluster depletion leading to a lack of available sulfur for the DTB to biotin conversion. Another possibly inhibiting factor could be the delivery of electrons, or SAM, which are also needed for the turnover of the enzyme.

Examination of the end ODs (620 nm) for the biotin production plate show that there is a significant difference within the ODs of strain 3-32-1 at 10 µM. One of the triplicates shows an OD of 0.045 compared to the other two triplicates, showing ODs of around 0.21. This low growth could be an explanation for why the biotin production titer is so low for this strain at 10 µM. This low growth could be caused by an error while transferring the 5 µL inoculum into the production plate using the multichannel pipettor. It is likely that a much smaller volume of inoculum was transferred into the plate in this row, leading to a low number of cells. This would also explain the low ODs in the higher concentrations for this strain, where the pipetting error likely continued. The abnormally high OD measurements for strain 3-32-1 at 100 µM are likely caused by an air bubble which disturbed the measurement.

It is likely that higher resistance to the oxidative stress induced by hydrogen peroxide and paraquat does not necessarily translate to higher resistance to BioB induction or that higher BioB induction does not lead to increased biotin production capabilities.

Table 9 - End ODs for biotin production assay

Strain	0 µM IPTG	0 µM IPTG	0 µM IPTG	5 µM IPTG	5 µM IPTG	5 µM IPTG	10 µM IPTG	10 µM IPTG	10 µM IPTG
1-23-1	0,29918404	0,30825023	0,29465095	0,30825023	0,29011786	0,31731641	0,353581142	0,349048051	0,358114234
3-31-1	0,24932004	0,27651859	0,28558477	0,28105168	0,27198549	0,28105168	0,312783318	0,299184044	0,303717135
BS07701	0,28558477	0,27651859	0,27651859	0,30371714	0,28558477	0,27651859	0,330915684	0,303717135	0,312783318
3-32-1	0,29011786	0,29011786	0,2674524	0,2674524	0,29918404	0,27198549	0,217588395	0,045330916	0,213055304
3-23-1	0,30825023	0,29918404	0,31278332	0,32638259	0,33091568	0,32638259	0,335448776	0,3762466	0,335448776
3-22-1	0,35811423	0,33998187	0,3218495	0,33544878	0,33998187	0,33091568	0,358114234	0,353581142	0,326382593
1-33-1	0,3218495	0,31278332	0,32638259	0,34904805	0,3218495	0,31731641	0,358114234	0,371713509	0,389845875
1-22-1	0,3762466	0,32638259	0,35811423	0,33091568	0,36264733	0,31278332	0,389845875	0,3762466	0,299184044

Strain	20 µM IPTG	20 µM IPTG	20 µM IPTG	50 µM IPTG	50 µM IPTG	50 µM IPTG	100 µM IPTG	100 µM IPTG	100 µM IPTG
1-23-1	0,18585675	0,2175884	0,19492294	0,17679057	0,1631913	0,19945603	0,081595648	0,077062557	0,077062557
3-31-1	0,19945603	0,23572076	0,2175884	0,16772439	0,1631913	0,19038985	0,176790571	0,149592022	0,185856754
BS07701	0,20398912	0,2130553	0,22212149	0,18132366	0,17679057	0,19038985	0,167724388	0,190389846	0,185856754
3-32-1	0,06799637	0,00906618	0,05893019	0,01359927	0,06799637	0,06346328	0,018132366	0,902085222	0,407978241
3-23-1	0,18132366	0,19492294	0,18132366	0,14959202	0,13145966	0,14052584	0,113327289	0,149592022	0,113327289
3-22-1	0,14505893	0,15412511	0,18585675	0,11332729	0,08159565	0,23118767	0,049864007	0,077062557	0,258386219
1-33-1	0,18132366	0,19038985	0,18585675	0,14052584	0,22212149	0,13145966	0,149592022	0,222121487	0,104261106
1-22-1	0,17679057	0,17225748	0,17679057	0,18132366	0,17225748	0,09519492	0,199456029	0,208522212	0,031731641

color key

cOD_{620}	0	0,1	0,2	0,3	0,4

4. Conclusion

The bottleneck in the biotin production using *E. coli* strains is the enzyme BioB, which catalyzes the last step of the biosynthetic pathway, the conversion of DTB to biotin. This enzyme was discovered, from previous research, to cause oxidative stress within the cells, inhibiting growth independent of biotin-forming activity when overexpressed.

This thesis used high-throughput genetic engineering, selection, and screening methods to investigate if the creation of strains that are more resistant to oxidative stress would translate into these strains being able to tolerate the expression of higher BioB levels and thus produce more biotin. This was done by introducing controlled levels of genetic mutations in Biosyntia's proprietary biotin-producing *E. coli* strains, selecting phenotypes with higher oxidative stress resistance, and screening those strains' biotin production capabilities.

It was found that the inactivation of the mismatch repair system led to higher levels of rifampicin resistance, indicating that higher rates of mutagenesis were produced using this system. The mutation rate of this system and the accuracy of the rifampicin readout could not be verified as next-generation sequencing of the strains was not performed due to time constraints. It would be beneficial to verify the number of mutations and relate them to the number of rifampicin-resistant colonies to develop a helpful readout for the mutagenesis rate.

The strains where the MMR system was deactivated were selected using 2D gradients with various concentrations of stressors. These selection strains did show higher growth rates and, thus, higher resistances towards these stressors when grown in media with different stressor concentrations; however, this did not translate to higher biotin production capabilities. To further investigate this hypothesis, another biotin production assay and growth assay should be performed with the selected strains, as the biotin production assay did not yield sufficiently reliable data to form a conclusion regarding the higher IPTG levels.

The creation of plasmid diversity using gDNA libraries was abandoned due to the low diversities achieved during this experiment. Optimizing the protocol and investigating whether introducing foreign or endogenous genes could confer beneficial properties to the strain and lead to higher biotin production would be beneficial. Using other high-throughput screening methods like droplet microfluidics fluorescence-assisted screening systems could further enhance the throughput of the screenings, increasing the number of strains that could be tested.

Using cobalt or IPTG-induced BioB expression as selection stressors could lead to different results, as it is unknown whether the chosen stressors in this experiment accurately mimic BioB-induced cellular stress. Using BioB toxicity as a selection stressor could be a more sensible way to generate more BioB-tolerant strains.

5. References

Acevedo-Rocha, C. G., Gronenberg, L. S., Mack, M., Commichau, F. M., & Genee, H. J. (2019). Microbial cell factories for the sustainable manufacturing of B vitamins. In *Current Opinion in Biotechnology* (Vol. 56, pp. 18–29). Elsevier Ltd. https://doi.org/10.1016/j.copbio.2018.07.006

Bali, A. P., Lennox-Hvenekilde, D., Myling-Petersen, N., Buerger, J., Salomonsen, B., Gronenberg, L. S., Sommer, M. O. A., & Genee, H. J. (2020). Improved biotin, thiamine, and lipoic acid biosynthesis by engineering the global regulator IscR. *Metabolic Engineering*, *60*, 97–109. https://doi.org/10.1016/j.ymben.2020.03.005

Brown, S. W., & Kamogawa, K. (1991). The Production of Biotin by Genetically Modified Micro-organisms. In *Biotechnology & genetic engineering reviews* (Vol. 9).

Burkholder, P. R., & Mcveigh, I. (1942). Synthesis of Vitamins by Intestinal Bacteria. *Ibid*, *28*, 285–298. https://www.pnas.org

Cohen, S. N., Chang, A. C. Y., Boyert, H. W., & Hellingt, R. B. (1973). *Construction of Biologically Functional Bacterial Plasmids In Vitro (R factor/restriction enzyme/transformation/endonuclease/antibiotic resistance)* (Vol. 70, Issue 11).

Cramer, J. D., & Jarrett, J. T. (2018). Purification, Characterization, and Biochemical Assays of Biotin Synthase From Escherichia coli. In *Methods in Enzymology* (Vol. 606, pp. 363–388). Academic Press Inc. https://doi.org/10.1016/bs.mie.2018.06.003

Dasgupta, A. (2019). Biotin. In *Biotin and Other Interferences in Immunoassays* (pp. 17–35). Elsevier. https://doi.org/10.1016/B978-0-12-816429-7.00002-2

Dunn, M. F. (2019). Vitamin Formation from Fatty Acid Precursors. In *Biogenesis of Fatty Acids, Lipids and Membranes* (pp. 259–271). Springer International Publishing. https://doi.org/10.1007/978-3-319-50430-8_24

Eggersdorfer, M., Laudert, D., Létinois, U., McClymont, T., Medlock, J., Netscher, T., & Bonrath, W. (2012). One hundred years of vitamins - A success story of the natural sciences. In *Angewandte Chemie - International Edition* (Vol. 51, Issue 52, pp. 41–41). https://doi.org/10.1002/anie.201205886

Fantino, J. R., Py, B., Fontecave, M., & Barras, F. (2010). A genetic analysis of the response of Escherichia coli to cobalt stress. *Environmental Microbiology*, *12*(10), 2846–2857. https://doi.org/10.1111/j.1462-2920.2010.02265.x

Fasnacht, M., & Polacek, N. (2021). Oxidative Stress in Bacteria and the Central Dogma of Molecular Biology. In *Frontiers in Molecular Biosciences* (Vol. 8). Frontiers Media S.A. https://doi.org/10.3389/fmolb.2021.671037

Fleming, A. M., & Burrows, C. J. (2020). On the irrelevancy of hydroxyl radical to DNA damage from oxidative stress and implications for epigenetics. *Chemical Society Reviews*, *49*(18), 6524–6528. https://doi.org/10.1039/d0cs00579g

Gironés-Vilaplana, A., Villanõ, D., Marhuenda, J., Moreno, D. A., & Garciá-Viguera, C. (2017). Vitamins. In *Nutraceutical and Functional Food Components: Effects of Innovative Processing Techniques* (pp. 159–201). Elsevier Inc. https://doi.org/10.1016/B978-0-12-805257-0.00006-5

Hassan, H. M., & Fridovich, I. (1978). Superoxide radical and the oxygen enhancement of the toxicity of paraquat in Escherichia coli. *Journal of Biological Chemistry*, *253*(22), 8143–8148. https://doi.org/10.1016/s0021-9258(17)34373-9

Hsieh, P. (2001). Molecular mechanisms of DNA mismatch repair. In *Mutation Research* (Vol. 486).

Ifuku, O., Koga, N., Haze, S. I., Kishimoto, J., Arai, T., & Wachi, Y. (1995). Molecular Analysis of Growth Inhibition Caused by Overexpression of the Biotin Operon in Escherichia Coli. *Bioscience, Biotechnology, and Biochemistry*, *59*(2), 184–189. https://doi.org/10.1271/bbb.59.184

Leonard, S., Gannet, P. M., Rojanasakul, Y., Schwegler-Berry, D., Castranova, V., Vallyathan, V., & Shi, X. (1998). Cobalt-mediated generation of reactive oxygen species and its possible mechanism. *Journal of Inorganic Biochemistry*, *70*, 239–244.

Majtan, T., Frerman, F. E., & Kraus, J. P. (2011). Effect of cobalt on Escherichia coli metabolism and metalloporphyrin formation. *BioMetals*, *24*(2), 335–347. https://doi.org/10.1007/s10534-010-9400-7

Nur, I., Munna, M. S., & Noor, R. (2014). Study of exogenous oxidative stress response in Escherichia coli, Pseudomonas spp., Bacillus spp., and Salmonella spp. *Turkish Journal of Biology*, *38*(4), 502–509. https://doi.org/10.3906/biy-1312-93

Parveen, N., & Cornell, K. A. (2011). Methylthioadenosine/S-adenosylhomocysteine nucleosidase, a critical enzyme for bacterial metabolism. In *Molecular Microbiology* (Vol. 79, Issue 1, pp. 7–20). https://doi.org/10.1111/j.1365-2958.2010.07455.x

Py, B., & Barras, F. (2010). Building Feg-S proteins: Bacterial strategies. In *Nature Reviews Microbiology* (Vol. 8, Issue 6, pp. 436–446). Nature Publishing Group. https://doi.org/10.1038/nrmicro2356

Reyda, M. R., Dippold, R., Dotson, M. E., & Jarrett, J. T. (2008). Loss of Iron Sulfur Clusters from Biotin Synthase as a Result of Catalysis Promotes Unfolding and Degradation †. In *Arch Biochem Biophys* (Vol. 471, Issue 1).

Sirithanakorn, C., & Cronan, J. E. (2021). Biotin, a universal and essential cofactor: Synthesis, ligation and regulation. In *FEMS Microbiology Reviews* (Vol. 45, Issue 4). Oxford University Press. https://doi.org/10.1093/femsre/fuab003

Soetaert, Wim., & Vandamme, E. J. (2010). *Industrial biotechnology : sustainable growth and economic success*. Wiley-VCH.

Streit, W. R., & Entcheva, P. (2003). Biotin in microbes, the genes involved in its biosynthesis, its biochemical role and perspectives for biotechnological production. In *Applied Microbiology and Biotechnology* (Vol. 61, Issue 1, pp. 21–31). Springer Verlag. https://doi.org/10.1007/s00253-002-1186-2

Wang, Y., Liu, L., Jin, Z., & Zhang, D. (2021). Microbial Cell Factories for Green Production of Vitamins. In *Frontiers in Bioengineering and Biotechnology* (Vol. 9). Frontiers Media S.A. https://doi.org/10.3389/fbioe.2021.661562

Weinstein, Z. B., & Zaman, M. H. (2018). *Evolution of Rifampin Resistance in Escherichia coli and Mycobacterium smegmatis Due to Substandard Drugs*. https://journals.asm.org/journal/aac

Woolston, B. M., Edgar, S., & Stephanopoulos, G. (2013). Metabolic engineering: Past and future. *Annual Review of Chemical and Biomolecular Engineering*, *4*, 259–288. https://doi.org/10.1146/annurev-chembioeng-061312-103312

Appendix

cPCR Protocol for gDNA Libraries

The colonies from the transformation were verified using the following DreamTaq Master Mix for colony PCR and PCR cycler program. After running the PCR program 8 µL of reaction mix for each colony were run on a 1% agarose gel at 100V for 40 minutes.

DreamTaq Master Mix for colony PCR

Components	Comment	Volume
FW primer (oBS570)	10 µM	1.5 µL
RV primer (oBS5762)	10 µM	1.5 µL
Template	Toothpick stab of colony or pre-diluted in sterile H2O	0-20 µL
Buffer/Polymerase/dNTP		20 µL
dH2O		20 µL
Total		50 µL

PCR cycler program

Step	T [°C]	Minutes:Seconds	Go to	Loops
1	95	3:00		
2	95	0:15		
3	57	0:20		
4	72	3:00	2	32
5	72	5:00		
6	8	hold		

Oligo List

Name	Sequence	Comment
oBS00570	GGATAACCGTATTACCGCCTTTG	RV primer cPCR for pBS1897
oBS05762	TCGAGGTCGACGGTATCG	FW primer cPCR for pBS1897

Strain List

Strain Name	Host	Description	Genotype	Plasmids
BS04755	*E. coli*	Strain used for gDNA libraries	WT(BW25113); bioB::[FRT]{bioA,bioF}; iscR(H107Y); [bioA,bioF,bioC,bioD]::[FRT]{ybhB,ybhU}; bioH::[FRT]{rpnA,yhgH}; metE(C645A); cysP(S216N); redacted mutation(s)	-
BS07084	*S. cerevisiae*	Strain used for gDNA libraries	WT(NCYC 3608); MATα SUC2 gal2 mal2 mel flo1 flo8-1 hap1 ho bio1 bio6	-
BS07701	*E. coli*	Strain used for MMR inactivation experiment (repeat experiment, shortened version)	WT(BW25113); bioB::[FRT]{bioA,bioF}; iscR(H107Y); [bioA,bioF,bioC,bioD]::[FRT]{ybhB,ybhU}; bioH::[FRT]{rpnA,yhgH}; redacted mutation(s); +[T5lacO/bioB,FRT,TetR,FRT]{torS,torT}; +[T5lacO/bioB,FRT,CamR,FRT]{icd,ymfD}; Δ[FRT,CamR]{torS,torT}; Δ[FRT,TetR]{icd,ymfD}; +[T5lacO/bioB,FRT,KanR,FRT]{ypjC,ileY};	pBS2215

			Δ[FRT,KanR]{ypjC,ileY}; FRT::[T5lacO/bioB,FRT,CamR,FRT]{rpnA,yhgH}; Δ[FRT,CamR]{rpnA,yhgH}; redacted mutation(s) [T5lacO/BioB,FRT,Kan,FRT]::ycgR; Δ[FRT,KanR]{emtA,ymgE}; fur(K14I)	
BS07210	*E. coli*	Bioassay strain	WT(BW25113); yigM::[FRT,KanR,FRT]{rraB,yigN}; BS01059_TBD; [bioA,bioF,bioC,bioD]::[FRT, camR, FRT]{ybhB,ybhU}	pBS3014
BS08072	*E. coli*	Strain used for transformation	WT(TOP10)	-
BS06757	*E. coli*	Strain used in MMR inactivation experiment	WT(BW25113); bioB::[FRT]{bioA,bioF}; iscR(H107Y); [bioA,bioF,bioC,bioD]::[FRT]{ybhB,ybhU}; bioH::[FRT]{rpnA,yhgH}; redacted mutation(s)+[T5lacO/bioB,FRT,TetR,FRT]{torS,torT}; +[T5lacO/bioB,FRT,CamR,FRT]{icd,ymfD}; Δ[FRT,CamR]{torS,torT}; Δ[FRT,TetR]{icd,ymfD}; +[T5lacO/bioB,FRT,KanR,FRT]{ypjC,ileY}; Δ[FRT,KanR]{ypjC,ileY}; FRT::[T5lacO/bioB,FRT,CamR,FRT]{rpnA,yhgH}; Δ[FRT,CamR]{rpnA,yhgH}; redacted mutation(s) [T5lacO/BioB,FRT,Kan,FRT]::ycgR; Δ[FRT,KanR]{emtA,ymgE}	pBS2215 pBS1565
BS04704	*E. coli*	Strain used for plasmid prep	-	pBS2215

Mutant Strain Key

The stressor plate concentration refers to the stressor concentration of the plate the strain was picked from in experiment 2.2.3 Selection for Improved Phenotypes. Selection round refers to the batch of selection plates, the strain was picked from. Round 1 plates were not induced, Round 3 plates were induced twice.

Label	**Parent strain**	**Selection Round**	**Stressor Plate Concentration**
1-22-1	BS07701	1	50 µM PQ, 0.5 mM H_2O_2
1-23-1	BS07701	1	50 µM PQ, 1 mM H_2O_2
1-32-1	BS07701	1	100 µM PQ, 0.5 mM H_2O_2
1-33-1	BS07701	1	100 µM PQ, 1 mM H_2O_2
3-22-1	BS07701	3	50 µM PQ, 0.5 mM H_2O_2
3-23-1	BS07701	3	50 µM PQ, 1 mM H_2O_2
3-31-1	BS07701	3	100 µM PQ, 0 mM H_2O_2
3-32-1	BS07701	3	100 µM PQ, 0.5 mM H_2O_2

Plasmid List

Name	Description; Genotype	Resistance	Replication Origin
pBS1897	Used as backbone for gDNA libraries; pZE21 plasmid in the pBR322 family.	Kan	pBR322
pBS2215	Plasmid carrying DAM (DNA methylase) for MMR inactivation experiment; pMA7-sacB, Amp, original Hao Luo pBS136 with introduced I-sceI selfcuring.	Amp	ColE1
pBS1565	DTB production plasmid; apFAB346/[bioFADGC]	Kan	p15A
pBS3014	Super resistance plasmid with high mCherry expression and kanamycin, ampicillin, spectinomycin, tetracyclin, chloramphenicol and zeocin resistances; apFAB071/mCherry	Amp, Spec, Cam, Zeo, Tet, Kan	pBR322

Media Recipes

10xmMOPS

For 5L:

In a 5 L plastic measuring cup with a stir bar add the following to ~1500 mL sterile milliQ H_2O:

Component	FW	grams	M
MOPS buffer[1]	2093	4186	4
Tricine[2]	1792	3585	4

Add 10 M KOH to a final pH of 7.4 (50 to 100 mL) and bring the total volume to 2200 mL. Use a 500 mL measuring glass to be sure that the volume is exact. Make fresh $FeSO_4$ solution and add it to the MOPS/Tricine solution:

Component	FW	grams	H2O vol (ml)	stock conc. (M)
FeSO4•7H2O	278	14	50	1

Add the following solutions to the MOPS/tricine/$FeSO_4$ solution (see below how to make each of these):

Component	Molarity	BS recipes	Volume (5L)
NH4Cl	1.9 M		250 ml
K2SO4	0.276 M		50 ml
CaCl2•2H2O	1 M		25 µl
MgCl2	2 M		13.125 ml
NaCl	5 M		500 ml
Micronutrient stock (50 ml tube in freezer)[3]	—	mBS043	1 ml
Autoclaved milliQ H2O	(55 M)		1875 ml
Add H2O up to 5 L	—		5000 ml

At the end filter sterilize the 10 X MOPS into autoclaved 500 mL bottles using the vacuum pump and filter: Nalgene: flow bottle top filter 0.2 µm membrane 75 mm diameter. Store at -20°C.

2x YP

Ingredients	Brand	Mass for 800 mL
Yeast Extract	(Bacto™)	20 g
Peptone	(Gibco Bacto™)	40 g

Add 600 mL of deionized H2O to a 1L plastic measuring cup. Stir with a magnet. Add the components one at a time and let them dissolve. Fill water up to 800 mL and distribute in 5x 160 ml in 250 mL bottles. Autoclave 20 min at 120°C to sterilize.

YPD medium

Volume [mL]	Components
160	2xYP
40	20% glucose solution
200	H2O

LB medium

LB Broth (Lennox) Vendor: Sigma- Aldrich Cat.no.: L3022

Note: Adapted from the protocol used at CFB, 160601

Mix 10 g of LB Broth (Lennox) in 400 mL dH_2O. Shake until the solutes have dissolved. Adjust pH to 7.0 with 2 M NaOH. Adjust the volume of the solution to 500 mL with dH_2O. Aliquot in 200 mL bottles or as needed. Sterilize by autoclaving for 20 min.

SOC medium

Compound	Per liter	Final concentration
Tryptone (CAT#)	20 g	2%
Yeast extract (CAT#)	5 g	0.5%
NaCl (CAT#)	0.5 g	10 mM

Add components and 950 mL deionized H_2O to a 1 L plastic measuring cup. Stir using a magnet until the solutes have dissolved. Add 10 mL of a 250 mM KCl solution (final concentration 2.5 mM). Adjust pH to 7.0 with 5 M NaOH (approx. 0.2 ml). Adjust the volume to 1 liter with deionized H2O using a 1 L measuring glass. Aliquot into 50 mL bottles. Sterilize by autoclaving for 20 min. After autoclaving, in a sterile environment, add 0.9 mL of 20 % glucose solution to every 100 mL SOC bottle (final concentration 10 mM).

Stock Solutions

500X Vitamin solution

Mix compounds together and filter sterilize. Store vitamin solution as aliquots at -20°C. KOH solution is prepared by dissolving 0.1247 grams of KOH (90% Reagent Grade from Sigma) in 100 mL of water.

Compound	Weight [g]	Volume [mL]	Directions
calcium pantothenate (C8731, Sigma-Aldrich)	238	25	Dissolve in water
p-aminobenzoic acid (A9878, Sigma-Aldrich)	69	25	Dissolve in 0.02 M KOH
4-hydroxybenzoic acid (240141, Sigma-Aldrich)	69	25	Dissolve in 0.02 M KOH
2,3-dihydroxybenzoic acid (126209, Sigma-Aldrich)	77	25	Dissolve in 0.02 M KOH

20% Glucose solution

Weigh out 20g of glucose (dextrose). Add to 70ml of H_2O. Dissolve by stirring, use heat if necessary. Once the sugar has dissolved, bring the volume of the mix up to 100ml total. Autoclave.

0.132 M K2HPO4

Add 400 mL deionized H_2O to the 1 L plastic measuring cup stir with magnet. Add 11,49 g K_2HPO_4 into the 400 mL and stir until K_2HPO_4 is completely dissolved. Adjust the volume to 500 mL by transferring the solution to the 500 mL measuring glass. Adjust with deionized H_2O. Transfer to 500 mL Blue Cap Bottle. Sterilize by autoclaving for 20 min. Store at RT or 5°C.

Compound	**Per 500 mL**	**Final concentration**
K2HPO4	11,5 g	0.132 M

0.5 M EDTA

This solution is prepared by dissolving 1.86 g of $Na_2EDTA•2H2O$ in 10 mL of distilled, deionized water and adjusting the pH to 8.0 with sodium hydroxide. Solution is subsequently filter sterilized and stored at or RT or 5°C.

cAMP 50mM

To prepare ten 1.8 mL cAMP stocks dissolve 0.296 g of cAMP in 18 mL of sterile water and transfer 1.8mL aliquots into 2 mL Eppendorf tubes. Store in -20°C freezer.

DTB solution

1 g/L solution of DTB was prepared by dissolving 0.1 g of DTB in sterile water. The stock solution was stored at 5°C. For use in experiments, the solution was diluted 1:1000, e.g. 300 µL in 300 mL.

HABA solution

The HABA solution is prepared from two individual solutions that are mixed. To make the solutions you need:

- NaOH solution - 10 mM
- HABA reagent (1,5 mg/mL of HABA (4'-hydroxyazobenzene-2-carboxylic acidin) in 10 mM NaOH)
- 1x PBS buffer solution (8 g/L NaCl; 2 g/L KCl; 2,68 g/L Na*2*HPO*4* × 7H*2*O; 0,24 g/L KH*2*PO*4*)

Solutions:

A:

- 25 mL of NaOH solution
- 125 mL 1× PBS buffer
- 50 mL HABA reagent

B:

- 125 mg of streptavidin dissolved in 62,5 mL PBS buffer

Once the solutions A and B are properly mixed and all ingredients are fully dissolved, they are combined to form the HABA solution.

Antibiotic Stocks

These are the antibiotic concentrations used for liquid and solid media in all experiments.

Antibiotic	final antibiotic conc.	Antibiotic stock to medium
ampicillin	100 µg/ml	1µl to 1 mL
kanamycin	50 µg/ml	1µl to 1 mL
rifampicin	100 µg/mL	4µL to 1 mL

Kanamycin stock solution

Kanamycin monosulfate C18H36N4O11 · H2O4S, MW 582.58, CAS Number 25389-94-0,
SDS: https://www.sigmaaldrich.com/DK/en/sds/SIAL/A1593

Vendor: Sigma-Aldrich, Cat.no.: BP861

Chemical	Per 20 mL	Final concentration
Kanamycin	500 mg	50 mg/mL

Add 50 mL Falcon tube to a 100 mL plastic measuring beaker and weigh 500 mg of kanamycin into the 50 mL Falcon tube. Add 10 mL dH_2O (sterile) using a 10 mL serological pipette. Mix thoroughly until kanamycin is completely dissolved (if needed use vortex). Use a 20 mL sterile syringe and a sterile syringe filter (w/0.2µm polyethersulfone to filter sterilize) to sterilize the 10 mL into a sterile 50 mL Blue Cap Bottle. Aliquote 1.1 mL solution into 1.5 mL sterile Eppendorf tubes. Store in -20°C freezer.

Ampicillin stock solution

Ampicillin C16H19N3O4S · 3H2O**,** MW 403.45**,** CAS Number: 7177-48-2
SDS: https://www.sigmaaldrich.com/DK/en/sds/SIAL/BP861

Vendor: Sigma-Aldrich, Cat.no.: A1593

Chemical	Per 20 mL	Final concentration
Ampicilin	2.0 g	100 mg/mL

Add the 50 mL Falcon tube to the 100 mL plastic measuring beaker and weigh 2 g of ampicillin into the 50 mL Falcon tube. Add 20 mL dH_2O using the 25 mL serological pipette. Mix thoroughly until ampicillin is completely dissolved (if needed use vortexer). Use a 20 mL sterile syringe and a sterile syringe filter (w/0.2µm polyethersulfone to filter-sterilize) to sterilize the 10 mL into a sterile 50 mL Blue Cap Bottle. Aliquote 1.1 mL solution into 1.5 mL sterile Eppendorf tubes. Store in -20°C freezer.

Rifampicin stock solution

Rifampicin C43H58N4O12, CAS Number 13292-46-1,
SDS: https://www.sigmaaldrich.com/DK/en/sds/sigma/r3501

Vendor: Sigma-Aldrich, Cat.no.: R7382

Chemical	Per 20 mL	Final concentration
Rifampicin	0.5 g	25 mg/mL

Add 0.5 g of Rifampicin to 20 mL of sterile water, add HCL solution as needed to lower pH (~6) until the powder is completely dissolved. Store away from light at -20°C.

Agar Plates

LB agar plates

LB Broth (Lennox) Vendor: Sigma- Aldrich Cat.no.: L3022

Note: Adapted from the protocol used at CFB, 160601

Preparing LB agar:

Mix 12 g of LB Broth (Lennox) in 400 mL dH_2O. Stir until the solutes have dissolved. Adjust pH to 7.0 with 2 M NaOH if necessary. Add 12 g Agar. Stir until the solutes have dissolved. Adjust the volume to 600 mL by adding the solution to the 500 mL measuring glass and adjusting the volume to 500 mL. Pour it back into the plastic measuring cup. Add 100 mL dH_2O. Distribute 300 mL into a 500 mL Blue Cap Bottle. Sterilize by autoclaving for 20 min.

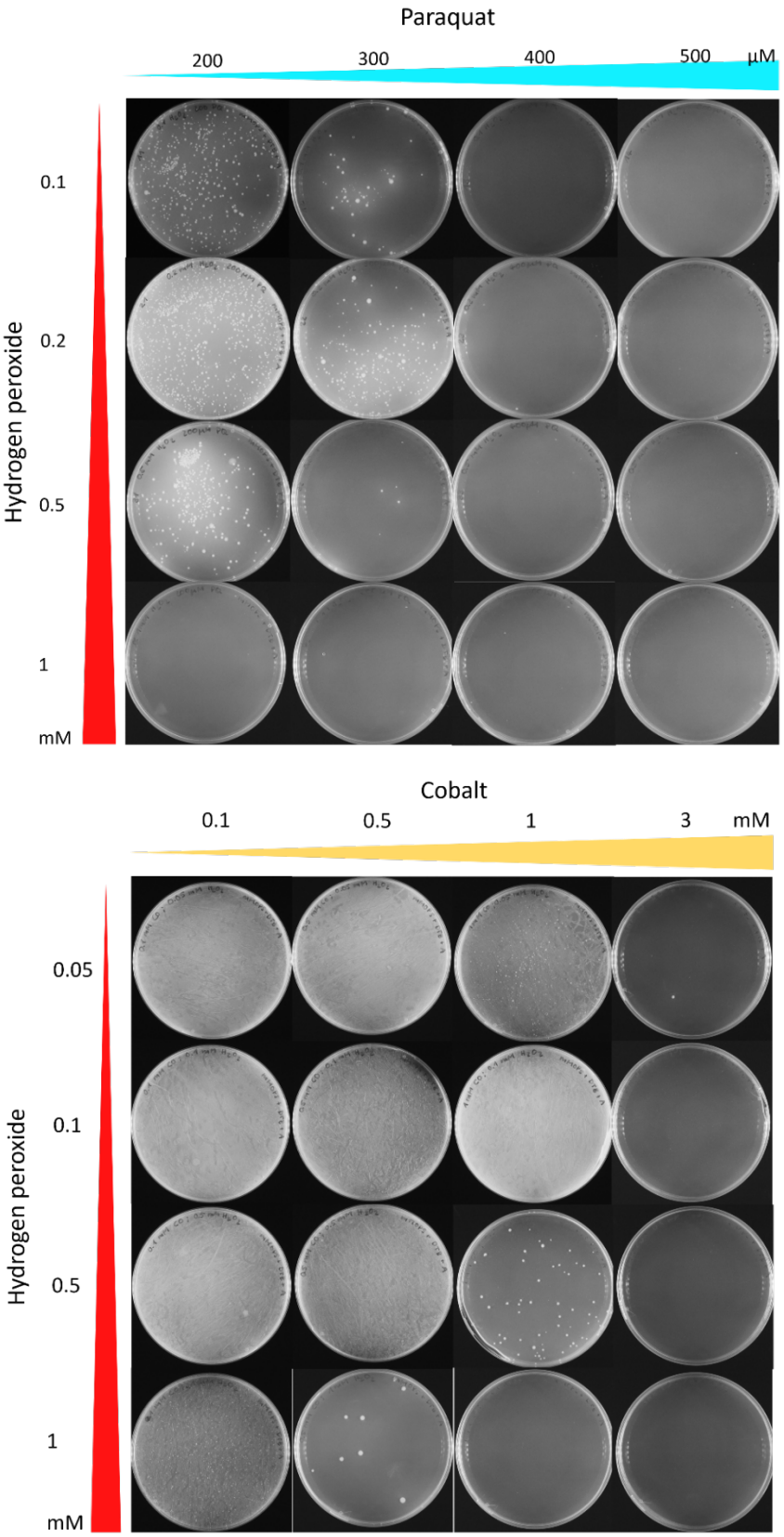
Preparing LB agar plates:

Melt LB agar in the microwave or use freshly autoclaved LB agar. The agar needs to be fully liquid (no clumps at all). Leave the fluid LB agar until a temperature of app. 50-60 degrees is reached. The temperature is about right when you are just able to hold on the LB agar for longer than 30 seconds. Mix the LB agar with the proper antibiotic. Shake the bottle carefully to avoid air bubbles and pour into sterile petri dishes. App. 15-20 mL LB agar should be used per petri dish. Leave the Petri dishes to dry with a half-open lid in the sterile clean bench. When the LB agar has solidified, label each petri dish with the date and antibiotic and store it in the bag at 5°C.

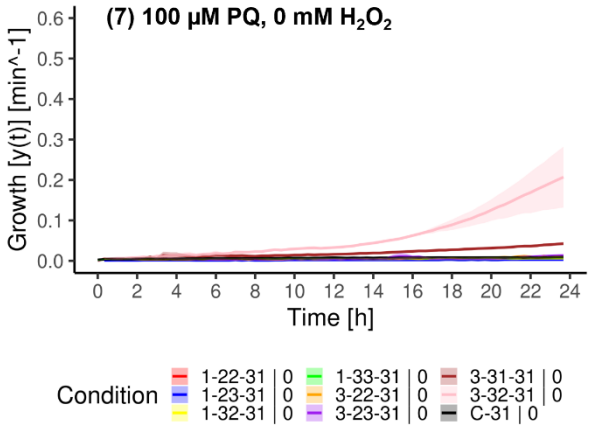
mMOPS plates

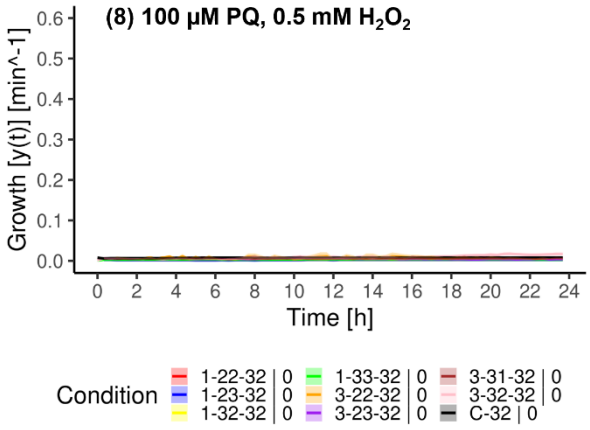
Heat 200 mL of 3% agar in the microwave until liquid, mix with 200 mL of 2xmMOPS solution, add antibiotics as needed, mix, and pour about 20 mL of the medium into a sterile petri dish (will yield about 20 Petri dishes) and let it solidify with a half-open lid under the lab bench in a sterile environment. Place the plates in a bag, label, and store at 5°C until use.

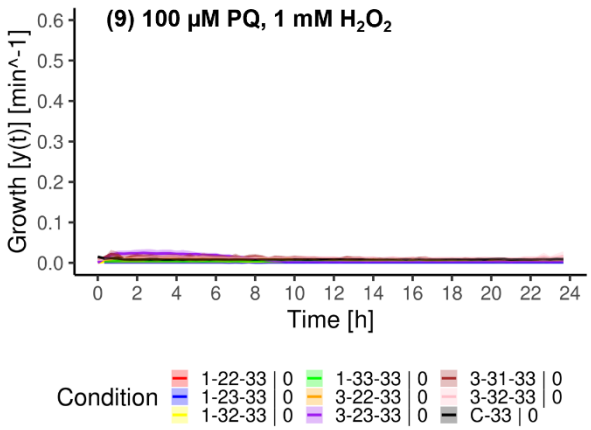
2D Selection Gradients



Growth Experiment with Stressors







Biotin Standard Preparation

Prepare a 100 mM biotin stock solution

- Weigh out 0.24431 g biotin

- Transfer carefully into 10 mL volumetric flasks

- Rinse the weighing cup with DMSO by pipetting and pour into the volumetric flask

- Fill the volumetric flask to the marking with DMSO

Prepare a 1 mM dilution from the 100 mM biotin stock solution

- Pipette 100 µL of the 100 mM biotin stock solution into a 10 mL volumetric flask

- Fill the volumetric flask to the marking with mMOPS

Prepare a 100 µM biotin working standard from the 1 mM biotin solution:

- Pipette 2.5 mL of the 1 mM biotin solution into a 25 mL volumetric flask

- Fill the volumetric flask to the marking with mMOPS

- Pipette 0.5 mL of this solution into an Eppendorf tube and hand over to the Analytical Chemistry department for confirmatory analyses before proceeding to the next step (store the solution in a blue cap bottle at -20 oC until the confirmation has been received)

Prepare the biotin standards for the biotin assay according to the biotin standards layout below (use 15 mL centrifuge tubes):

- Take 24 centrifuge tubes (á 15 mL) and label them with the numbers 1-24

- Pipette the correct volume (according to the biotin standards layout) of 1× MOPS into each tube

- Pipette the correct volume (according to the biotin standards layout) of the 100 µM biotin stock solution into each tube - throw out the remaining amount of the 100 µM biotin stock solution

- Shake each tube well

- Validate the standards by running a Bioassay test comparing the new standards (in triplicates) with the old standards

- Aliquot 1 mL of each standard into a DWP according to Scheme 1

- Aliquot 100 µL of each standard into PCR plates (requires 12 PCR plates) according to Scheme 1

- Seal the PCR plates with aluminum foil and store at -20°C.

Biotin Standards Layout

Std number	BTN (µM)		Stock (100 µM) (mL)	Pipet type/setting	MOPS 1x (mL)	Pipet type/setting
1	0	0	0.000	-	5.000	P5
2	2	0.1	0.100	P200	4.900	P5
3	3	0.15	0.150	P200	4.850	P5
4	4	0.2	0.200	P200	4.800	P5
5	5	0.25	0.250	P1000	4.750	P5
6	6	0.3	0.300	P1000	4.700	P5
7	8	0.4	0.400	P1000	4.600	P5
8	10	0.5	0.500	P1000	4.500	P5
9	12	0.6	0.600	P1000	4.400	P5
10	14	0.7	0.700	P1000	4.300	P5
11	16	0.8	0.800	P1000	4.200	P5
12	18	0.9	0.900	P1000	4.100	P5
13	20	1	1.000	P1000	4.000	P5
14	22	1.1	1.100	P5	3.900	P5
15	25	1.25	1.250	P5	3.750	P5
16	30	1.5	1.500	P5	3.500	P5
17	35	1.75	1.750	P5	3.250	P5
18	40	2	2.000	P5	3.000	P5
19	45	2.25	2.250	P5	2.750	P5
20	50	2.5	2.500	P5	2.500	P5
21	60	3	3.000	P5	2.000	P5
22	70	3.5	3.500	P5	1.500	P5
23	80	4	4.000	P5	1.000	P1000
24	90	4.5	4.500	P5	0.500	P1000

1	9	17	1	9	17	1	9	17	1	9	17
2	10	18	2	10	18	2	10	18	2	10	18
3	11	19	3	11	19	3	11	19	3	11	19
4	12	20	4	12	20	4	12	20	4	12	20
5	13	21	5	13	21	5	13	21	5	13	21
6	14	22	6	14	22	6	14	22	6	14	22
7	15	23	7	15	23	7	15	23	7	15	23
8	16	24	8	16	24	8	16	24	8	16	24