Long-term response to recombinant gene expression on chromosome-level of genome-integrated *E. coli* expression systems

Master thesis

Christoph Köppl, BSc

Matr. Nr.: 01440539

Supervisor:
Assoc. Prof. Dr. Gerald Striedner

Universität für Bodenkultur, Wien
Institute of Bioprocess Science and Engineering

Wien, 11.08.2020

Eidesstaatliche Erklärung

Ich erkläre eidesstattlich, dass ich die Arbeit selbständig angefertigt, keine anderen als die angegebenen Hilfsmittel benutzt und alle aus ungedruckten Quellen, gedruckter Literatur oder aus dem Internet im Wortlaut oder im wesentlichen Inhalt übernommenen Formulierungen und Konzepte gemäß den Richtlinien wissenschaftlicher Arbeiten zitiert, durch Fußnoten gekennzeichnet bzw. mit genauer Quellenangabe kenntlich gemacht habe.

_____

Unterschrift

Acknowledgements

First, I want to thank my parents. Without their continuous support and encouragement this master thesis would have not been possible. A big “Thank you” also to Assoc. Prof. Dr. Gerald Striedner and Dipl.-Ing. Dr. Monika Cserjan for their guidance and the possibility to write my master thesis in this working group. Especially, I want to thank Artur for his excellent supervision and advice working on this thesis. Last but not least, I want to thank all of my laboratory colleagues for always having an open ear for my questions.

Table of Content

1 Abstract

State of the art pharmaceutical production of recombinant proteins in bacteria rarely features continuous fermentation due to instability of production strains. Studies have shown that the overall implementation of continuous processing has the potential to significantly decrease production costs and environmental footprint of biotechnological products. Main bottleneck on the way to a fully continuous process is the genetic stability of the bacterial production systems. During protein production the cells are exposed to a high metabolic burden, which in turn leads to high mutation rates. This master thesis evaluates the potential of genome integrated *E. coli* expression systems for continuous fermentation and production. Evaluation was performed with the model proteins Green Fluorescent Protein, an easy to produce protein, and the antigen binding fragment FTN2, a challenging protein. Three BL21 strains were compared in chemostat cultivations, two of them utilizing phage derived T7 RNAP in combination with T7 promotors, while the third tested variation was BL21 using host native RNAP and a novel promotor system. In addition, one of the strains using T7 RNAP was mutated through directed evolution to improve its genetic stability. The results indicate that the novel production system is very stable and suited for long term continuous cultivation and production. The mutant strains obtained via directed evolution show potential for the production of challenging proteins in continuous fermentation. Due to inherent differences between B and K *E. coli* strains it was not possible to implement the production system in HMS174. This master thesis highlights the need for further research in this field, as there is urgent need for universally applicable production systems with high genetic stability. Additional studies have to be performed to further evaluate the potential of the generated stable mutant production strains for continuous and standard production process operation.

Trotz des immensen Einsparungspotentials hinsichtlich der Produktionskosten und des ökologischen Fußabdrucks werden in der rekombinanten Proteinproduktion mit bakterieller Zellkultur selten kontinuierliche Fermentationsprozesse eingesetzt. Dies ist hauptsächlich durch die genetische Instabilität konventioneller Produktionsstämme bedingt, die während der Produktion hohem metabolischen Stress ausgesetzt sind und erhöhte Mutationsraten aufweisen. Diese Masterarbeit evaluiert das Potenzial von genomintegrierten *E. coli* Expressionssystemen für die kontinuierliche Produktion. Die Beurteilung erfolgte anhand der Modelproteine Green Fluorescent Protein, einem leicht zu produzierenden Protein, und dem Antigen Binding Fragment FTN2 eines humanen Antikörpers, das schwer zu produzieren ist. Durch Chemostat-Fermentationen wurde die Eignung von drei *E. coli* Produktionsstämmen für die kontinuierliche Produktion evaluiert. Verglichen wurden drei BL21 Stämme, zwei davon verwenden die aus dem T7 Phagen stammende T7 RNAP in Kombination mit dem T7 Promotor, während die dritte getestete Variante native RNAP und ein neuartiges Promotorsystem anwendet. Einer der zwei verwendeten T7 RNAP Stämme wurde zuvor durch gelenkte Evolutionsexperimente mutiert, um seine genetische Stabilität zu verbessern. Die Ergebnisse lassen darauf schließen, dass die mutierte Variante ausgezeichnete Stabilität aufweist und gut geeignet für die kontinuierliche Fermentation von schwer zu produzierenden Proteinen ist. Aufgrund von inhärenten Unterschieden zwischen B und K Stämmen war es nicht möglich das neuartige Produktionssystem in den K Stamm HMS174 zu implementieren. Diese Masterarbeit hebt die Notwendigkeit für weitere Forschung in diesem Gebiet hervor, da Bedarf für universell einsetzbare Produktionssysteme mit hoher genetischer Stabilität besteht. Weiters müssen zusätzliche Studien durchgeführt werden, um das Potential der generierten stabilen Mutanten für die kontinuierliche Fermentation zu evaluieren.

2 Introduction

Since the first recombinant therapeutic proteins have been produced in *E. coli* more than 30 years ago, they have acquired enormous importance in clinical applications (Graumann and Premstaller, 2006). From 2011 to 2016, 62 new recombinant therapeutic proteins have been approved by the FDA (Lagassé et al., 2017), expressing the ever growing applications and demand for recombinant therapeutic proteins. Despite the advance of mammalian cell culture for the production of post-translationally modified products, *E. coli* is still the production organism of choice for non-glycosylated proteins due to its fast growth, high product yields and easy manipulation (Graumann and Premstaller, 2006). In addition to these characteristics, *E. coli* can be cultivated easily in high cell density cultures with a theoretical limit of cell densities of up to 200 g cell dry mass per litre (CDM/L)(Rosano and Ceccarelli, 2014). Contrary to animal cell culture, where complex media formulations and serum-containing media are still in use (Yao and Asayama, 2017), *E. coli* can be cultivated in media produced from readily available and inexpensive components (Rosano and Ceccarelli, 2014). This makes *E. coli* a reliable and easy to handle production platform for a broad range of recombinant therapeutic proteins.

2.1 Commonly used expression systems in *E. coli*

Despite the numerous promotor systems available for protein production in *E. coli*, only a handful of them are broadly used. This is due to characteristics that effective promotor systems must fulfil. In principle, effective promotor systems should be sufficiently strong, tightly regulated, easily introducible into the host organisms, effectively and cheaply inducible but should not be influenced by the compounds present in the culture media. (Terpe, 2006)

2.1.1 Strong expression systems

The *E. coli* pET expression system is the most commonly used bacterial system. It is based on the phage-derived inducible T7 RNA polymerase (RNAP) that is more than five times faster than the *E. coli* host RNAP. The pET system allows for fast and high-level production of recombinant protein. The gene for the T7 RNAP is located in the *E. coli* chromosome (e.g.: BL21(DE3)) and under control of the *lac* promotor derivative L8-UV5. This version of the *lac* promotor carries multiple point mutations which make it strongly inducible with IPTG, decrease its dependency on cAMP and therefore facilitate its induction in the presence of glucose. Upon induction, the *lac* promotor is activated and production of T7 RNAP is started. This RNAP then starts to produce RNA transcripts of the target gene, which is under control of a T7 promotor. Despite all efforts to create a tightly regulated expression system, the T7 system shows basal expression, which can be a problem when expressing toxic proteins. (Briand et al., 2016) (Terpe, 2006)

Another disadvantage when using the T7 system is the T7 RNAP itself. Because it is only needed for the production of the recombinant protein and not necessary for the transcription

of any host cell proteins, it is a hotspot for mutations. It has been observed, that decreasing target protein production levels are attributable to loss of function mutations in the genomic sequence of the T7 RNAP (Vethanayagam and Flower, 2005). Another flaw with the use of very strong expression systems like the T7 RNAP is that it is often not possible to express challenging proteins like antibody fragments (Schuller et al., 2020). Due to the strength of the system, especially when paired with high-copy-number plasmids, a high metabolic load is exerted on the host cells, which can lead to reduced yield, decreased cell growth and might ultimately even lead to cell death (Schuller et al., 2020).

In addition, inclusion body (IB) formation may pose a problem in high level protein expression. This occurs due to interactions between hydrophobic stretches of protein, when present in high concentrations, or the conditions in the *E. coli* cytoplasm itself, which may divert significantly in pH, redox potential, available folding mechanisms, osmolarity and cofactors from the organism the protein of interest (POI) is derived from. All these factors can lead to aggregation or even degradation of the POI. (Rosano and Ceccarelli, 2014)

In addition to exogenous expression systems like the T7 RNAP, there exist a number of different native host RNAP expression systems that have been utilized for recombinant protein production. For instance, the synthetic *tac* or *trc* promotor utilize host RNAP and can be used for the effective production of recombinant proteins. They consist of the -35 region of the *trp* promotor combined with the -10 region of the *lac* promotor. Both of these promotors are reasonably strong and allow accumulation of recombinant protein to levels of up to 15-30% of the total cellular protein. (Terpe, 2006)

2.1.2 Weak expression systems

However, high level POI production may especially become problematic when expressing proteins, which are toxic for the host cells. Expression of toxic proteins may cause target protein degradation or misfolding and accumulation, leading to high levels of cellular stress (Quick and Wright, 2002, Tomatis et al., 2019). To avoid this problem, weaker promoter systems can be applied. One commonly used promotor system is based on the *lac* promotor. This promotor is very weak and leaky in its unaltered form and rarely applied to overexpress recombinant proteins. But because of its characteristics, it can be used in an altered form to facilitate production of recombinant proteins which are toxic for the host cells. This is possible because the weakness of the system and therefore low protein expression rates reduce the cellular stress levels compared to stronger promotor systems. (Quick and Wright, 2002, Terpe, 2006)

The *lac* promotor system characteristics can be changed by the number of lac operators responsible for regulation, their symmetry and the distance between the operators, as these characteristics strongly influence *lacI* binding. One proven strategy is to only apply one *lacO* between the -35 bp region and the -10 bp region of the promotor (Schuller et al., 2020).

Figure 1: Schematic representation of a 1lacO promotor system

To obtain a tightly repressed and controllable system utilizing the *lac* promotor in combination with only one lac operator, one must tune the level of LacI molecules in the cell to a level which facilitates effective repression (Schuller et al., 2020). For this purpose, the *lacI*^{Q} promotor can be utilized. It carries a single base change in the -35 promotor region, which leads to 10 times higher LacI expression rates (Glascock and Weickert, 1998). These heightened LacI levels used in combination with one *lacO* yield a *lac* promotor system, which exhibits reasonably high expression rates, shows low basal expression and is tightly controllable (Schuller et al., 2020).

2.1.3 The role of transcription factor sigma 70

Initiation of transcription at the *lac* promotor occurs, when its unrepressed form is recognized by the sigma 70 transcription factor, bound to the bacterial RNAP (Lodish et al., 2000). This transcription factor is especially essential for the transcription of genes during the exponential growth phase of *E.* coli (Jishage and Ishihama, 1995).

While this is true for all *E. coli* strains, the sequence of the sigma 70 transcription factor can show slight changes in its active centre for different strains. For instance, there are numerous naturally occurring differences in amino acid sequence on position 571 in the sigma 70 factor. Most *E. coli* strains (including BL21) carry a histidine at this position, while alternatively BW25113 incorporates glutamic acid. (Tomatis et al., 2019)

Another possible amino acid at this position is tyrosine, which is for example naturally occurring in the *E. coli* K12 strain (UniProtKB, 1986). It has been reported, that changes on this position can have substantial effects on the binding affinity of the sigma 70 factor to the RNAP core (Jishage et al., 2002). For instance, *rpoD*-Y571H exhibits reduced binding affinity, which impairs its ability to compete with other sigma factors for the RNAP core and therefore reduces transcription initiation rates (Jishage et al., 2002).

Another position in the sigma 70 subunit, which was found to have a substantial impact on RNAP core binding, is 575. It was discovered by directed evolution experiments, that a *rpoD*-E575V mutation lowers the transcription rate of recombinant proteins, but ultimately quadrupled the production of a human membrane protein, which is toxic for *E. coli*. This was mediated by the lowered transcription rates experienced, as well as a changed expression rate of numerous other host genes, which is hypothesized to bring the cellular stress for the host organism to an acceptable level. (Tomatis et al., 2019)

This emphasizes, that even minor changes introduced to a single cellular protein, as well as small differences in the inherent nature of various *E. coli* strains, can have dramatic effects on recombinant protein production.

2.2 *E. coli* K- and B-strains – a short comparison from an industrial point of view

Two of the most commonly used *E. coli* production strains for the expression of recombinant proteins on an industrial scale are HMS174 (K-strain) and BL21 (B-strain) (Marisch et al., 2013b).

On a genetic level, the basic genomes for K- and B-strains show typically an alignment of around 99% bp identity, suggesting only minor variations in the form of small deletions and poorly matched regions. Predicted from these results, it is estimated, that more than half of the 3793 proteins responsible for the basic functions of the cells are identical. The other half of the proteins carry mutations, but it is estimated that only about 310 proteins are exclusively functional in either B- or K-strain but not in both. (Studier et al., 2009)

The genetic differences of these strains manifest themselves in profoundly different behaviour regarding metabolic physiology and gene expression, which has implications on product yield, titre and productivity (Monk et al., 2016). Especially because interferences in cellular processes caused by the recombinantly produced protein are little understood, comprehension of the applied host system is of utmost importance in developing a new industrial production process for a recombinant protein. When HMS174 and BL21 are cultivated in large scale fermentation conditions, distinct differences emerge. For one, when cultivated in batch fermentations, the B strain does not only achieve higher growth rates and higher biomass yields, but also shows lower acetate accumulation in the cultivation media. It has been reported, that up to 16% of the available carbon is converted into acetate in HMS174. This leads to the effect of this strain being less efficient regarding glucose utilisation for biomass build-up. Also, on transcriptome level, HMS174 and BL21 show distinct differences. It was found that 155 genes were expressed differentially in BL21, 50 in HMS174 and only one was commonly differentially expressed during a standard batch cultivation. One of the most striking differences found is the upregulation of genes involved in motility in HMS174 in contrast to the BL21 strain, that carries a deletion of the *fli*-operon, making it unable to produce flagella. (Studier et al., 2009, Marisch et al., 2013b)

Analysis of protein expression data in combination with transcriptomics can yield valuable information on the comparison of the two strains, as well as help identify possible targets for further strain modification and optimisation. Marisch et al. found among others that in BL21 the carbon transport activity was higher compared to HMS174, due to a higher expression of MalE and GatD, two proteins important in maltose transport and galactose metabolism. This is in harmony with the higher observed growth rates in BL21 and makes it a possible target for strain modification in HMS174. (Marisch et al., 2013b)

In an industrial production process, system stability and recombinant protein production capabilities are of utmost importance. It has been shown that stability of the systems and production can greatly differ in different strains. Marisch et al. used an T7-based, plasmid production system to produce the model protein human superoxide dismutase (SOD) in fed-batch fermentations to compare three different production strains: two K-strains (HMS174 and RV308) and the B-strain BL21. They found, that while BL21 produced the highest specific amount of recombinant protein, HMS174 showed the best stability and retained the plasmid-based production system over the whole process time. This stands in contrast to BL21 and

RV308 exhibiting plasmid loss, while maintaining cell growth after induction, which was not observed for HMS174. While RV308 produced the least amount of recombinant protein, it showed the lowest levels of cellular stress as well as the lowest rate of degradation of soluble protein. Based on the characteristics of the different strains, a targeted strain selection is highly important when choosing a suitable production system for successful recombinant protein production. (Marisch et al., 2013a)

2.3 Model protein green fluorescence protein (GFP)

One commonly used, easy-to-produce recombinant protein for the evaluation of novel production systems is GFP which was originally discovered by Shimomura in 1962 in *Aequorea victoria* (jellyfish). It exhibits a strong green fluorescence under UV light, with an emission spectrum peaking at 508nm. Especially due to its characteristics as efficiently emitting internal fluorophore, GFP has become a widely used marker protein of gene expression. (Tsien, 1998, Shimomura, 2005)

The wild type protein has a length of 238 amino acids and a molecular mass of 26886 Da (UniProtKB, 1995). It forms an ap-hydoxybenzylideneimidazolinone chromophore constituted of the residues 65-67, which are serine, tyrosine and glycine respectively (Tsien, 1998). The chromophore develops by autocatalytic backbone condensation between Ser-65 and Gly-67, which is followed by oxidation of Try-66 to didehydrotyrosine (Cody et al., 1993). For this process molecular oxygen is needed (Cody et al., 1993).

One especially interesting variant of GFP, when *E. coli* is used as production organism, is GFPmut3.1. Its region around the chromophore has been subjected to oligo-directed mutagenesis, leading to two single amino acid substitutions, namely Glycine 65 and Alanine 72, whereas in the wild type protein Serine is incorporated at both positions. This results in a red-shifted excitation maximum of the mutant to 488 nm, allowing for an approximately 21-fold increase in fluorescence intensity when compared to the original protein. An additional benefit of using the GFP mutant for production in *E. coli* is, that it shows enhanced folding efficiency compared to the wild type protein. (Cormack et al., 1996)

2.4 Fab – Antibody fragments

Since the rise of full-length monoclonal antibodies (mAB) as effective therapeutic agents with 22 approved mAB for therapeutic use in the United States of America and over 200 mAB candidates in the clinical pipeline, research has been focussing on diversification of molecular forms of antibodies to study (Nelson and Reichert, 2009).

One of the possibilities to reduce the size of the mAB but still retain its antigen binding capabilities is the use of fragments, like the antigen-binding fragment (Fab). The Fab consists of the light chain (LC) and a fragment of the heavy chain (HC) linked via a disulfate bond. The LC is constituted of the full variable and constant region and the HC consists of the variable and first constant region of the full-length HC. As of 2010, there were 3 Fab therapeutics approved in the United States of America. (Nelson, 2010)

Fabs offer some distinct advantages in manufacturing and as therapeutics compared to conventional mABs, which makes them especially interesting alternatives. One key advantage is their relatively small size, which enables them to permeate tissues and tumours more rapidly and thoroughly than mABs. Furthermore, it has been hypothesized, that Fabs might enable binding to cryptic epitopes, which is not possible for full length antibodies. In addition to the positive therapeutic properties of Fabs, it has been proven, that they can be efficiently produced in microorganisms such as *E. coli,* because of their non-glycosylated nature. This makes production much easier and reduces costs significantly compared to mAB, which have to be produced in mammalian cell culture. (Nelson and Reichert, 2009)

Unfortunately, there are also several disadvantages of Fabs. For one, Fabs lack the Fc site, which facilitates antibody-dependent cell-mediated cytotoxicity and complement-dependent cytotoxicity, making binding to the epitope the only possible mode of action. The absence of the Fc domain also leads to lower stability of the Fab and rapid degradation in humans. To increase the half-life of the molecules, special modifications were applied such as conjugation to other proteins like albumin. PEGylation is another option to increase stability and has been applied in case of the anti-TNFα Fab, certolizumab pegol, an FDA approved therapeutic Fab. Another disadvantage arising due to the lack of the Fc site is the increased risk of aggregation. This might occur during production and purification and possibly raises the risk of immunogenicity in patients. (Nelson, 2010)

To obtain a functional Fab, the formation of a disulphide bond between the LC and HC is necessary. Therefore, the immature proteins have to be transported into an oxidising environment, where the formation of the disulphide bond is favoured. In bacterial expression systems such as *E. coli*, this is facilitated by the export of the LC and the HC into the oxidising environment of the periplasm. (Kasli et al., 2019)

In the periplasm also Dsb enzymes that catalyse formation, correction and maintenance of disulphide bonds facilitate correct Fab folding (Ito and Inaba, 2008). One example of the Dsb enzymes is DsbA, which plays the prime role in introducing disulphide bonds into general exported proteins (Ito and Inaba, 2008). Additionally, the periplasm offers further benefits for recombinantly produced proteins. For one, it contains only around 4 – 8% of the total proteins present in *E. coli* and has a reduced protease concentration compared to the cytosol, therefore significantly lowering the risk of proteolytic degradation (Kasli et al., 2019). Another advantage would simply be the spatial separation from the cytoplasm, which reduces downstream purification demands, if cell disruption methods are applied, which selectively disrupt the outer cell membrane (Kasli et al., 2019). However, this might not always be possible, as other studies have found that supposed selective extraction of proteins from the periplasm is rather unselective and extraction efficiency is highly dependent on the size of the protein as well as its concentration (Schimek et al., 2020).

2.5 Translocation in *E. coli*

Transport of proteins into the periplasm in *E. coli* is mainly mediated by two different pathways, namely the secretory (Sec) and the Tat pathway. While proteins in the Tat pathway are fully folded upon translocation, proteins in the Sec pathway translocate unfolded and are

subsequently processed in the cytoplasm (Santini et al., 2001). The main pathway used in *E. coli* is the Sec pathway, responsible for the export of 96% of the exportome (Tsirigotaki et al., 2017). Essential components of the bacterial Sec pathway are the SecYEG translocon, as well as the ATPase motor protein SecA, which plays a role in recognition of the proteins to be exported and mediates the ATP dependent posttranslational export into the periplasm (Tsirigotaki et al., 2017). Proteins that are determined for export through the Sec pathway carry an N- terminally linked, highly hydrophobic signal peptide (Kasli et al., 2019).

Export via the Sec pathway can be mediated in two different ways: either post-translationally or co-translationally. In the co-translational way, the signal peptide is recognized by the signal recognition particle (SRP) and then targeted to its membrane receptor FtsY. There, transport through the SecYEG translocon occurs simultaneously to translation. Contrary to this, in the post-transcriptional pathway the pre-mature protein's signal peptide is recognized by export-factors such as SecA or the chaperones SecB and trigger factor. Subsequently, the protein is then targeted to the SecYEG translocon bound to SecA and export of the protein commences. In both pathways, the signal peptide (commonly PelB or DsbA) is cleaved off after translocation, leaving the protein successfully exported into the periplasm and ready for disulphide bond formation and folding. (Tsirigotaki et al., 2017)

2.6 Continuous bioprocessing for production of biopharmaceuticals

The most common cultivation method in industrial microbial production of biopharmaceuticals is the fed-batch fermentation method, whereby recombinant proteins can be produced in high cell density cultures. However, in order to meet the constantly rising demand for recombinant proteins, more efficient and innovative strategies for the production of large amounts of high-quality proteins are needed. (Graumann and Premstaller, 2006)

Already observed trends in the natural evolution of other industry-sectors show a transition from different forms of batch processes to continuous production, which would also show significant benefits in the biotechnological industry. One of the most important advantages from an economical point of view is certainly cost reduction. It has been shown that for an integrated continuously run biotechnological manufacturing process, average cost can be reduced by approximately 55% compared to conventional processing on a ten-year portfolio. According to the same economic model, an additional savings potential of up to 25% under elevated product demand situations can be predicted. (Walther et al., 2015)

The comparison of fed-batch and continuous fermentation systems reveals that essential process parameters such as temperature, pH, dissolved oxygen concentration and nutrient feed rate can be precisely controlled in both modes of action. This results in both cases in tight control on growth rate and biomass concentration. However, the main difference between fed-batch fermentation and continuous fermentation is the constant supply and removal of medium and fermentation broth in continuous fermentation, resulting in a constant dilution of the culture. Because of this dilution, growth rate, pH, biomass concentration, dissolved oxygen concentration, as well as cellular proteins and metabolites can reach a dynamic equilibrium over extended periods of time. (Matteau et al., 2015)

This equilibrium can be exploited for the efficient production of recombinant proteins, as the possibility to elongate the production phase has the potential to significantly increase process productivities while simultaneously lowering the cost (Peebo and Neubauer, 2018). Nevertheless, continuous cultures also have some disadvantages over the fed-batch cultivation approach. For instance, genetic stability of the production strain becomes a major concern over longer production periods, especially when cells are exposed to high metabolic stress levels (Vethanayagam and Flower, 2005). This phenomenon has been observed while working with systems based on the common T7 phage expression scheme, where reduced amounts of active T7 RNAP were produced due to mutations in its chromosomally encoded gene (Vethanayagam and Flower, 2005).

The vast majority of recombinant proteins nowadays is produced using plasmid-based expression systems, which offer the advantage of high expression rates, due to higher copy numbers of the gene of interest compared to chromosomally integrated systems. This and especially the ease as well as relative speed of genetic manipulations on plasmid-based systems makes them the construct of choice for many production processes. (Waegeman and Soetaert, 2011)

However, in case of long-term continuous fermentation processes, complications such as plasmid loss and consequently the outgrowth of plasmid carrying cells can be an unpleasant scenario. It has been proposed that issues arising especially during continuous fermentation, such as high gene doses, plasmid loss and elevated metabolic load can be evaded when the target gene is chromosomally integrated into the production strains. It has been reported, that the *E. coli* K-strain HMS174(DE3) was successfully modified to contain the gene encoding GFP under control of a T7 promotor integrated into the host cell genome at the TN7 site. This newly created strain was then cultured in a fed-batch fermentation in parallel to a production system carrying the analogous construct on a plasmid. Upon induction it has been observed that the plasmid-based system reached a very high product formation rate, however the plasmid copy number simultaneously more than tripled. This trapped the cells in a vicious cycle of amplification of the gene dose and metabolic stress. Due to this cycle, cell metabolism was irreversibly damaged shortly after induction and cell growth could not be maintained at the predefined rate. To the contrary, in the genome integrated system cell growth did not come to a halt and comparable product formation rates and a high product end titre were reached. The same system was then applied to further continuous fermentation experiments carried out in chemostat culture. This experiment showed high system stability over more than 40 doublings in a non-induced state and more than 10 doublings in induced culture, resulting in high product concentrations. Unfortunately, the system turned instable 70 h after induction. This is likely due to cellular responses mediated by high stress levels over extended periods of time. (Striedner et al., 2010)

A more specific explanation for this phenomenon could be the production of non-functional T7 RNAP as a result of chromosomal mutations in the gene encoding the T7 RNAP. It has been shown that protein production systems based on T7 RNAP can show a decrease in productivity due to these mutations (Vethanayagam and Flower, 2005). Especially the continuous production of recombinant intracellular proteins is challenging because the retention times of host cell and product itself are identical and can therefore not be separated, which leads to a

constant removal of production cells (Schmideder and Weuster-botz, 2017). Additionally, the cells are exposed to a considerable metabolic burden, when strong promotors like in the inducible T7 RNAP system are used (Striedner et al., 2003).

This may lead to genetic instability of the production organism and eventually to outgrowth of the producing population by emerging non-producing mutant populations. Mutation hotspots may include the gene for the T7 RNAP since it is not essential for the survival of the host itself and leads to high metabolic stress due to its immense activity. Furthermore, it is only used for the transcription of the recombinant protein sequence and therefore expendable for all regular cell functions. (Vethanayagam and Flower, 2005)

For plasmid-based expression systems, a common way to shut off recombinant protein production by the host cells is the loss of the plasmid. This scenario is highly unpleasant, since non-plasmid carrying cells can grow approximately 30% faster than the original population producing the target protein. Outgrowth of the original production host can then commence until eventually no producing cells are left. (Sieben et al., 2016)

One way to prevent these problems is to spatially separate biomass production and the recombinant protein formation phase. Schmideder et al. (2017), proposed the use of a two-vessel cascade system of continuously stirred tank reactors (CSTR) to achieve this goal. They used the recombinant photoactivatable protein mCherry (PAmCherry) expressed by *E. coli* BL21(DE3) under the control of the T7 RNAP system. The first vessel used in this approach was implemented solely for the production of biomass, while the second CSTR was used for the formation of recombinant protein. This allowed for a constant feed of fresh cells into the second protein production reactor. Using this setup, they were able to outperform a previously described optimized fed-batch process for the production of mCherry regarding space time yield. However, product concentration, product selectivity and product yield coefficient of the fed-batch process could not be reached, but it is suggested, that by increasing the biomass concentration in the cascade system, further improvements to this system can be achieved. (Schmideder and Weuster-botz, 2017)

Another approach to guarantee the stability of a plasmid-based system regarding plasmid loss is the application of antibiotics to the culture medium, when selection based on antibiotic resistance is used. Unfortunately, antibiotic levels need to be kept at a constant level over the whole production phase to efficiently prevent plasmid loss. This poses a problem since antibiotics are generally expensive compounds and in some cases as for example the commonly used antibiotic ampicillin, leakage of antibiotic degrading enzymes out of producing plasmid bearing cells can protect plasmid free cells to a certain extent. (Sieben et al., 2016)

An additional problem that goes hand in hand with the use of antibiotic resistance genes is the rise of antibiotic resistant pathogenic bacteria that can emerge by horizontal gene transfer of resistances (Walsh, 2000). Although, this is a natural phenomenon, it has been undoubtedly accelerated by the use, misuse and abuse of antibiotics (Walsh, 2000). Additionally, to all these disadvantages the use of antibiotics in fermentation processes pose, the application of antibiotic compounds in pharmaceutical industry is often prohibited by health authorities. While the use of Kanamycin and Tetracycline may be allowed under some circumstances, the

use of β-lactams, compounds which are often affiliated with allergic reactions, is forbidden (Peubez et al., 2010).

Alternatively, to antibiotic based selection of plasmid bearing cells, Selvamani *et al.* 2013, proposed a different mechanism applicable to long term continuous fermentation. For this purpose, the *leuB* gene, a crucial component of the leucine biosynthetic pathway, was introduced into the same plasmid, harbouring the gene of interest and transferred into an *E. coli leuB* knockout strain. This strain was grown in a chemostat cultivation for more than 200 h and showed a stable maintenance of the plasmid in the absence of antibiotics. Through comparison with a control strain it could be shown that the plasmid stability was close to 100% over the entire period of 200 h, which is similar to experiments performed with antibiotic selection pressure. Under the absence of selection pressure in the same control strain, the population harbouring the plasmid was reduced to 79% after 190 h. (Velur Selvamani et al., 2014)

This emphasises, that continuous culture is a viable approach for the production of recombinant proteins and can be performed with plasmid-based, as well as chromosomally integrated systems.

2.7 Stress induced mutations in *E. coli*

Every biological population naturally undergoes constant selection pressure, which is made evident by the formation of spontaneous mutations eventually resulting in evolution. While the mutation rates can vary widely from 10^{-2} to 10^{-10} per base pair per generation, it has been determined that wild type *E. coli* show a spontaneous mutation rate of about 10^{-3} per genome per generation. (Rugbjerg and Sommer, 2019, Lee et al., 2012)

However, in the case of recombinant protein production, certain genomic regions triggering cellular stress may develop to mutation hotspots. For instance, base substitutions in the sequence of the T7 RNAP or the lac repressor can lead to lower recombinant protein production, resulting in relieve of cellular stress. (Rugbjerg and Sommer, 2019, Vethanayagam and Flower, 2005)

To study the underlying general mechanisms of stress related mutations, Maharjan et. al. (2017), conducted mutagenesis studies using chemostat cultures under different nutritional limitations. The emerging mutation patterns under single nutrient limitation (P, O, C, N and Fe) and a constant growth rate of 0.1 h^{-1} leading to high cellular stress were studied. Surprisingly, mutation rates were not always elevated, but the change of the mutational pattern was significant. Fe-, O- and N- limitations did not increase the overall mutation rate, while P- and C- limitations did. Analysis of the mutational pattern showed that mutations requiring transposons were highest under Fe- and O- limitation, while base-pair substitutions and indels were highest under P limitation. This suggests that stressful conditions in nature strongly influence the pattern of genetic variation and that the evolvability of specific traits is dependent on particular mutations. However, since a clear link between increased rates of individual mutation types and their benefit in those specific environments cannot be seen, it is suggested, that environmental conditions do not provide directed mutations. Nonetheless,

a connection between nutrient limitation and selectivity of DNA repair mechanisms was revealed. DNA repair system expression patterns changed in each nutritional state, with the most mutagenic limitation exhibiting the lowest rate of MutS and MutY DNA repair levels. (Maharjan and Ferenci, 2017)

When a similar study was performed with double nutrient limitations, it could be observed, that the rate of mutations is not cumulative in double limitations but falls in-between the value of the two single mutations. Double limitations are therefore not synergistic or additive but behave as if one of the two limitations dominate over the other. This suggests interactions at the regulatory level, possibly of DNA repair systems affecting mutation rates. These double limitation findings are not entirely consistent with Liebig's law of the minimum, which suggests, that a single dominant limitation determines biological behaviour. Nonetheless, studies with additional different double limitations suggest that combined limitations can differ from individual stresses in physiological and evolutionary outcomes, hinting that environments not only provide selection conditions but also particular patterns of genetic variation. (Maharjan and Ferenci, 2018a, Warsi and Dykhuizen, 2017)

However, nutrient based limitations are not an issue in industrial fermentations due to the possibility to optimize fermentation media. Other influencing factors are for instance the growth rate, which is relatively easy to regulate and can have a significant positive impact on recombinant protein production when it is reduced during expression (Lim and Jung, 1998). But this could pose a problem especially in continuous cultivation, as it has been reported that mutation rates can be raised significantly with decreasing growth rates. Already a small shift from 0.7 h^{-1} to 0.6 h^{-1} in aerobic culture was found to double the mutation rate per generation. At a growth rate of 0.1 h^{-1} the mutation rates per generation for aerobic cultures further increased by a factor of approximately 4 compared to nutrient rich conditions, suggesting that growth rate significantly influences mutation rates. (Maharjan and Ferenci, 2018b)

To overcome the issues of metabolic load and the thereby related genetic escape, the development of stable, high producing and reliable cell lines is the key to success. By means of deep DNA sequencing, the analysis of evolving possible failure modes can be analysed. With this technique, genes linked to metabolic load and high probabilities of a mutation can be identified and thus enable targeted genetic edits to yield stable production strains for long term commercial fermentations. (Rugbjerg and Sommer, 2019)

3 Aim of thesis

The main goals of this master thesis are to investigate the genomic stability of genome-integrated *E. coli* expression systems in long-term production conditions and to investigate the resulting chromosomal response of the host organism due to metabolic load triggered by recombinant gene expression. In the course of the thesis, two different promotor systems will be compared, using the *E. coli* strain BL21. On the one hand the commonly used T7 expression system, utilizing a highly active phage derived RNAP will be applied and on the other hand a much weaker novel promotor system exploiting host native RNAP. This weaker system consists of the A1 promotor, one of the main early stage promotors of the bacteriophage T7, in combination with a single *lacO* which has been inserted between the -35 and the -10 region of the promotor. Through the use of a host native RNAP in this system, the risk of mutations within the RNAP itself can be circumvented, since the functioning host RNAP is needed for all cellular proteins alike.

These two systems will be analysed regarding their stability and productivity while expressing an easy-to-produce model protein, GFPmut3.1, and a challenging protein, Fab dFTN2. The high cellular stress caused by recombinant protein production will be utilized as a tool to mutate the host's genome during several serial cultivations in a high throughput fermentation environment. Potential mutants, which exhibit good growth characteristics and reasonable product titres, will be isolated. These cell lines are hypothesised to be mutated in order to circumvent or reduce cellular stress upon recombinant protein production, which makes them ideal candidates for long term continuous production processes.

Another goal of this thesis is to incorporate the 1lacO-A1 promotor system into HMS174^{Q} with GFPmut3.1 and dFTN2 as target recombinant proteins respectively and to evaluate the performance of this expression system in a K12 strain compared to the B strain BL21^{Q}.

4 Materials and methods

All methods were processed according to the manufacturer's protocols or standard operating instructions available in the laboratory, if not stated otherwise.

4.1 Strain generation

4.1.1 Strains used

For genome integration of the expression cassettes housing the target genes, *E. coli* K-12 strain HMS174 (genotype: F^- *recA1 hsdR*(r_{K12}^- m_{K12}^+) (Rif ^{R})) was used. Prior to genome integration, the strain was modified to carry the $lacI^Q$ promotor. For testing and comparison, the K-12 strain *E. coli* RV308 (genotype: *lacI* q-, su-, Δ*lacX74*, *gal*, IS II::OP308, *strA*), as well as the B strain BL21 (genotype: *fhuA2 [lon] ompT gal [dcm] ΔhsdS*) (New England BioLabs (NEB), C2530H), modified to carry the $lacI^Q$ promotor, was used. All these cultures were sourced from in-house cell banks stored at -80 °C. Transformation of non-methylated newly constructed plasmids, as well as their storage was facilitated using the *E. coli* K12 strain NEB® 5-alpha (genotype: *fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80 Δ(lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17*) (NEB, C2987H).

4.1.2 Genome integration and selection of positive recombinants

Subsequently to PCR and purification, the genome integration cassettes were integrated into *E. coli* $HMS174^Q$ using an altered version of the protocol provided by Sharan *et al.* 2009. For this purpose, an overnight culture of the bacterial strain, carrying a pSIM plasmid was prepared in M9ZB medium supplemented with the antibiotic used for plasmid selection (chloramphenicol 30 µg/mL, ampicillin 30 µg/mL). The pSIM plasmid carries an antibiotic resistance gene as well as three genes of the λred phage essential for genome integration: Exo, Beta and Gam. Exo has a 5' to 3' dsDNA exonuclease activity, which generates 3' overhangs on linear DNA. Beta binds the single stranded DNA (3' overhangs), promotes ss-annealing and generates recombinant DNA. Gam prevents RecBCD nuclease from degrading double-strand linear DNA fragments and therefor plays an important role in dsDNA recombineering. Due to the temperature sensitive origin of replication of the pSIM plasmid, the above-mentioned overnight culture was grown at 30 °C and 180 rpm on an orbital shaker. On the next day, 0.5 mL of the culture were transferred into 35 mL fresh M9ZB medium supplemented with the adequate antibiotic and left to grow at 30 °C and 180 rpm until an OD_{600} of one was reached. Then the cells were placed in a 42 °C water bath and shaken at 200 rpm for 15 min. Since the lambda red genes have temperature sensitive promotors, Exo, Geta and Gam are produced in this phase. After the induction phase, the cells were chilled on ice for 5 min and 17.5 mL of the culture was centrifuged at 4600 g for 7 min at 4 °C in a precooled centrifuge. The supernatant was poured off and the cells were resuspended in 30 mL of ice-cold dH_2O. Then, centrifugation as described above was carried out. Again, the supernatant was poured off, the cells were resuspended in 1 mL of ice-cold dH_2O and transferred to a pre-

chilled microcentrifuge tube. In this precooled microcentrifuge, the cells were spun at 10000 g and 4 °C for 30 s. The procedure of washing the cells in 1 mL of RO water was repeated once more and finally the cell pellet was resuspended in 200 µL of ice-cold dH_2O. The electroporator was turned on and set to 1.80 kV, 25 µF and 200 Ω. For introduction of the integration cassette into the cells, 45 µL of cell suspension and 5 µL of solution harbouring the linear DNA fragment were transferred into a 0.1 mm precooled electroporation cuvette and mixed. The cell-DNA suspension was then immediately pulsed and resuspended in 950 µL SOC outgrowth medium (New England BioLabs, B9020S). The cells were then grown overnight at 30 °C and 750 rpm in a Thermomixer and the whole suspension was afterwards plated on Vegitone agar plates supplemented with the antibiotic suited for the selection of positive recombinants and incubated at 30 °C overnight. Adapted from (Sharan et al., 2009)

The grown colonies were then picked with a sterile pipette tip and transferred to a fresh Vegitone agar plate supplemented with the same antibiotics. After incubation overnight at 37 °C, the grown clones were screened via colony PCR. Then, the resulting DNA fragments from positive recombinants were sequenced (outsourced to Microsynth AG) consequently to gel extraction and purification.

4.1.3 Electroporation of plasmid DNA

For the introduction of plasmid DNA into *E. coli* strains, a slightly altered version of the electroporation protocol for genomic integration provided by Sharan *et al.* 2009 was applied to produce electrocompetent cultures. Cells were grown overnight in 10 mL M9ZB medium optionally containing the desired antibiotic for selection at 37 °C and 180 rpm on an orbital shaker. On the next day 0.5 mL of the overnight culture were transferred into 35 mL of fresh M9ZB medium and grown at 37 °C and 180 rpm until an OD_{600} of around 0.6 was reached. Then, 17.5 mL of the culture were centrifuged in a precooled centrifuge at 4 °C with 4600 g for 7 min. The supernatant was disposed of, then the pellet was resuspended in 30 mL of ice-cold dH_2O and centrifuged again. After disposal of the supernatant, the cells were resuspended in 1 mL of ice-cold dH_2O, transferred into a precooled microcentrifuge tube and centrifuged in a precooled microcentrifuge at 4 °C and 10000 g for 30 s. This wash step was repeated once more and then the cells were resuspended in 200 µL of ice-cold dH_2O and kept on ice until used. For electroporation, the electroporator (BTX Harvard Apparatus: ECM630 Electro Cell Manipulator) was set to 1.80 kV, 25 µF and 200 Ω. Then, 49 µL of cell suspension and 1 µL of plasmid DNA solution were mixed in a 0.1 mm precooled electroporation cuvette and pulsed immediately. The cells were then mixed with 950 µL SOC outgrowth medium (New England BioLabs, B9020S) and left to recover for 3 h at 37 °C and 750 rpm on a Thermomixer. Then 100 µL of the suspension were plated on Vegitone agar containing the desired antibiotic for selection and incubated at 37 °C.

4.1.4 Introduction of plasmid DNA into chemically competent *E. coli*

For the purpose of producing a collection of all constructed plasmids, as well as to have them readily available at all times, constructed plasmids were immediately after completion introduced into bacterial hosts for storage. Chemically competent *E. coli* (NEB® 5-alpha,

C2987H) from New England Biolabs (NEB, USA) were used for this purpose. The introduction of plasmid DNA and cultivation of cells was performed according to the manufacturer's instructions.

4.2 Generation of plasmids and integration cassettes

4.2.1 *lacI_opt* plasmid for GFPmut3.1 expression

The plasmid harbouring the *lacI_opt* gene with the W220F mutation was created via PCR, using the already existing pETk1lacOA1tZ.c-GFPmut3.1 plasmid, which carries a copy of the *lacI* gene. For this purpose, two primers have been designed to bind within the *lacI* gene of the plasmid, with one of them carrying the desired mutation. After successful PCR, the product was DpnI (NEB, R0176S) digested to remove the template plasmid and subsequently purified via agarose gel electrophoresis followed by gel extraction and PCR clean-up (NEB, Monarch® Kits). The linear PCR product was then 5' phosphorylated via a T4 Kinase reaction and subsequently ligated with T4 ligase. All enzymes were supplied by NEB and the procedures were performed according to the manufacturer's instructions. Then the plasmid was purified using the Monarch® PCR & DNA Cleanup Kit (NEB) and electroporated into competent *E. coli* HMS174^{Q}.

4.2.2 *rpoD_opt* integration cassette

The genome integration cassette harbouring the alteration leading to the H571Y mutation in the sigma 70 factor (see Appendix: Figure 22), was designed using the software "CLC Genomics Workbench 12, Version 12.0.3" and ordered from Integrated DNA Technologies, Inc. in gBlock® format. It features a 52 bp long homologous overhang on the 5' end of the cassette and a 61 bp long overhang on the 3' end. The changed base, leading to the desired mutation is highlighted in Figure 22 and marked as BL21_sequence. Additionally, the integration cassette contains a chloramphenicol resistance gene for selection of positive recombinants, which is flanked by two FRT sites. Design of the integration cassette was carried out to minimize changes in the host's chromosomal structure. For this purpose, the natural sequence of the host's DNA was analysed using ARNold, an online tool to predict possible Rho-independent transcription termination sequences. These sequences are a major mechanism for transcription termination in bacteria, leading to 3' termination and a variety of additional regulatory attenuation events (Naville et al., 2011). One such possible sequence is proposed to start 44 bp downstream of the *rpoD* gene and has a length of 42 bp. Consequently, this sequence was kept unaltered in order to not interfere in cellular transcription mechanisms. The second rpoD_opt integration cassette was designed analogous and featured an elongated 152 bp long homologous 5' overhang. Construction of this cassette was performed via extension PCR, using a 120 bp long homologous primer.

4.2.3 *rpoD_E575V* integration cassette

For the construction of the *rpoD_E575V* genome integration cassette, a plasmid housing the desired cassette was created first. For this purpose, the last 360 bp of the *rpoD* gene, as well as 97 bp downstream of the gene, were amplified via colony PCR from *E. coli* HMS174. For reliable and correctly oriented ligation with the backbone, the forward primer was designed carrying a SphI cutting site on its 5' end, while the reverse primer carried an EcoRI cutting site on its 3' end. The backbone itself was prepared from an already existing plasmid carrying the desired chloramphenicol resistance gene flanked by two FRT sites. This was done analogous to amplification of the target gene, with the forward primer carrying the EcoRI cutting site on its 5' end and the backward primer being designed with a SphI cutting site on its 3' end. After amplification and isolation of the two fragments, they were digested according to the manufacturer's instructions using the suiting enzymes (NEB, SphI-HF® R3182S and EcoRI-HF® R3101S). Then the backbone was DpnI digested and phosphorylated with Antarctic Phosphatase (NEB, M0289S). Subsequently, the two fragments were ligated according to the supplied instructions using T4 DNA ligase NEB, M0202S). This plasmid was then introduced into chemically competent *E. coli* NEB® 5-alpha (NEB, C2987H) and after the successful growth of plasmid carrying cells, the plasmid was extracted from an overnight culture using the Monarch® Plasmid Miniprep Kit (NEB). For the purpose of inserting the desired mutation, a forward and reverse primer was designed to bind exactly next to each other, with one carrying the desired altered sequence. Then a PCR was performed to introduce the mutation. After subsequent agarose gel electrophoresis and gel extraction, the isolate was DpnI digested to remove all template plasmid and then 5' phosphorylated using T4 Polynucleotide Kinase (NEB, M0201S). Afterwards, ligation with T4 DNA ligase was performed and the plasmid was introduced into chemically competent *E. coli* NEB® 5-alpha. For the construction of the genome integration cassette, the plasmid was isolated, and a PCR performed with a forward primer binding to the 5' end of the insert and a reverse primer carrying a 50 bp homologous overhang, which binds downstream of the second FRT site. After successful PCR, the genome integration cassette (see Figure 23) was purified via agarose gel electrophoresis and gel extraction (Monarch® DNA Gel Extraction Kit, NEB).

4.3 General molecular biology procedures

4.3.1 Polymerase chain reaction (PCR)

All PCRs have been performed using Q5® High – Fidelity DNA polymerase and PCR reagents and buffers supplied by NEB. PCRs have been performed according to the manufacturer's instructions and recommendations. Reagent mixes for PCRs contained 1x diluted Q5 reaction buffer, 0.2 mM of each nucleoside triphosphate, 0.5 µM of each of the primers, varying concentrations of template DNA (according to New England BioLab's recommendations), 1x diluted Q5® High – Fidelity DNA polymerase, 1x diluted Q5 High GC enhancer and nuclease free water. In case of colony PCRs, instead of template DNA solution, a few cells from a single colony were picked with a sterile pipette tip and suspended into the reaction mix. Settings for the thermocycler were as follows: initial denaturation at 98 °C for 30 s (2 min in case of colony PCR), followed by 30 cycles of each denaturation at 98 °C for 5 s, annealing at the calculated

annealing temperature dependent on the primers used for 20s, as well as elongation of the product DNA at 72 °C for 30 s per 1 kilobase pair length of the product. Subsequent to the 30 cycles, a final elongation of 2 min at 72 °C was carried out. After completion of the PCR, the resulting DNA products were analysed via agarose gel electrophoresis.

4.3.2 Agarose gel electrophoresis

PCR products were stained using the Gel Loading Dye (6x) from NEB (B7024S). 7 µL of Quick-Load Purple 1 kb Plus DNA ladder (N3200S) from NEB was used in all cases as standard. Prior to electrophoresis, the sample chamber was filled up to the mark with electrophoresis buffer and a fresh gel (stored at 4 °C in electrophoresis buffer) was carefully laid on the bottom of the chamber. Then the standard and all the samples were loaded into the gel and the chamber was closed and connected to the voltage source. Analytical gels were run for 20min at a constant voltage of 120 V, while preparative gel electrophoresis was carried out at 90V for 80 min. Afterwards, the gels were imaged using the Bio-Rad ChemiDoc™ XRS+ System. In case of preparative agarose electrophoresis, the relevant bands were excised using a clean scalpel for further processing.

4.3.3 Agarose gel electrophoresis running-buffer

Table 1: Contents of Agarose Gel Electrophoresis Buffer

RO-H_2O	9.8 L
TAE buffer (50x)	200 mL
Ethidium bromide [10 mg/mL]	300 µL

All contents listed in Table 1 were measured out and mixed in the prepared 10 L dispensing unit.

4.3.4 Preparation of agarose gels (1% agarose content)

Table 2: Contents of 1% agarose gels

dH_2O	490 mL
50x TAE buffer	10 mL
Ethidium Bromide [10mg/mL]	20 µL
Agarose	5 g

All substances according to Table 2 were mixed in a clean 500 mL flask, heated until dissolved and poured into suited moulds. After solidification, they were stored in agarose gel electrophoresis buffer.

4.3.5 Preparation of plasmid DNA

Extractions of plasmid DNA from overnight cultures of *E. coli* were performed using the Monarch® Plasmid Miniprep Kit from NEB. All procedures were carried out according to the manufacturer's instructions.

4.3.6 Extraction of purified DNA from agarose gels

Extraction of purified DNA after agarose gel electrophoresis was performed with the Monarch® DNA Gel Extraction Kit from NEB according to the manufacturer's instructions.

4.3.7 Purification of DNA

Purification of PCR products was carried out using the Monarch® PCR & DNA Cleanup Kit from NEB according to the manufacturer's instructions.

4.4 General cultivation media

4.4.1 Agar plates

All agar plates have been prepared from Vegitone Agar (Nutrient Agar No 2, Vegitone - 04163, Merck), according to the manufacturer's instructions. For this, 11.5 g Vegitone Agar were dissolved in 500 mL of dH_2O and autoclaved for 20 min at 121 °C. In case of media with 2% glucose as supplement, the 11.5g of Vegitone Agar were dissolved in 480 mL of RO water and the needed amount of glucose sterilized separately in 20 mL of RO water. After sterilisation, the agar was cooled in a water-bath to 50 °C, combined with the glucose if necessary and supplemented with antibiotics if needed. Afterwards, the plates were poured in the laminar flow cabinet, left to solidify for 30 min and stored at 4 °C until use.

Table 3: Components of nutrient agar plates

Vegitone Agar	11.5 g
dH_2O	500 mL

Table 4: Components of nutrient agar supplemented with 2% glucose

Vegitone Agar	11.5 g
dH_2O	500 mL
D(+)-Glucose	10 g

Table 5: Volumes of antibiotic stocks added to 500 mL agar if necessary

Ampicillin stock [100 mg/mL]	150 µL (end concentration in agar: 30 µg/mL)
Kanamycin stock [50 mg/mL]	300 µL (end concentration in agar: 30 µg/mL)
Chloramphenicol stock [25 mg/mL]	200 µL (end concentration in agar: 10 µg/mL)

4.4.2 M9ZB medium

Preparation of the medium was performed according to Table 6.

Table 6: Components of M9ZB medium

Sequence	Component	Amount	
1	Tryptone	10.00	g/L
	Yeast extract	5.00	g/L
	NaCl	5.00	g/L
	NH_4Cl	1.00	g/L
	KH_2PO_4	3.00	g/L
	Na_2HPO_4	6.00	g/L
2	Glucose monohydrate	4.00	g/L
3	1 M $MgSO_4$-Solution	1.00	ml/L

For preparation of this medium, components under sequence one and two in Table 6 were dissolved in separate flasks and autoclaved. After sterilisation, both solutions were combined, and the needed amount of $MgSO_4$-Solution was added. For preparation of the 1 M $MgSO_4$-Solution, 246.47g $MgSO_4*7H_2O$ were dissolved in 1 L of dH_2O and sterilized.

4.4.3 LB medium

Preparation of the LB medium was performed according to Table 7.

Table 7: Composition of LB medium

Sequence	Substance	Amount	
1	NaCl	10.00	g/l
2	Tryptone	10.00	g/l
3	Yeast extract	5.00	g/l

For the preparation of this medium all components were dissolved in dH_2O in the sequence depicted and the bottle was subsequently filled to the desired volume with dH_2O. The medium was then autoclaved and stored at 4 °C prior to use.

4.5 BioLector® cultivations

4.5.1 BioLector® - device

The BioLector® is a high throughput micro-bioreactor device, capable of carrying out up to 48 cultivations in parallel. Due to its capability to regulate temperature, humidity and the gas atmosphere during cultivation, the micro-scale fermentations are highly reproducible. Additionally, the device gives insight in real time data on pH, dissolved oxygen concentration (DO), biomass concentration and has the ability to measure fluorescent molecules if needed. (m2p-labs, 2020a)

Cultivations are carried out in 48-well flower plates (m2p-labs), capable of effective mixing via the three mm orbital shaking motion of the BioLector®, high mass transfer of oxygen (>0.11 mol/L/h), a wide cultivation volume range of 800-1900 µL and allow for scalability to production bioreactors. (m2p-labs, 2020b)

4.5.2 Medium

For all BioLector® cultivations except for pre-cultures, an in-house produced medium was used. This synthetic feed in time (FIT) medium, capable of simulating fed-batch like cultivation conditions, utilizes two different carbon sources. For the batch cultivation glucose is used, while the fed batch is simulated by a glucose releasing enzyme mix reacting with dextran. During the time course of the fermentation, the hydrolytic enzyme glucoamylase cleaves off glucose residues from the non-metabolizable polysaccharide, which feed the culture. This process is capable of imitating the lab-scale substrate feed during fed batch cultivations (Toeroek et al., 2015). Stock solutions of the media components were prepared according to Table 16 and mixed to obtain the final medium as is shown in Table 17. The polysaccharide solution for this medium was taken from the m2p-labs Media development kit and not produced in-house. Prior to inoculation the medium was supplemented with sterile glucose stock (100 g/L) to an end concentration of 1 g/L glucose in the medium. To facilitate fed batch like characteristics during cultivation, 0.6% of the glucose releasing m2p-labs Enzyme-Mix 100 U/mL were added. Since all cultivations were carried out in a fully induced environment, IPTG was introduced to the medium to an end concentration of 0.5 mM. For all precultures, M9ZB Medium supplemented with 30 µg/mL kanamycin was used.

4.5.3 Calculation of cell biomass

During BioLector® cultivations, bacterial cell growth was monitored via backward scattered light measurement at 620 nm. The biomass concentration was calculated from light scatter signals with calibration settings obtained by linear regression analysis up to 50 g/L (Kensy et al., 2009, Toeroek et al., 2015). To calibrate the light scatter signal, an overnight culture of the *E. coli* strain BL21(DE3) was prepared in LB medium and grown on an orbital shaker at 37 °C and 180 rpm. The next day, the cells were harvested and centrifuged at 8500 g for 9 min. Afterwards, they were washed with dH_2O and resuspended in 13 mL m2p medium without glucose and enzyme mix. Different dilutions (undiluted, 1:1.5, 1:3, 1:6 and 1:12) were

prepared and measured alongside a blank sample in the BioLector®. The light scatter measurements were performed at 620 nm until constant values were reached. (1400 rpm, 37 °C, 85% humidity and all available gains of 20, 25, 30 and 35) To obtain calibration curves, the CDM of the overnight culture was determined. For this purpose, 2 mL of sample was drawn twice into separate Eppendorf tubes and centrifuged at 10,000 g and 4 °C for 10 minutes. After two washing steps with dH_2O, the pellets were resuspended in dH_2O, poured into separate pre-weighted beakers and placed at 104 °C to dry. After the cell mass was completely dry, the beakers were placed in the autoclave for cooling and were subsequently weighed. After calculation of the CDM, the calibration curves for each separate gain could be determined by linear regression analysis. The results of this analysis including offset and slope of the calibration were then transferred to the BioLector® software BioLection 2 (m2p-labs) for accurate measurements of online biomass concentrations.

4.5.4 Cultivation setup - BioLector®

All cultivations in the BioLector® were performed, using the *E. coli* B strain BL21 (genotype: F^-, *dcm*, *ompT*, *hsdS*(r_B- m_B-), *gal*) with *lacI*^{Q} promotor. The production cassette carrying the gene for both HC and LC of the Fab FTN2 with dsbA leaders, as well as a Kanamycin resistance gene, has been integrated into the host genome at the TN7 site.

The methodology of fully induced serial cultivations to exert high metabolic stress on cells was adapted from Schuller et al. (unpublished data). For all cultivations in the BioLector® fermentation device, 48-well FlowerPlates without optodes (MTP-48-B, m2p-labs) were used. The plates were sealed using sterile gas permeable plate covers with evaporation reduction layer (F-GPR48-10, m2p-labs). For inoculation of the first cultivation plate of the serial passage, 770 µL of combined culture medium was dispensed in each well and inoculated with 30 µL of preculture obtained from an overnight cultivation in M9ZB medium supplemented with 50 µg/mL kanamycin. The following passages were performed similarly, but 790 µL medium and 10 µL inoculum from the previous plate were used. During the course of the cultivation, the biomass concentrations were calculated using the scattered light signal at 620 nm and the previously determined calibration curve. Standard settings for all cultivations were 1400 rpm, 30 °C, 85% humidity and a measurement cycle time of 20 min. The first six passages were performed as described above. For the seventh passage, six selected cultures from the sixth passage were taken and plated on Vegitone agar supplemented with glucose (2%) and kanamycin (50 µg/mL). After colonies grew overnight (37 °C), eight clones from each plate were picked and transferred to a fresh agar plate. This plate was again incubated over night at 37 °C. The following day, a preculture was created from each colony in M9ZB medium supplemented with kanamycin (50 µg/mL) and run in the BioLector®. The seventh passage was then inoculated from these precultures and performed as described above. After the end of the passage, samples calculated to contain 1 mg of CDM were drawn and centrifuged at 10000 g and 4 °C to separate the cell pellet from the supernatant. The supernatant was then drawn off and transferred to a new sample tube, while the tube containing the cell pellet was dried using a paper tissue. All samples were then stored at -20 °C until used for further analysis.

4.6 Continuous cultivations in the DASGIP system

4.6.1 Equipment

The continuous cultivations were carried out in the computer-controlled fermentation system DASGIP® Bioblock from Eppendorf. Four autoclavable bioreactors covering a working volume up to 2 L allow the cultivation of 4 different cultures in parallel in a fully controlled environment. In order to control temperature and stirrer speed, as well as pH and pO_2, two additional control boxes (DASGIP® TC4SC4 and DASGIP® PH4PO4L) were used. Control of air supply to the reactors, as well as composition of the sparge gas, was facilitated by the DASGIP® MX4/4 gas mixing unit from Eppendorf. For media, base and antifoam supply, the Dasbox MP8 pump system was used. To control the fermentation environment, the dissolved oxygen sensor Visi Ferm DO ECS 325HO, as well as the pH sensor EasyFerm Plus PHI K8 325 both supplied by Hamilton were employed. To ensure a sterile environment inside the bioreactors, an inlet filter (Sartopore® Air Midisart® 0.2 µm, Sartorius), as well as an outlet filter (Sartofluor 300, Sterile Capsule, Sartorius) was used on all bioreactors separately.

4.6.2 Media

In the course of the fermentation, three separate media for batch, fed-batch and continuous cultivation were used. The composition of the semisynthetic batch medium (see Table 18) was calculated according to the expected end biomass of either the batch phase (4 g) or the fed-batch process (21 g). The following compounds were calculated referring to the end biomass of the fed-batch: KH_2PO_4, 85% H_3PO_4 (buffer system and P and K source), $C_6H_5Na_3O_7$ * $2H_2O$ and $(NH_4)_2SO_4$. Additionally, the batch medium contained yeast extract, $MgCl_2$ * $6H_2O$, $CaCl_2$ * $2H_2O$, trace element solution (produced in-house, see Table 21) and $C_6H_{12}O_6$ * H_2O. All these components were calculated according to different yield coefficients (e.g.: glucose yield coefficient $Y_{X/S}$ = 0.3 g/g) referring to the end CDM of the batch phase. For preparation of the medium, the required amounts for one bioreactor were calculated and the multiplied by the number of used reactors with an additional safety factor of 0.1. Then, all the components of the medium except of the glucose were weighed in together, dissolved and directly added to the clean fermentation devices. The glucose was then separately prepared for each reactor and connected to a closed inlet with silicone tubing. All the reactors were then autoclaved for 20 min at 121 °C and after cooling, the glucose was complemented into each bioreactor by simply opening the connected inlet port.

Fed-batch media (300 mL per reactor) were prepared similarly, where each of the components was calculated according to the amount of biomass produced in the phase (17 g). The salt solution containing magnesium chloride-hexahydrate, calcium chloride-dihydrate and trace element solution were weighed in and dissolved separately from the glucose. Again, the solutions were autoclaved in separate containers, connected via a clamped off silicone tube and combined after cooling. For detailed composition see Table 19.

The continuous feed medium (see Table 20) was calculated according to the planned constant biomass concentration of 30 g/L after the fed-batch. Additionally, to the already listed compounds above, this medium contained magnesium sulphate-heptahydrate, as well as 20

µmol/gCDM IPTG for the induction of recombinant protein production. This medium was prepared in multiple 20 L batches and sterile filtered prior to use. For filtration of the medium a 3M LifeASSURE™ PDA series disposable capsule filter with a pore size of 0.2 µm was used. After completed sterile filtration into an autoclaved 20 L bottle, the medium was stored at 4 °C until needed.

During fermentation 12.5% (w/w) ammonium hydroxide solution (diluted 1:2 from 25% (w/w) ammonium hydroxide solution, Merck) was used as nitrogen source, as well as to regulate the pH during cultivation.

The trace element solution was prepared in 5 M HCl according to Table 21 and stored at room temperature prior to use.

4.6.3 Fermentation plan and procedure

After sterilisation and complementation of the batch media, the bioreactors were assembled in the DASGIP® Bioblock for a 24h sterility test. To conduct this test, stirrer and gas inlet were switched on, the temperature was set to 37 °C and the pH in all reactors was raised to 7.0. After successfully tested sterility in all bioreactors (observable by constant pO_2 values over 24 h), inoculation was performed. For this, overnight cultures of the desired production strains were prepared in M9ZB medium supplemented with 50 µg/mL kanamycin. The strains used are specified in Table 8.

Table 8: Production strains used for continuous cultivation

Strain	POI	Genotype prior to manipulation	Source	Comments
BL21(DE3)	GFPmut3.1	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS*	NEB, C2527H	T7 system
BL21(DE3)	GFPmut3.1	*fhuA2 [lon] ompT gal (λ DE3) [dcm] ΔhsdS*	NEB, C2527H	T7 system - mutant
BL21^{Q}	GFPmut3.1	*fhuA2 [lon] ompT gal [dcm] ΔhsdS*	NEB, C2530H	1lacO-A1 system

The two BL21(DE3) strains both carry the GFPmut3.1 gene under control of the T7 promotor integrated at the TN7 site into the host's genome. The only difference between these two production strains was, that one of them was a mutant strain, which has been selected from a long-term production experiment, due to its good expression characteristics. The third strain cultivated was *E. coli* BL21 carrying the *lacI*^{Q} promotor additionally to the expression cassette housing the GFPmut3.1 gene under the control of the one *lac* operator – A1 promotor combination, utilizing host RNAP. All bioreactors were inoculated with 25 OD_{600} units of their preculture. The inoculation ports were then connected to the antifoam bottle and all remaining cell suspension was flushed into the reactor.

After inoculation, all control systems were switched on to keep the bioreactors at a constant pH of 7.0, the dissolved oxygen level at 30% and the temperature at 37 °C. After the end of the batch phase, a scale controlled exponential feed with a constant growth rate of 0.1 h^{-1} was started. Aim of the non-induced batch and fed-batch phase was to produce a cell dry mass

concentration of 30 g/L. The cultivations in continuous mode were performed with a constant dilution rate of 0.1 h^{-1} at 30 °C. Recombinant protein production was fully induced with IPTG (20 µmol/g CDM) after switch to continuous cultivation.

In Figure 2, the calculated CDM curve for the first representative 180 h of the fermentation subsequent to the exponential feed start can be seen. After completion of the exponential feed, the cultures were induced with 20 µmol IPTG/g CDM. To keep this induction level at a constant concentration of 20 µmol IPTG/g CDM, the appropriate amounts of IPTG were sterile filtered into the bioreactors as well as into the feed medium for the constant feed. Following the induction, the feed was switched to continuous mode with a dilution rate of 0.1 h^{-1}. For this purpose, all media pumps were set to an hourly flow rate corresponding to a tenth of the volume present inside the bioreactors at the end of the exponential feed (in this case 67.28 mL/h). Simultaneously, the outlet pumps of the bioreactors were activated. These pumps drew cell suspension out of the reactors via a dip tube, which was positioned to just touch the surface of the liquid, therefore keeping the bioreactor volumes constant. The continuous fermentation was continued for a total of 312 h, in which corresponds to 45 generations.

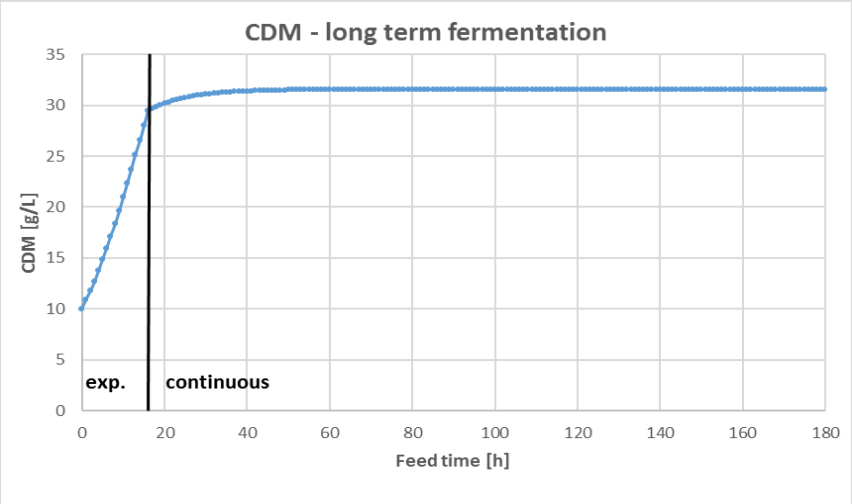

Figure 2: Calculated CDM curve long-term fermentation

4.6.4 Sampling of standard analytics

During continuous cultivation, samples were drawn at 0 h, 5 h, 24 h and following in 24 h intervals over the entire process time. For each bioreactor, 3 mL of cell suspension was sampled, diluted as needed in PBS buffer and its OD_{600} determined (Ultraspec 500 pro, Amersham Biosciences). Following this measurement, an estimation of the actual CDM concentration was calculated, and 5 samples of cell suspension predicted to contain 1 mg CDM were drawn. These samples were centrifuged at 10000 g and 4 °C for 10 min. Afterwards, the supernatant was poured off and the walls of the sample tubes were carefully dried with a paper tissue in order to just leave the dry cell pellet behind. These samples were then immediately frozen at -20 °C for later analysis.

Additionally, a CDM determination was carried out. For this purpose, 1 mL of fresh cell suspension was transferred into pre-weighed sample tubes and centrifuged to separate the supernatant from the cell pellet. The supernatant was decanted off the cell pellet and frozen

at -20 °C for further analysis. Then the cell pellet was washed with 1.6 mL of dH_2O and centrifuged again. After the supernatant was discarded, the cell pellet was resuspended in 1.6 mL of dH_2O, spun down and left to dry at 104 °C, before it was transferred to the exicator and subsequently weighed. The fluorescence of the fresh cell suspension was measured using the TECAN Infinite® 200 Pro plate reader controlled with the Tecan i-control (version 1.10.4.0) software. For the measurement, 100 µL of the cell suspensions (or dilutions where necessary) were transferred to 96 well flat bottom microplates with clear bottom (TECAN, Art.No.: 30122298) and measured at an excitation wavelength of 485 nm and an emission wavelength of 520 nm. From these measurements, the amount of produced GFP could be determined via a previously created calibration curve, which was obtained by measuring serial dilutions of a GFP standard the same way. Additionally, to the fluorescence measurements, the fresh cell suspension was analysed with a flow cytometer CytoFLEX S (Beckman Coulter). For this purpose, a 1:10 dilution of the culture broth was diluted serially 1:45 twice in sterile filtered PBS buffer. This 1:20250 dilution was then measured in the CytoFLEX S flow cytometer and forward and sideward light scatter signal, as well as the fluorescence of the cells at 525 nm after excitation at 488 nm was recorded. To obtain data representative of the whole population, the measured cell count was set to 15000 for all samples. The entire cell population was thereafter selected in a forward- versus sideward- light scatter diagram and analysed using the Kaluza Analysis software version 2.1 (Beckman Coulter).

4.7 Analytic procedures

4.7.1 Cell lysis – BioLector® plates

After the end of fermentation, the BioLector® cultivation plate was centrifuged for 10 min at 4 °C to separate biomass and supernatant. Thereafter, the supernatant was carefully removed, and the cells were resuspended in 200 µL lysis buffer (100 mM Tris/HCl, pH 8.2 supplemented with 0.25% (w/w) sodium deoxycholate). Then the cells were left to lyse while gently shaken on an orbital shaker for 1.5 h. The samples were then again centrifuged, and the supernatant was immediately used for dot plot analysis.

At the end of the seventh passage, a different cell lysis protocol was applied. For this, the remaining culture broth after sampling was adjusted to contain a calculated amount of 1 mg of CDM and centrifuged in the cultivation plate. Then the supernatant was discarded, and the cell pellets were resuspended in 100 µL of EDTA cell lysis buffer (50 mM TRIS/HCl pH 8 supplemented with 10 mM EDTA). The plate was then placed on a heated orbital shaker and left for one hour at 60 °C while being gently shaken. Afterwards, the samples were centrifuged once more, and the supernatant was used immediately for dot plot analysis.

4.7.2 Dot plot

For dot plot analysis, 3 µL of each sample was spotted on a nitrocellulose membrane (PROTRAN® Nitrocellulose Transfer Membrane, BA 83, pore size 0.2 µm) and left to dry for 10 min. Afterwards, the membrane was submerged in a blocking solution made of T-PBS buffer (9.55 g/L PBS Dulbecco, 0.5 mL/L Tween® 20) supplemented with 3% (w/w) of milk powder and left gently shaken for 1 h at room temperature. The membrane was then washed three times with T-PBS buffer for 5 min on an orbital shaker. In the next step, the membrane was incubated in an antibody solution specific against the kappa light chain of the Fab. For this, the antibody stock solution (Anti-Human Kappa Light Chains – Alkaline Phosphatase, Sigma A3813) was diluted 1:5000 fold in T-PBS supplemented with 1% (w/w) of milk powder. This solution was then poured on the membrane and left on an orbital shaker for one hour. Afterwards, the membrane was again washed three times with T-PBS buffer and left to develop for 10-15 min with 10 mL of staining solution prepared from SigmaFAST BCIP/NPT tablets (Sigma). During this process, special attention was payed to reduce the exposure to light, as the staining solution is light sensitive. The reaction was then stopped by washing the membrane with dH_2O multiple times and thereafter, the plot was dried and scanned for further analysis.

4.7.3 Cell lysis – 1 mg CDM samples

Cell pellets were thawed and resuspended in 200 µL of cell lysis buffer (30 mM Tris/HCl pH 8.2, 30 mM EDTA, 10 mM $MgCl_2$). In case of reducing cell lysis for samples containing GFP, NuPAGE® Sample Reducing Agent was added to the lysis buffer. Then 50 µL of lysozyme (2 mg/mL, 10000 U, AMRESCO) and 50 µL of benzonase solution (250 U/µL, Sigma Aldrich) were added. The mixture was incubated for 10 min on a shaker and 1.5% of Triton™ X-100 (Sigma-

Aldrich) were added before further shaken incubation for 10 min. Cell fragments were then separated from the soluble fraction via centrifugation at 15000 g at 4 °C for 10 min. The soluble fraction was afterwards immediately decanted into new reaction tubes and stored at -20 °C until needed for further analysis. The pellet containing the cell debris, as well as insoluble proteins in the form of inclusion bodies (IB), was then washed twice with 1 mL of washing buffer (100 mM Tris/HCL, pH 8.2) and then centrifuged as described above. The supernatant was then poured off and the remaining IB pellet was resuspended in 400 µL of IB-dissolution buffer (8 M Urea in 100 mM Tris/HCl, pH 8.2). In case of the reducing cell lysis protocol for GFP, NuPAGE® Sample Reducing Agent was added to the buffer. The IBs were left to dissolve on a shaker at room temperature for 30 min, before centrifugation as previously described. The supernatant containing the dissolved IBs was then decanted off the remaining pellet into a fresh tube and stored at -20 °C.

4.7.4 NuPAGE® electrophoresis

For electrophoresis, samples after cell lysis were thawed and diluted if necessary. Then a loading sample, consisting of 13 µL standard/sample, 5 µL NuPAGE® LDS Sample Buffer (4x) and 2 µL NuPAGE® Sample Reducing Agent (10x), was created in a fresh reaction tube. In case of non-reducing electrophoresis 2 µL of dH_2O were added instead of the reducing agent. The samples were then placed in a heating block for 10 min at 70 °C and 550 rpm. In the meantime, the electrophoresis buffer (NuPAGE® MES running buffer) was prepared from a stock solution and the electrophoresis chamber was assembled. For this, the electrophoresis gel (NuPAGE® 4-12% BisTris-Gel) was taken out of its package, rinsed with dH_2O and assembled into the chamber after removal of the comb. Then the chamber was filled with NuPAGE® MES running buffer. The gel was then loaded with 15 µL of the samples and 7 µL of standard. For standard electrophoresis, the Mark 12™ Unstained Standard (Invitrogen™) was used, while for western blots, a pre-stained protein ladder (PageRuler™ Prestained Protein Ladder 10-180 kDa, Thermo Scientific™) was loaded onto the gel. The electrophoresis chamber was then closed and connected to the power supply. The gels were run for 45 min at 200 V and 400 mA. After electrophoresis, the gel was washed with dH_2O and removed from its casing. Then it was transferred into a container with fresh fixing solution (50% (v/v) ethanol and 10% (v/v) acetic acid) and left for 30 min while gently shaken. After fixing, the solution was discarded and the gel was transferred into staining solution (1.16 g/L Coomassie Brilliant blue R250, 25% (v/v) ethanol, 8% (v/v) acetic acid) and left on a shaker to develop for 30 min. After staining, the gel was placed in de-staining solution (25% (v/v) ethanol, 8% (v/v) acetic acid) and gently shaken for an hour. Then, the de-staining solution was exchanged for a fresh one and the process repeated. After the second change of solution, the gel was left overnight to discolour completely and then scanned.

4.7.5 Fab - Western blot

For western blots, the samples were treated as described in the section NuPAGE® electrophoresis. After the electrophoresis, the samples were transferred out of the gel and onto a nitrocellulose membrane using the iBlot® Dry Blotting System (Invitrogen™). For this

purpose, iBlot® Transfer Stacks (Invitrogen™) were used. The stack was assembled in the iBlot® device according to the manufacturer's instructions and the membrane was blotted at 20 V for 8 minutes. After the disassembly of the stack, the membrane was retrieved and all further steps were performed as described in the section Dot plot, starting with blocking of the membrane.

4.7.6 Fab – ELISA

For the Fab - Enzyme Linked Immunosorbent Assay (ELISA), samples prepared via the non-reducing cell lysis protocol were thawed and diluted as needed with dilution buffer. This buffer consisted of PBS (0.92 g/L Na_2HPO_4, 0.2g/l KH_2PO_4, 0.2 g/L KCl, 8.0 g/L NaCl) supplemented with 0.1% (v/v) Tween® 20 and 1% (w/w) Bovine Serum Albumin. To minimize errors in sample preparation due to inaccurate pipetting, all sample dilutions were performed gravimetrically. For the preparation of the assay, a microtiter plate (Nunc MaxiSorp™, Invitrogen™) was coated with the capture antibody (Anti-Human IgG (Fab spec.), I5260, Sigma -Aldrich). This was done by diluting the antibody 1:400 in coating buffer (8.4 g/L $NaHCO_3$, 4.2 g/L Na_2CO_3) and transferring 100 µL of the antibody solution into each well of the plate. Afterwards, the plate was incubated over night at 4 °C. Then the plate was washed three times with washing buffer (PBS with 0.1% (v/v) Tween® 20) using the HydroFlex™ microplate washer (TECAN). To remove any access buffer, the plate was then dried by gently beating it onto a stack of paper tissues. The samples were prepared in another microtiter plate (Nunc™ MicroWell™ 96-well, Thermo Scientific™) by pipetting 140 µL of dilution buffer in each well of the rows A-G. Then 300 µL of the samples and standards (100 ng/mL, purified human Fab/Kappa, Bethyl P80-115) were applied in row H of the plate. Starting from row H, serial 1:2 dilutions were made by transferring 140 µL of each well to the next row. This was repeated for all rows before subsequently transferring 50 µL out of each well onto the coated plate. Then, the plate was incubated at room temperature for 1 hour on an orbital shaker. Afterwards, the plate was washed three times, before the second antibody (Anti-Human IgG (hinge region), 2A11, abcam ab7497) was applied. This antibody was diluted 1:1000 in dilution buffer and 100 µL of the solution were transferred into each well. Then the plate was left to incubate for 1 hour as described before. After incubation the plate was washed again three times and 100 µL of the third (enzyme-marked) antibody (Anti-Mouse IgG (Fab spec.) + Peroxidase, Sigma A2304) diluted 1:1000 in dilution buffer were pipetted into each well. Once more, the plate was left to incubate for 1 hour on an orbital shaker at room temperature. Then, the plate was washed again three times and the staining solution was prepared. For this purpose, a buffer tablet, supplied with the o-Phenylendiamindihydrochloride kit (Thermo Scientific™) was dissolved in 22 mL of dH_2O before addition of the 5 mg o-Phenylendiamindihydrochloride tablet. After all components were dissolved, 100 µL of the solution was added to each well and left to incubate until a colouration in the highest dilution of the standard could be clearly seen. Then the reaction was stopped by the addition of 100 µL of 1.25 M H_2SO_4 to each well. The plate was then measured in a multi-channel photometer (Infinte® M200, Tecan) at 492 nm and a reference wavelength of 620 nm. The data was then evaluated, and the sample concentrations were determined using the Magellan™ 6 (Tecan) program.

4.7.7 GFP – ELISA

GFP concentrations were determined additionally to fluorescence spectroscopy with an ELISA specific for GFP. This procedure was carried out analogous to the Fab – EISA, therefore only differences in the procedure as well as the used antibodies and dilutions are described here. Microtiter plates were first coated with the capture antibody (mAb mouse α GFP, G6539, Sigma) 1:4000 diluted in dilution buffer (for composition see section Fab – ELISA). After incubation and washing, 100 µL of dilution buffer (composition see Fab – ELISA) were applied to each well to block the unspecific binding sites of the capture antibody. During incubation of the plate for 2 hours on an orbital shaker at room temperature, the samples and standards were prepared. As standard, a 2 µg/mL GFP solution (recomb. GFP, ab 84191, Abcam) stored at -80 °C in 20 mM TRIS/HCl supplemented with 40% glycerol was used, which was subsequently 1:100 fold diluted with dilution buffer. After blocking of the antibody, and washing of the plate, all samples and standard dilutions were transferred onto the plate. After incubation and washing, 50 µL of the second - biotinylated - antibody (Polyclonal AB to GFP – Biotin, R1091B, Acris) diluted 1:50000 in dilution buffer were added to each well and left to incubate. The incubated plate was afterwards washed again and 100 µL of Streptavidin-Horseradish Peroxidase Conjugate (GE Healthcare, RPN4401V) diluted 1:1600 in dilution buffer were added to each well. This solution was left to incubate on the plate for one more hour, while gently being shaken. Staining and the rest of the procedure was carried out as described in the section Fab – ELISA.

5 Results and discussion

5.1 Chemostat cultivations

Aim of the chemostat cultivations was to evaluate the variation in long term performance of different genome integrated *E. coli* production strains. In total, three strains have been compared. Two of which utilizing the conventional T7 system and one carrying the novel 1lacO-A1 promotor system. Additionally, one of the strains using the T7 system was previously mutated and selected according to two parameters: Possession of good growth characteristics under long term production conditions and reasonable product formation. The hypothesis is, that the mutations in this strain allow it to outperform the unmutated T7 reference strain, regarding the stability of the production system while still producing reasonable amounts of product. This long-time cultivation was performed with GFPmut3.1 as an easy to produce model protein. BL21(DE3) strains were used in combination with the T7 system, while a BL21^{Q} variant carried the 1lacO-A1 construct. After an initial uninduced fed-batch with a calculated end biomass of 30 g/L CDM, the host cells were grown under carbon limitation in the chemostat at a constant growth rate of 0.1 h^{-1} for 45 generations. Throughout the chemostat cultivation, the cells were fully induced with IPTG, which was supplied via the feed medium to maintain a constant IPTG to CDM ratio of 20 µmol IPTG per g of CDM.

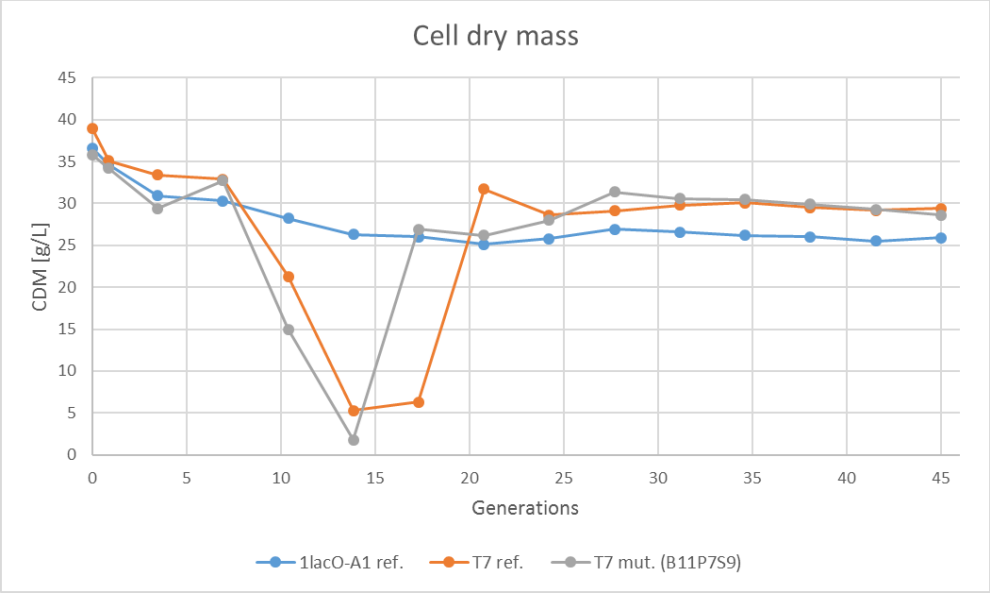

Figure 3: Cell dry mass concentrations over the course of the continuous fermentation

As can be seen in Figure 3, the cell dry mass after the fed-batch phase was slightly higher than expected, but stabilized after 3.5 generations into the continuous feed around the desired level of 30 g/L. In case of the T7 ref. and the T7 mut. (B11P7S9) strain, CDM started to decrease after 7 generations, dropping to 5.3 g/L for the reference strain and 1.8 g/L for the mutant.

However, the cells recovered in the course of the fermentation and remained stable at a density of around 30 g/L CDM thereafter.

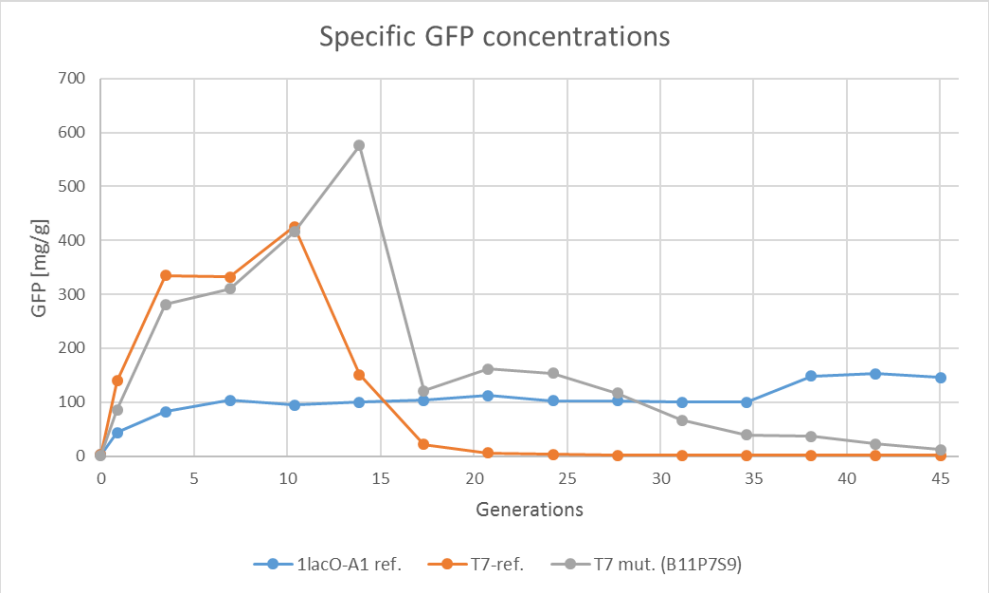

Figure 4: Specific GFP concentrations over the course of the continuous fermentation

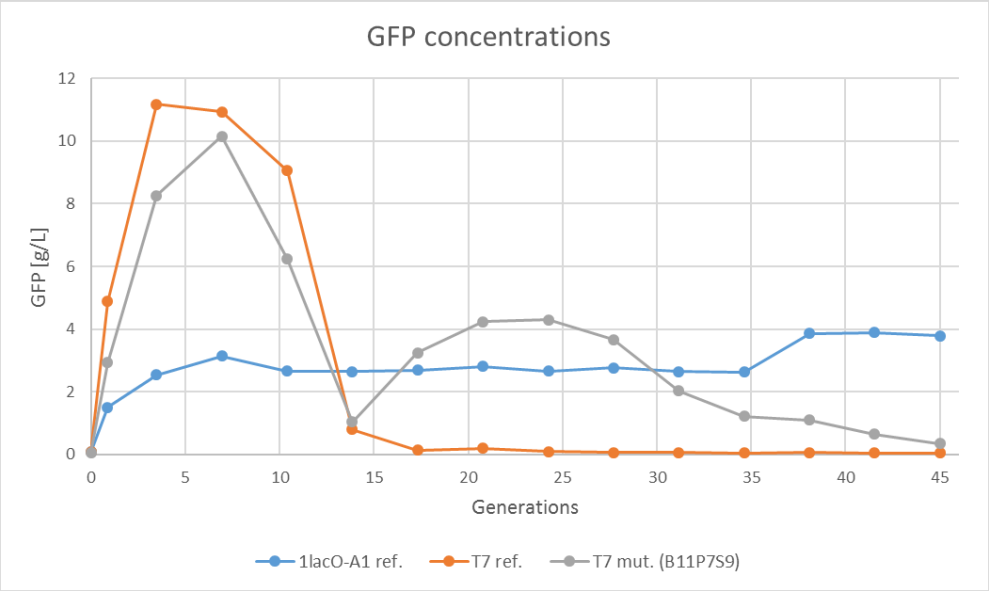

Figure 5: GFP concentrations over the course of the continuous fermentation

As can be seen in Figure 4 and Figure 5 (this data was obtained by fluorescence photometry measurements of the fermentation broth) in the T7 based reference production system, induction led to a rapid increase in GFP concentrations up to 11.2 g/L and 425 mg/g CDM, while the reference strain with 1lacO-A1 promotor system showed a product titre of 2.8 g/L and 100 mg/g CDM. The mutant strain showed an analogous response to induction to the T7 reference strain with similar product titres of 10.2 g/L and 567 mg/g CDM. This T7 mutant strain showed constantly dropping product titres after the first initial production spike and after 45 generations, the concentration of GFP in the fermentation broth dropped to 0.4 g/L. On the other hand, it could be observed, that the concentration of recombinant protein in the

fermentation broth of the unmutated T7 reference strain dropped to 0.1 g/L after only 17 generations. In case of the 1lacO-A1 production strain, the GFP concentration was held constantly at around 2.7 g/L after induction throughout all the 45 generations. Additionally, to fluorescence measurements of the fermentation broth, measurements of the supernatant were performed at selected timepoints.

Table 9: Measurement of the GFP concentration in the supernatant

	Generations	GPF in supernatant [g/L]	Total GFP concentration [g/L]
T7 ref.	10.4	0.8	9.1
	13.8	0.6	0.8
	17.3	0.1	0.1
T7 mut. (B11P7S9)	10.4	1.3	6.2
	13.8	0.7	1.0
	17.3	0.2	3.2
	20.8	0.2	4.2
	24.2	0.1	4.3
	27.7	0.1	3.7

As can be seen in Table 9, at some timepoints throughout the fermentation, substantial amounts of product could be detected in the supernatant. It should be noted that this table only holds results of timepoints, where the GFP concentration in the supernatant was greater or equal to 0.1 g/L. All other measured samples (data not shown in Table 9) had concentrations lower than 0.1 g/L.

Regarding total production of recombinant protein, a rough estimate was calculated utilizing the areas under the respective graphs in Figure 5. This was possible using the constant titre of A1 ref. between 13.8 and 24.2 generations combined with the known media flow for calibration. Via this rough approximation, the total product generated by the T7 ref., A1 ref. and T7 mut. strain was calculated to be 50 g, 57 g and 72 g respectively.

For further analysis, certain timepoints in the fermentation run have been chosen for SDS page analysis. This was done to estimate the ratio of soluble product in the cell lysate (CL) to insoluble inclusion bodies (IB) formed during the production phase.

In Figure 6 and Figure 7, the results of the SDS Page can be seen, where strong bands around 27 kDa are detectable, which is approximately the size of monomeric GFPmut3.1 (UniProtKB, 1995).

Table 10: Sample description of SDS Page 1

Lane	Sample
M	MW Standard: Mark 12™
2	1lacO-A1 ref. - 3.5 Generations CL
3	1lacO-A1 ref. - 3.5 Generations IB
4	T7 ref. - 3.5 Generations CL
5	T7 ref. - 3.5 Generations IB
6	T7 mut. - 3.5 Generations CL
7	T7 mut. - 3.5 Generations IB
8	1lacO-A1 ref. - 13.8 Generations CL
9	1lacO-A1 ref. - 13.8 Generations IB
10	T7 ref. - 13.8 Generations CL
11	T7 ref. - 13.8 Generations IB
12	T7 mut. - 13.8 Generations CL
13	T7 mut. - 13.8 Generations IB
14	1lacO-A1 ref. - 24.2 Generations CL
15	1lacO-A1 ref. - 24.2 Generations IB

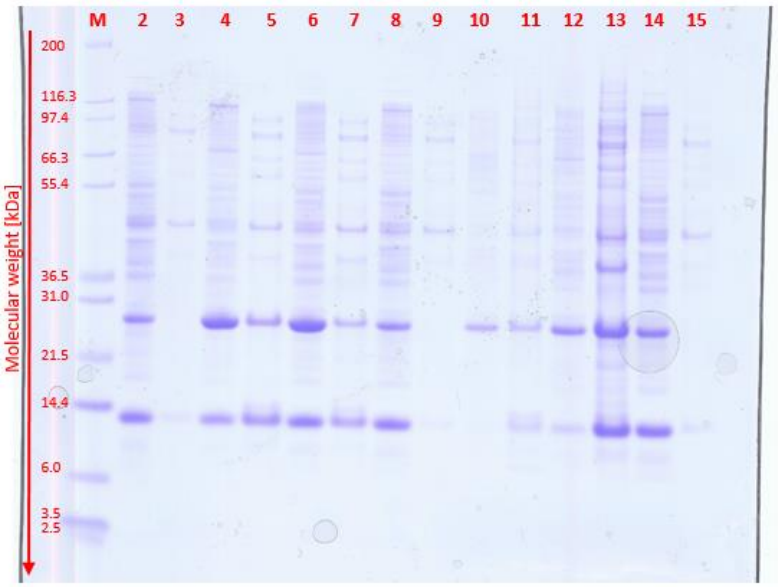

Figure 6: SDS Page 1 of selected points during the fermentation

Table 11: Sample description of SDS Page 2

Lane	Sample
M	MW Standard: Mark 12™
2	T7 ref. - 24.2 Generations CL
3	T7 mut. -24.2 Generations CL
4	T7 mut. -24.2 Generations IB
5	1lacO-A1 ref. - 34.6 Generations CL
6	1lacO-A1 ref. - 34.6 Generations IB
7	T7 ref. - 34.6 Generations CL
8	T7 mut. - 34.6 Generations CL
9	1lacO-A1 ref. - 45 Generations CL
10	1lacO-A1 ref. - 45 Generations IB
11	T7 ref. - 45 Generations CL
12	T7 mut. - 45 Generations CL

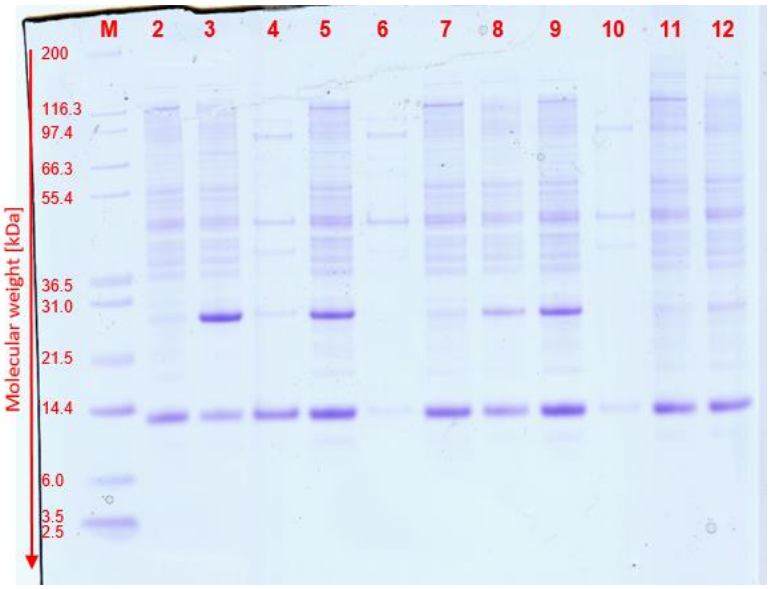

Figure 7: SDS Page 2 of selected points during the fermentation

To confirm the product concentrations previously shown, which have been obtained by fluorescence photometry, selected points in time of the continuous fermentation were chosen and analysed via ELISA.

Table 12: Results of ELISA

Generations	Strain	Product titre	
3.5	**1lacO-A1 ref.**	88.4	mg/g CDM
		2.73	g/L
	T7 ref.	376.5	mg/g CDM
		12.58	g/L
	T7 mut. (B11P7S9)	337.5	mg/g CDM
		9.92	g/L
13.8	**1lacO-A1 ref.**	104.3	mg/g CDM
		2.74	g/L
	T7 ref.	34.5	mg/g CDM
		0.18	g/L
	T7 mut. (B11P7S9)	95.2	mg/g CDM
		0.17	g/L
24.2	**T7 mut. (B11P7S9)**	149.1	mg/g CDM
		4.17	g/L
45.0	**1lacO-A1 ref.**	103.0	mg/g CDM
		2.67	g/L

Throughout the course of the fermentation, the cells were also analysed using a Flow Cytometer to detect the formation of subpopulations and assess the stability of the host strain.

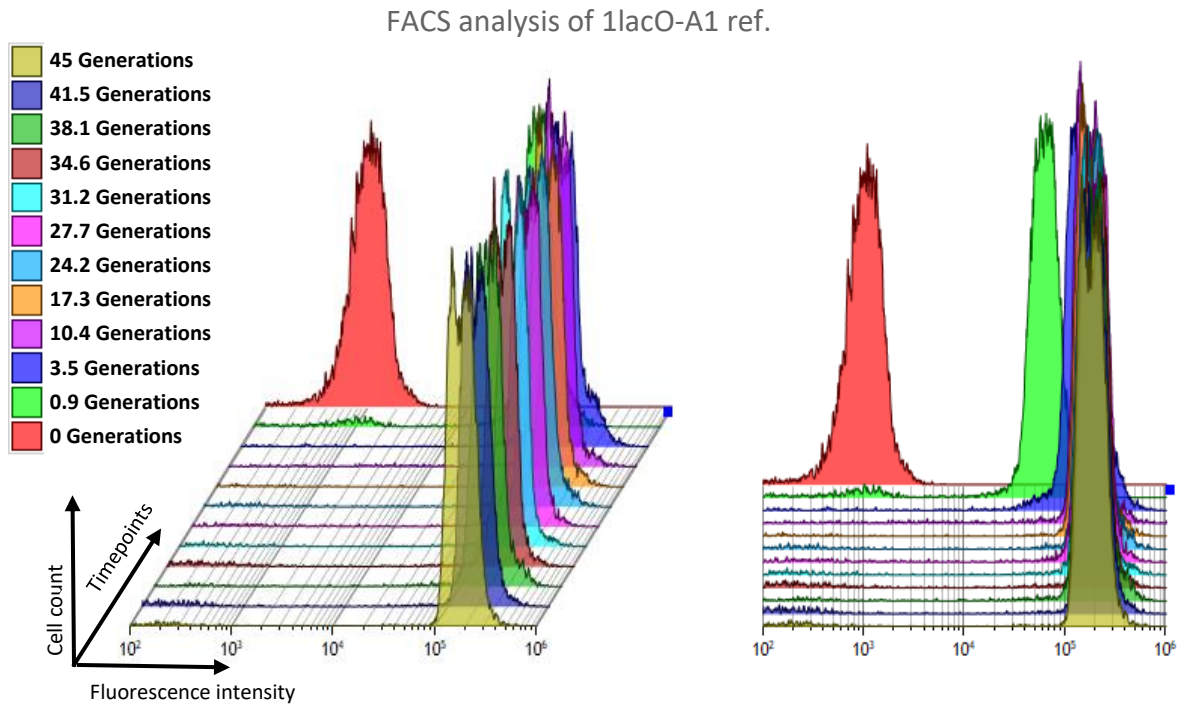

Figure 8: FACS analysis of 1lacO-A1 reference strain in the course of the continuous fermentation

Figure 8 shows the FACS analysis of the 1lacO-A1 reference strain during the continuous fermentation. The diagram here is shown in duplicates with another orientation to enhance the ability to observe shifts in fluorescence intensity and the formation of subpopulations. Upon analysis of Figure 8 it can be seen, that there is an initial shift in fluorescence intensity

occurring at the first three measurement points as a reaction to induction. The resulting induced cell population has a fluorescence intensity of around $2*10^5$ arbitrary units at the peak maximum. Furthermore, for the rest of the continuous fermentation the population of cells is very homogenous, and no subpopulations are formed.

Figure 9 shows the FACS analysis of the T7 reference strain during the continuous fermentation phase. As already described above, the diagrams to be seen here are duplicates with different orientations. In the FACS analysis of the T7 ref. strain, a shift to higher fluorescence intensity upon induction can be observed as expected. At the peak maximum, the fluorescence intensity observed is around 10^6 arbitrary units. The population then stays homogenous for only a very short period, forming a less producing subpopulation already in generation 7. This subpopulation then proceeds to grow and around 14 generations after induction, the population of non-producing cells is larger than the group of GFP-producing cells.

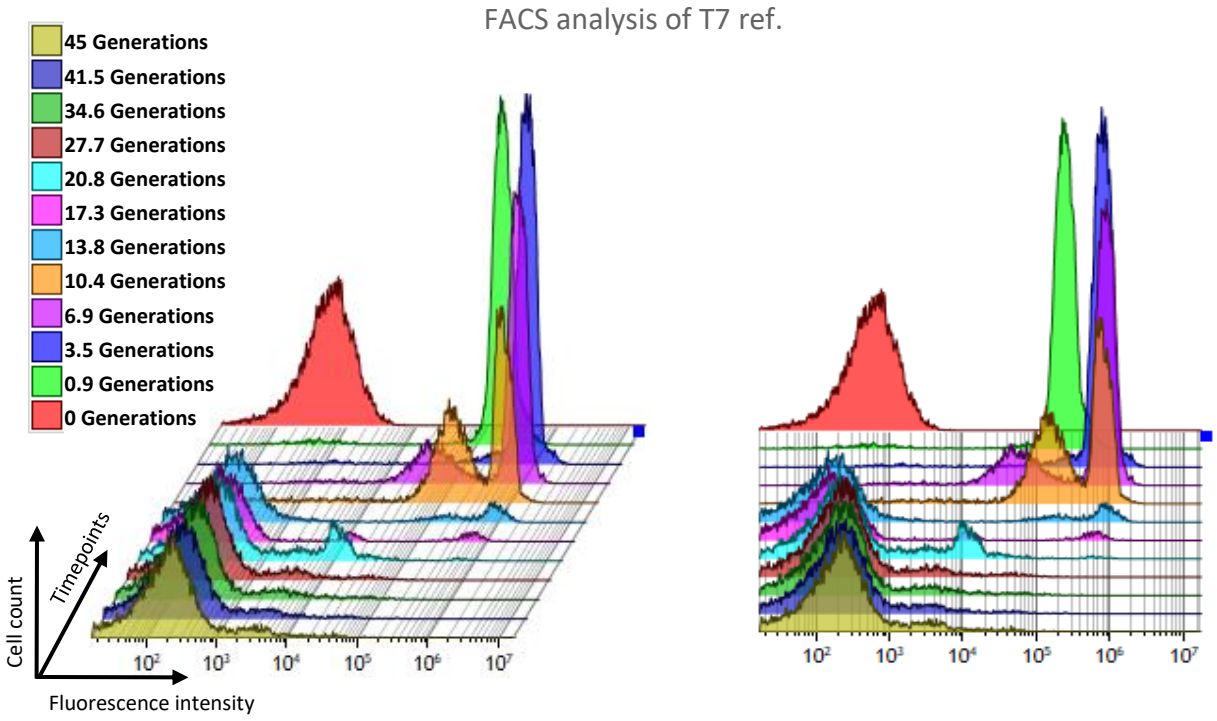

Figure 9: FACS analysis of T7 reference strain in the course of the continuous fermentation

In Figure 10, the FACS analysis of the strain T7 mut. (B11P7S9) during continuous fermentation can be seen. The same expected shift in fluorescence intensity, as has been already observed in the two other production systems can be noticed at the first few timepoints. The resulting induced cell population has a fluorescence intensity of around $8*10^5$ arbitrary units at its peak maximum. After induction, the producing population stays homogenous and stable for about 14 generations, where for the first time, a non-producing subpopulation can be detected. Interestingly, the main producing population then shows a small shift to lower fluorescence intensities, but then jumps again to its original place over the course of the fermentation. From generation 17 up to generation 45, a steady growth of the non-producing subpopulation

can be observed, going hand in hand with a declining number of producers. Finally, after 45 generations, almost only non-producers can be observed in the FACS diagram.

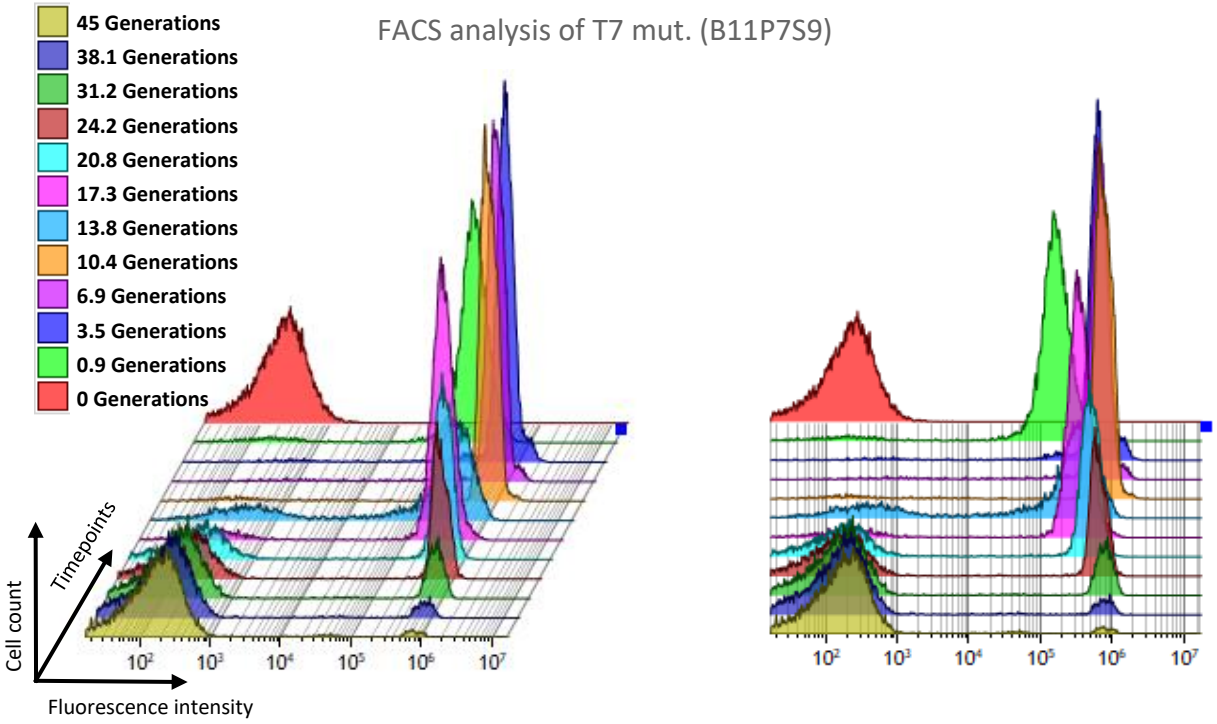

Figure 10: FACS analysis of T7 mut. (B11P7S9) strain in the course of the continuous fermentation

Upon direct comparison of the long-term stability under production conditions of the three tested strains, it can be clearly seen, that the 1lacO-A1 system is the most stable, showing constant biomass, product concentration and a very homogeneous cell population over the course of the continuous fermentation. On the contrary, both the T7 ref. and the T7 mut. strain could not hold the biomass concentration of 30 g/L and started to drop simultaneously after about 7 generations in the chemostat. This could be due to high cellular stress experienced by recombinant protein producing cells, which impedes their ability to grow at the required growth rate of 0.1 h^{-1}. Consequently, these cells are washed out of the bioreactor and subpopulations with low or no product formation and therefore low cellular stress start to outgrow the producing population. Shortly after the beginning of the decrease in cell density, a spike of extracellular product was detected for both strains, ranging from 0.8 g/L for the T7 ref. strain to 1.3 g/L for the T7 mut. culture, suggesting cell lysis took place at the same timepoint. In the FACS data, one can also simultaneously see the subpopulations of cells producing less product, which were formed during this decline in cell density of both T7 strains. For the T7 ref. strain, this is already observable at around 7 generations into continuous fermentation, while the T7 mut. strain maintained a uniformly producing population until generation 14. In this sense, the mutant strain outperforms the reference strain, but ultimately suffers the same fate. Also, in length of the production, the T7 mut. strain

outruns the T7 ref. strain, as production comes to a halt at around 17 generations for the T7 system and at 45 generations for the mutant strain. Regarding overall production, the T7 mutant also clearly surpasses both strains producing about 1.4 times as much product as the T7 ref. strain and 1.2 times as much as the A1 ref. strain. On the other hand, it can be stated, that while the T7 mutant showed a steadily increasing subpopulation of non-producing cells, outgrowing the producing ones, the A1 reference strain showed no signs of dwindling production or the formation of subpopulations. These results show that this strain can be cultivated under production conditions continuously for even more than 45 generations, without forming non-producing subpopulations. Due to this characteristic, the system would have eventually produced more recombinant protein than the T7 mutant. Moreover, the weakness of this promotor system averts inclusion body formation, while both other strains showed significant amounts of target protein in IBs, which could hinder downstream processing.

5.2 BioLector® cultivations with BL21^{Q}TN7::<1lacO-A1-dFTN2-t.Z>

One of the main problems in the production of toxic or difficult to produce recombinant proteins is to maintain a stable production system, whilst still reaching sufficient product titres (Saida et al., 2006). The genome integrated weak 1lacO-A1 promotor system, which utilizes host RNAP, was developed according to the theory, that it would minimize cellular stress upon recombinant protein production in relation to the widely used T7 system. This makes it an excellent candidate for long term cultivations, as the production system is thought to show enhanced stability. Furthermore, it is thought, that one could enhance the production strain's performance by cultivating it for extended periods of time in a fully induced state to let it mutate spontaneously (Schuller et al., unpublished data). Afterwards, clones with good growth characteristics and reasonable product formation rates can be selected for further experiments. For this purpose, six sequential passages with 48 parallel microbioreactors were

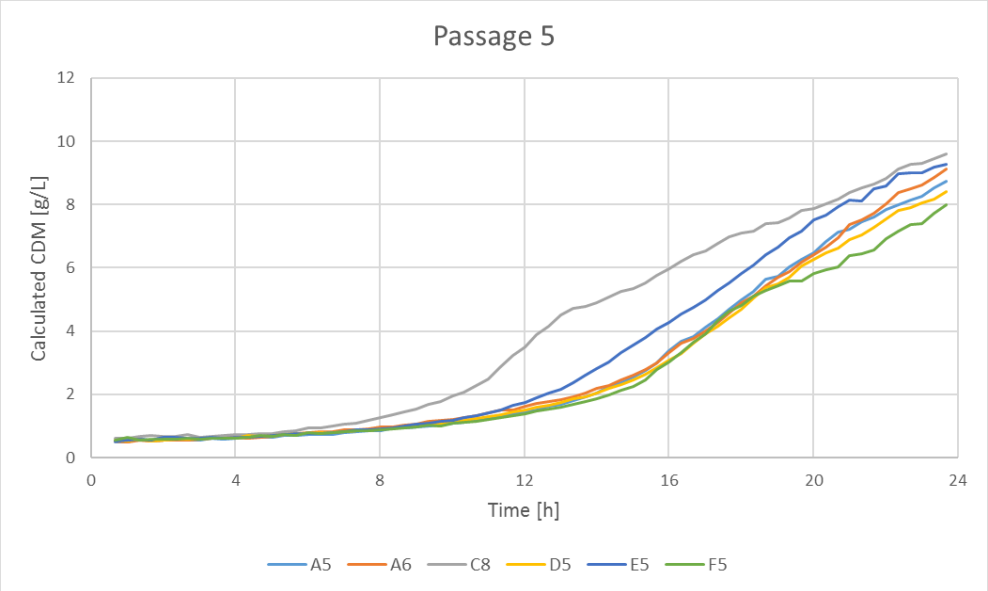

Figure 11: Growth curve of the selected mutants in passage 5

carried out in the BioLector® fermentation device using a fully induced fed-batch like synthetic medium. The results were then analysed regarding the growth behaviour of the obtained clones and their production capabilities. Single colonies of promising cultures were then isolated and tested in an additional seventh cultivation passage.

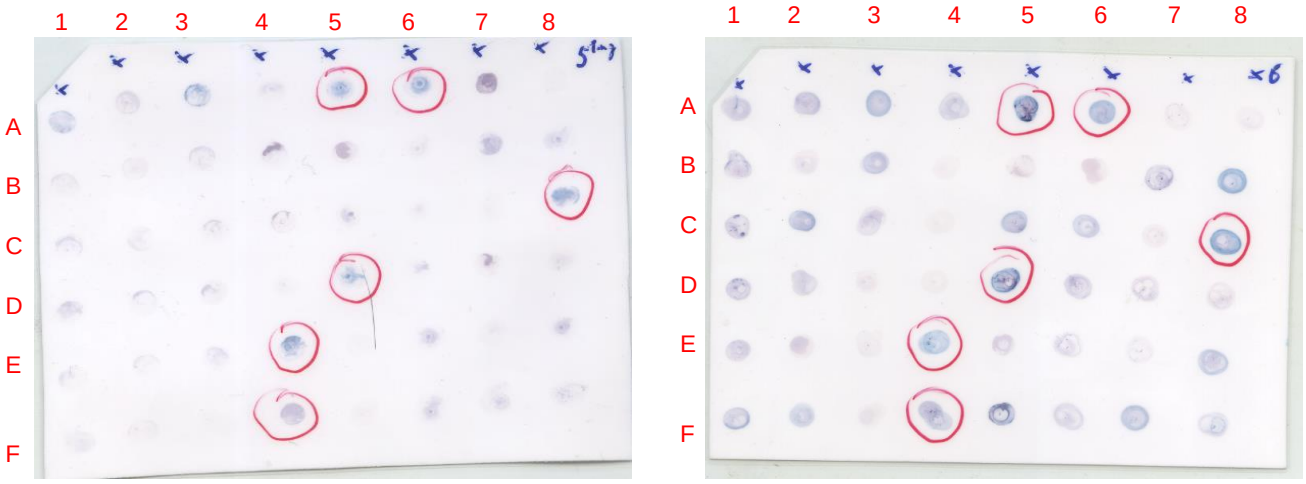

Figure 12: Dotplots of passage 5 and 6 with indication of the selected clones

Figure 12 shows the dotplots of the fifth and sixth serial cultivation. The indicated clones with good production capabilities throughout passage 5 and 6 were chosen, as their growth behaviour was satisfactory as well (see Figure 11 and Figure 13).

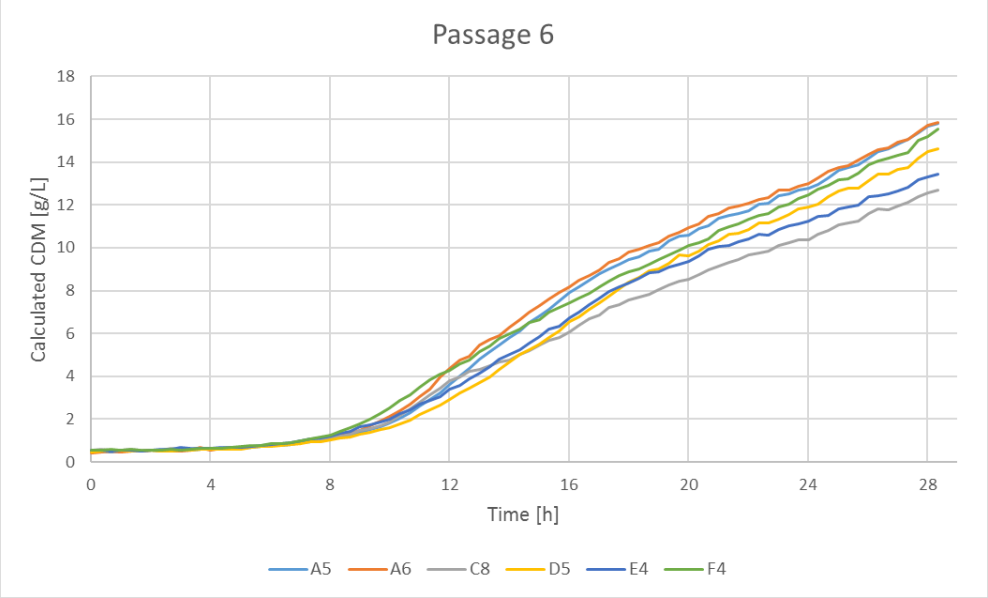

Figure 13: Growth curve of the selected mutants in passage 6

Results of the cultivation of selected single colonies of the beforementioned chosen mutants of passage 6 can be seen in data from passage 7 (Figure 14 and Figure 15). Resulting from the dotplot analysis, four clones with reasonable growth and production characteristics were

chosen from passage seven for further analysis. This was performed via Western Blot and ELISA (see Figure 16 and Table 13).

Table 13: Results of Fab ELISA of the selected clones

Sample	Fab [mg/g CDM]	Fab [mg/L]
C6 - CL	0.21	1.35
D7 - CL	0.64	6.92
E6 - CL	0.71	7.75
F2 - CL	0.55	6.08

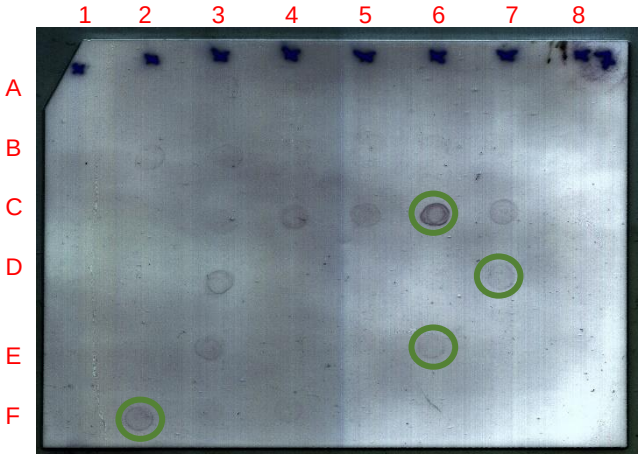

Figure 14: Dotplot of passage 7 with indication of selected clones for further analysis

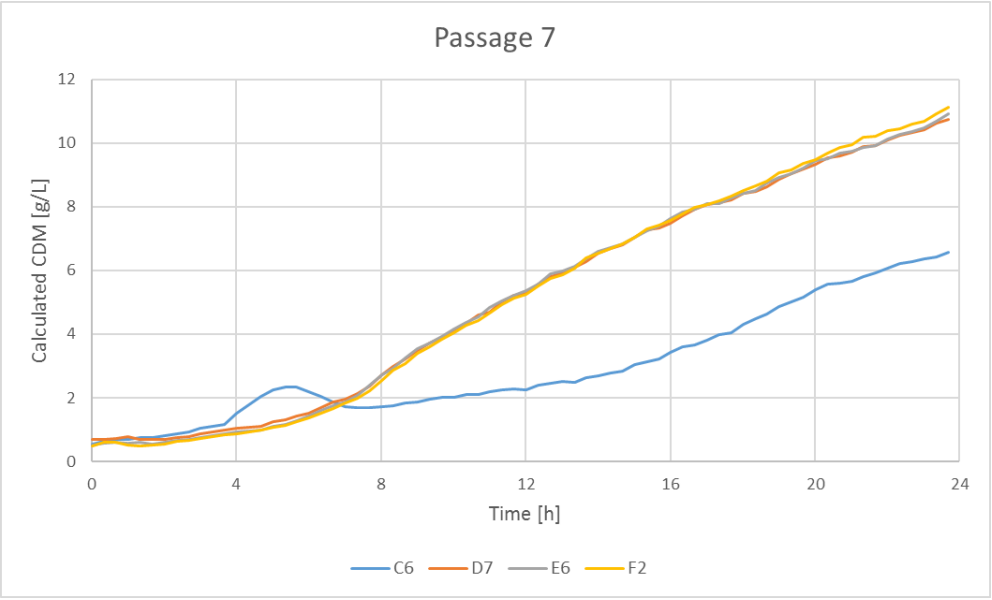

Figure 15: Passage 7 of selected clones for further analysis

Table 14: Sample description of Western Blot

Lane	Sample
1	Page Ruler™ Prestained Protein Ladder
2	Fab Standard Stock [5 µg/mL]
3	C6 - CL
4	C6 - IB
5	C6 - Supernatant
6	D7 - CL
7	D7 - Supernatant
8	E6 - CL
9	E6 - Supernatant
10	F2 - CL
11	F2 - IB
12	F2 - Supernatant

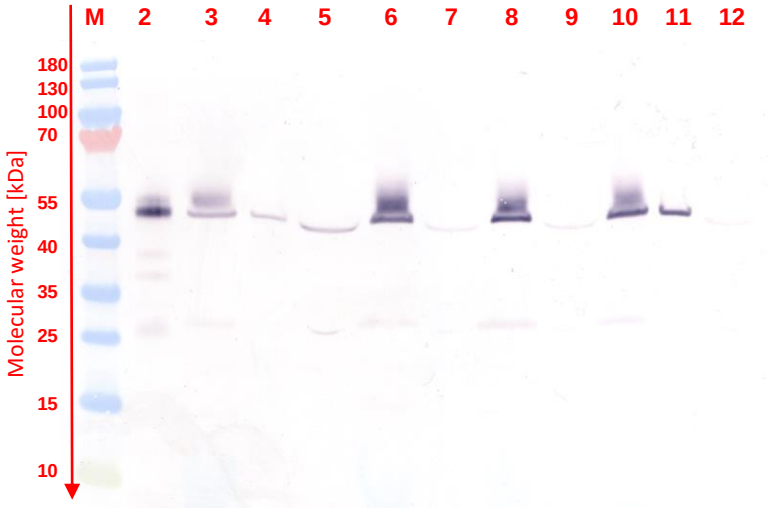

Figure 16: Western Blot of selected clones of passage 7

As can be seen in the Western Blot Analysis of passage 7 (see Figure 16), clone number D6 (lane 3 to 5), shows less intense bands than the other screened mutants. Furthermore, substantial amount of product can be observed in the culture supernatant of this clone (lane 5). As for mutants D7, E6 and F2, the Western Blot shows approximately the same product concentration in the cell lysate and only a very faint band in all supernatant samples. Free light chain could also be detected via the Western Blot, but the bands at approximately 25 kDa in lanes 3, 6, 8 and 10 are very weak. Results of the ELISA (Table 13) show, that clone E6 outperformed all other mutants with a specific product titre of 0.71 mg/g CDM, followed by D7, F2 and C6.

Upon direct comparison of the four obtained production strains, it can be clearly seen, that C6, which was only chosen because it shows the most intense dot in the dotplot of passage 7, produces the least amount of protein. Furthermore, lane 5 of the Western Blot suggests, that cell lysis took place during fermentation, as a portion of the produced protein was detected outside of the cells. This is likely caused by high cellular stress and would explain the poor growth of C6 in passage 7. All the other selected strains show unhindered cell growth peaking at a calculated CDM of about 11 g/L, which would suggest they mutated in such a way to be able to cope with the cellular stress experienced. It is also remarkable, that the Western Blot shows very little amount of free light chain in all clones, which proposes, that nearly all of the produced Fab is present in the correctly folded and active configuration. On the other hand, it can be seen, that even by expression via the weaker 1lacO-A1 promotor system, inclusion bodies are formed (lane 12 of Figure 16). With clone E6 producing a specific fab concentration of 0.71 mg/g CDM and possessing good growth characteristics, a reasonable production platform was created, which needs to be further tested in a fully controlled fermentation environment to assess its production characteristics and long-term stability.

5.3 Trial of 1lacO-A1 promotor system in HMS174^{Q}

To further research into strain dependent differences in production performance of recombinant proteins utilizing the 1lacO-A1 promotor system, a K-Strain of *E. coli* - HMS174^{Q} - was chosen as candidate for comparison to the B-Strain BL21^{Q}. As this new genome integrated production platform has not been generated yet, the main aim was to successfully integrate the 1lacO-A1-dFTN2-t.Z and 1lacO-A1-GFPmut3.1-t.Z expression cassettes into the genome of the host strain at the TN7 site.

5.3.1 Preliminary tests

Upon repeated endeavour to integrate the expression cassettes into the genome of the host strain and three negative results in a row, further testing of the strain and other K-Strains for comparison and troubleshooting was carried out. The first hypothesis for the negative results was a mutation in the used HMS174^{Q} cells or the integration cassettes. This was quickly ruled out as all sequencing results of the 1lacO-A1-GFPmut3.1-t.Z integration cassette, as well as the tested TN7 and lacI^{Q} site showed no alterations. However, there was a mutation found in the -35 region of the lacO in the 1lacO-A1-dFTN2-t.Z integration cassette, resulting from amplification of an already altered plasmid.

Furthermore, plasmid trials with the original, unaltered HMS174, as well as HMS174^{Q} were carried out. For this purpose, the plasmids pETk1lacOA1tZ.c-GFPmut3.1, pETk1lacOA1tZ.c-dFTN2 and pETk2lacOT5tZ.c-GFPmut3.1 were successfully electroporated into the cells. Upon induction it could be seen that the production system utilizing the T5 promotor showed good growth as well as production of GFPmut3.1 in both strains, while the 1lacO-A1 system hardly grew nor produced anything. Additionally, loss of plasmid was detected in both strains. In case of the 1lacO-A1 system with dFTN2, both host strains grew normally. Upon sequencing of the plasmids, the already discovered mutation in the -35 region of the lacO was found again, but all other results showed no alterations.

To gain insight whether this problem is K-strain specific, plasmid trials with *E. coli* RV308 were carried out. Once again, the same pET plasmid was used with GFPmut3.1 as target gene. The only difference in systems being the promotors, namely 1lacO-T5, 2lacO-T5, 1lacO-A1 and 2lacO-A1. Transformation of all the strains was successful, but upon induction all variants showed little or no production coupled with little growth and plasmid loss.

In a final pET plasmid trial with HMS174^{Q} and GFPmut3.1, the 1lacO-T5, 2lacO-T5, 1lacO-A1 and 2lacO-A1 promotor systems were tested. After successful transformation and induction, only the T5 promotor-based systems showed growth and production with the 2lacO variant outperforming the 1lacO variant. All A1 promotor systems showed no production, little growth and plasmid loss in induced culture.

As this seem to be a promotor specific problem in which neither the 1lacO-A1 or 2lacO-A1 variant are working in HMS174 and RV308, a closer look at the underlying mechanisms may provide a solution. The 1lacO-A1 system has the A1 promotor embedded into the lac operator itself. This results in a competition between the lac repressor and the sigma 70 factor for the

same binding site. If the binding strength of the transcription machinery is now greater than that of the lac repressor, the expression system cannot be controlled and continuously produces transcripts, stressing the host organisms. If that is the case, a system is needed, where the binding strength of the lac repressor outperforms that of the transcription factor and RNAP core to yield a tight and inducible expression system. Several mutations in the sigma 70 factor gene *rpoD* have been reported, which change transcription initiation rates. Among these is a Y571H mutation, which reduces the ability of the sigma 70 factor to bind the RNAP core (Jishage et al., 2002). This exact alteration occurs in BL21, therefore reducing the binding affinity of the sigma 70 factor to the RNAP core, which yields a tight, inducible expression system. On the other hand, HMS174 as well as RV308 have a Tyrosine at position 571, resulting in higher binding affinity of the sigma 70 factor to the RNAP core and yields an uncontrollable expression system. This provides a possible explanation for the observed behaviour of insufficient cell growth and production in induced culture, as high cellular stress is thought to be experienced by the host organism, resulting in plasmid loss.

5.3.2 lacI_opt trials

To counterfeit this promotor specific problem and to flip the balance from a highly leaky system to a controllable one, the lacI_opt plasmid was created. The goal was to increase the binding affinity of the lacI molecule to the operator and therefore limit sigma 70 binding to the promotor. It has been reported, that a W220F mutation in the lac repressor leads to 10-fold tighter repression, while still being inducible with IPTG (Gatti-Lafranconi et al., 2013). It was sought to exploit this phenomenon by creating a variant of the pETk1lacOA1tZ.c-GFPmut3.1 carrying the W220F mutation. After successful creation of this plasmid, it was transformed into HMS174^{Q} for further testing. Upon induction, a slightly enhanced product formation and growth could be observed.

While cells carrying the W220F mutation in the lac repressor showed slightly better production and growth than those with the wt. lac repressor, the results were still not satisfying. It is possible, that the enhanced binding strength of the lac repressor is still too low to yield a stable expression system.

5.3.3 Trials with alterations in the *rpoD* gene

After unsatisfying results with the W220F lacI_opt alteration, the next effort to obtain a working 1lacO-A1 promotor system in HMS174^{Q} included changing the *rpoD* gene itself, which codes for the sigma 70 factor. The hypothesis behind this idea was, that the promotor system would work, if one changes the sequence of the *rpoD* gene to make it identical with the sequence known from the strain BL21. The sequence of *rpoD* in HMS174 and BL21 is only different in one single position, being an alteration on position 1711 in the *rpoD* gene from cytosine in BL21 to thymine in HMS174, which leads to an modification in the amino acid sequence on position 571, namely histidine in BL21 and tyrosine in HMS174 (see Figure 17).

Figure 17: Difference in the rpoD gene BL21 vs. HMS174

To achieve this, a genome integration cassette containing the changed *rpoD* gene targeted to the original *rpoD* genomic site was created. This integration cassette featured approximately 50 bp long overhangs on the 5' and 3' end. Upon integration of the construct and colony screening, no positive transformants could be found. It is especially remarkable, that while the antibiotic resistance gene together with the FRT sites were correctly integrated, the base that was originally sought to be exchanged, remained unaltered in every single colony that was sequenced. Figure 18 shows an example of the sequencing results, which were analogous for every colony screened.

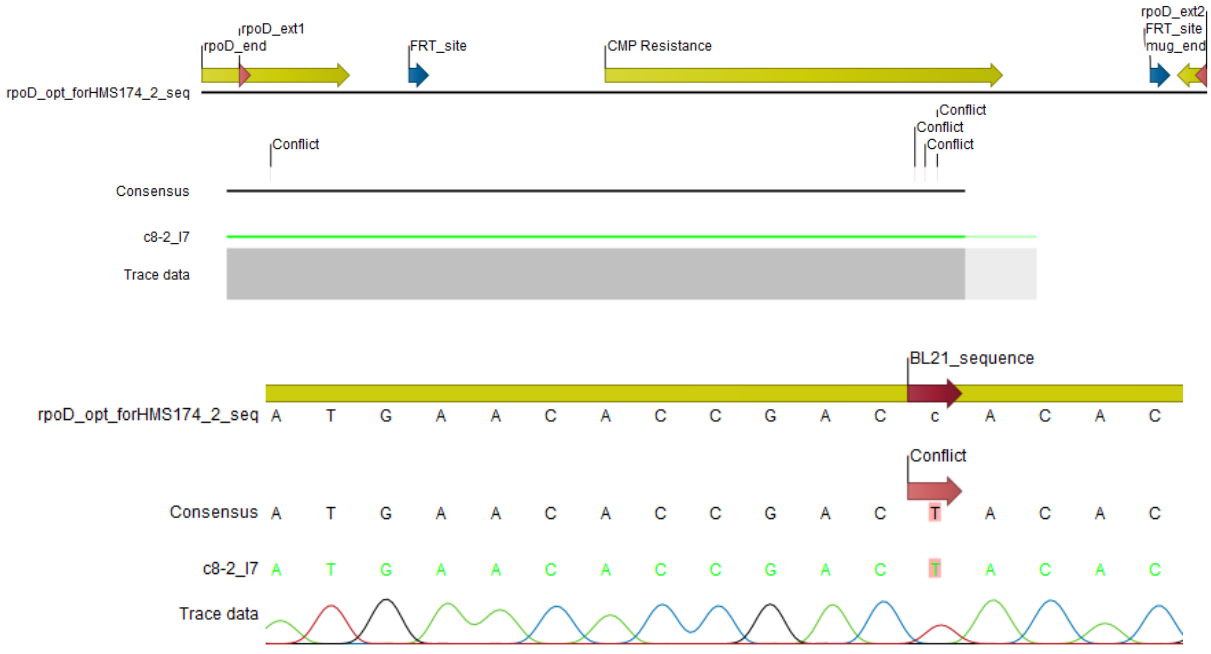

Figure 18: Sequencing result of an unsuccessfully transformed colony

To change these negative results, the genome integration cassette was elongated on the 5' end to enhance the base pairing upstream of the alteration. This new integration cassette with a 152 bp overhang was then integrated into HMS174^{Q}. Although, the same phenomenon as previously observed occurred, some colonies showed a mixed signal on the position in question, as can be seen in Figure 19. But all efforts to isolate a successful recombinant failed.

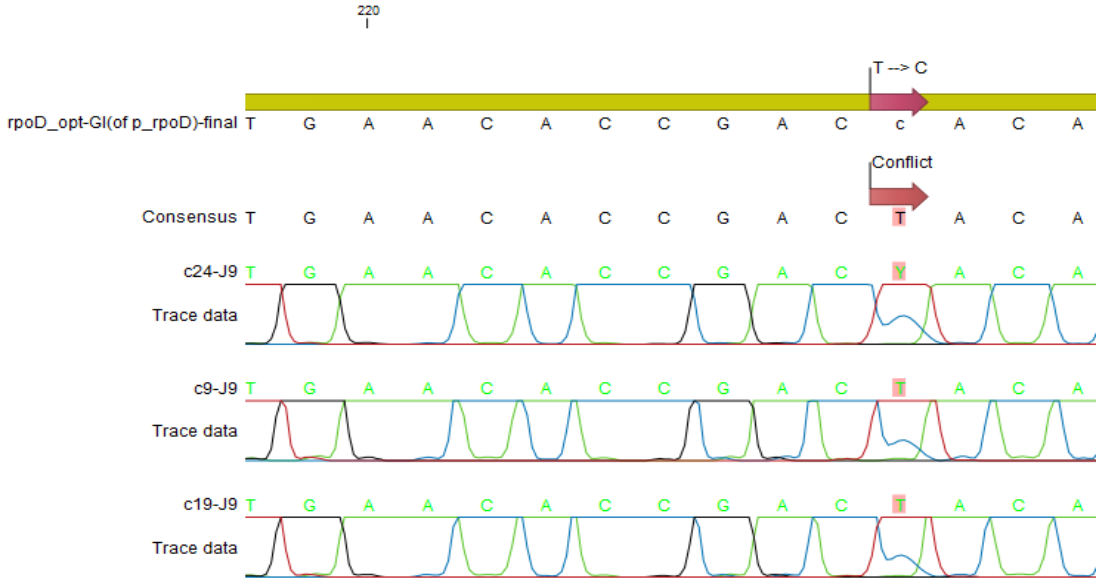

Figure 19: Sequencing results with elongated 5' end

As described by Tomatis *et al*., a by directed evolution discovered mutation in *rpoD* leads to a significant improvement of GPCR production, a G-protein coupled receptor, in *E. coli* which is toxic for the cells. They found, that an E575V mutation in *rpoD* was responsible for the host overcoming the stress imposed by recombinant protein production. As the position in question plays an important role in interaction between RNAP and promotors, this mutation might affect transcription initiation rates. They found that the mutant strain expressed the target protein at a lower rate but was able to produce it over elongated time periods. This elongated expression time and cell growth suggests a lowered transcription initiation frequency facilitated by the mutation in question (Tomatis et al., 2019).

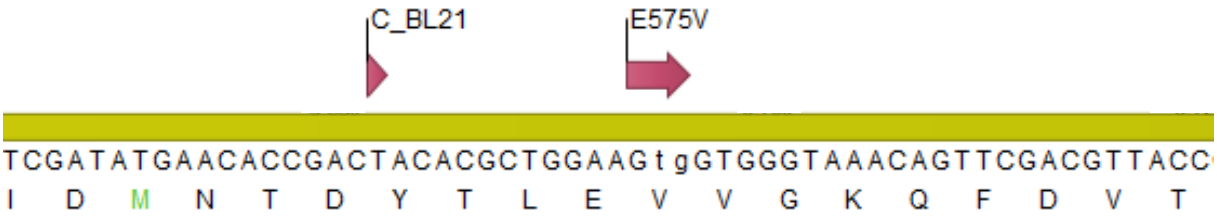

Figure 20: Sequence of E575V alteration in the rpoD gene (changed bases shown in small letters)

As our hypothesis is based on the theory, that the interaction between the sigma 70 factor and the RNAP core is too strong in HMS174^{Q} for the 1lacO-A1 promotor system to be tightly controlled, the E575V mutation in the *rpoD* subunit of the sigma 70 factor could provide a solution. For this purpose, another genome integration cassette was created carrying the E575V alteration, which was aimed to change the bases at position 1724 and 1725 in the *rpoD* gene. This cassette had an even further elongated 5' end of 241 bp upstream of the alterations. When the integration was performed with this new integration cassette, once again the alterations have not been integrated into the cells (see Figure 21).

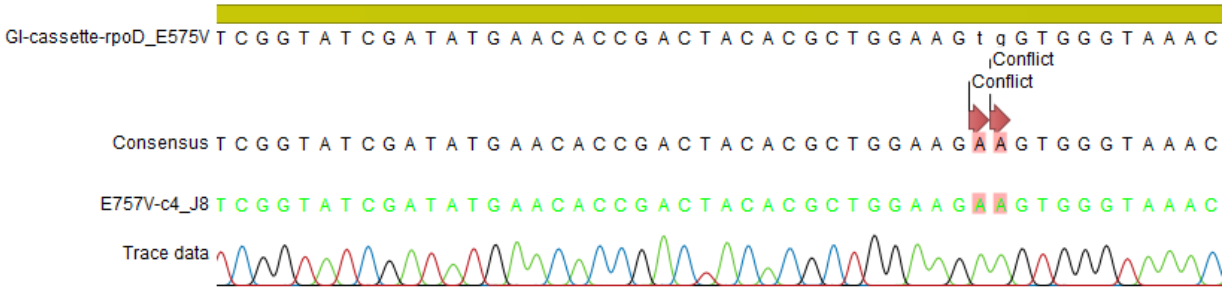

Figure 21: Sequencing result of rpoD E575V integration

Contrary to published data, where the *rpoD* gene was successfully altered and the desired mutation even discovered during a directed evolution experiment (Tomatis et al., 2019), all attempts to alter the sequence of the *rpoD* gene failed. It could be possible, that this is a strain specific phenomenon, as Tomatis *et al*., 2019 used the *E. coli* K-strain BW25113 for their experiments while HMS174, a different K-Strain, was used here. Another explanation for the observed behaviour would be, that the genome integration was indeed successful, but the recombinants could simply not grow well due to the altered sigma 70 factor, which is the main transcriptional factor in the exponential growth phase of *E. coli* (Jishage and Ishihama, 1995). This would also explain the mixed sequencing results observable in Figure 19, where a small subpopulation was found to have the desired alteration, but subsequently all efforts to isolate a successful recombinant of these mixed colonies failed.

6 Summary and Outlook

6.1 Chemostat cultivations

The production system based on the novel 1lacO-A1 promotor, utilizing the host RNAP showed greatly enhanced stability compared to T7 RNAP based constructs. Utilizing the host RNAP, it was possible to produce the recombinant protein throughout the entire 45 generations of the cultivation at a constant titre. On the other hand, the T7-system-based production strains exhibited much higher production rates, but genetic stability was limited to 7 generations under production conditions. At this timepoint producing cells lost their ability to maintain the constant growth rate of 0.1 h^{-1} and due to wash out the cell density decreased. In parallel the first non-producing cells started to grow and finally formed a non-producing population in the bioreactor. However, the genetic stability of genome integrated systems, independent from the promotor/RNAP combination used in this study is vastly superior to any plasmid-based *E. coli* system.

Based on the generated results, the weak 1lacO-A1 promotor system utilizing host RNAP with lower production rates but superior stability is thought to be especially suitable for continuous production of challenging proteins. In case of easy to produce proteins the T7 strains especially the mutant are attractive candidates, as they allow for extremely high product formation rates and titres for at least 7 doublings which can be a big gain compared to conventional fed-batch

production processes. The mutated T7 strain outperformed the reference T7 strain in both stability and recombinant protein production, making it also a better candidate for fed-batch cultivations. Since in all these cultivations GFPmut3.1, an easy to produce protein, was used, more research with different recombinant proteins has to be done to evaluate the concept of continuous production in chemostat-mode on a broader base.

6.2 BioLector® cultivations with BL21^{Q}TN7::<1lacO-A1-dFTN2-t.Z>

The isolation of Fab producing mutants from serial passages in fed-batch like conditions was successful. Four producer candidates with good growth performance have been isolated, with the best performing strain showing a titre of 0.71 mg POI/g CDM. Despite this relatively low titre, the isolated mutant strains could prove to be a valuable production system in continuous fermentation, as they have also shown to produce very low amounts of free light chain. As could be seen in the performed reference fermentation with GFP, these mutated strains may as well be more stable production hosts than the parental strains. To prove this hypothesis, a continuous fermentation with the found mutants should be carried out confirm this theory. Furthermore, this novel method for screening and strain generation under production conditions could be generally applicable. Using high throughput fermentation combined with product formation triggered cellular stress allows for generation and identification of mutant strains with beneficial production characteristics. The mutations identified in this experiment might be characteristic for the recombinant protein produced. If this is the case, production strains for difficult to produce proteins, could be subjected to directed evolution experiments to obtain mutant strains with improved performance. The hypothesis behind this is that every recombinant protein stresses the host cell in a different way and through long term exposure to these specific stressors, the cell eventually mutates in a way to overcome the stress. As growth behaviour is the only selection criteria in this system, mutations which lead to reduced product formation and therefore reduced cellular stress are often selected. This makes this selection system especially interesting for the generation of strains for continuous fermentation, where genetic stability of the production strain is a prerequisite for the process and can be traded for slightly lower product formation rates.

6.3 Trial of 1lacO-A1 promotor system in HMS174^{Q}

Despite many different approaches and trials, the 1lacO-A1 promotor system could not be brought to function in HMS174^{Q}. The suspected cause of this is a single amino acid difference between BL21 and HMS174 in the *rpoD* gene. As this gene codes for the sigma 70 factor, it is of central importance for the cellular transcription machinery during exponential growth (Jishage and Ishihama, 1995). The naturally occurring difference between the two strains in the sigma 70 factor occurs at position 571, where BL21 incorporates a histidine and HMS174 a tyrosine. It has been reported, that this Y571H alteration reduces the sigma 70 factor's ability to bind to the RNAP core (Jishage et al., 2002). This leads to a problem when utilizing the 1lacO-A1 promotor system. In this setup, the operator sequence is located between the -35 and -10 elements of the promotor, leading to a competition of lac repressor and sigma 70

factor for the binding site. The hypothesis is, that because of the higher binding affinity of the sigma 70 factor to the RNAP core in HMS174 and therefore higher transcription initiation rates, the lac repressor cannot efficiently repress translation and the result is a highly leaky system. To turn this fragile balance, several different approaches including the incorporation of a variant of the lac repressor with higher binding affinity to the lac operator, as well as multiple approaches to modify the *rpoD* gene have been tested. In the end, only the trial to incorporate the Y571H mutation into HMS174^{Q} was partially successful. Upon sequencing of some of the recombinants, subpopulations with the desired mutation were found. But as it was not possible to isolate them, it is believed, that HMS174^{Q} cannot grow sufficiently with alterations in its sigma 70 factor. To test this hypothesis, the *rpoD* gene could be knocked out and subsequently the altered version would be inserted. It would also be possible to test the effect of additional different mutations on the sigma 70 factor in its active centre. Alternatively, it would also be a possibility to test different host RNAP-based promotor systems in HMS174, as it is not possible to incorporate the 1lacO-A1 promotor system into the strain without substantially altering the host cells. This could in turn lead to unexpected effects on the strain and hinder the comparability to other production hosts carrying the same construct.

7 References

BRIAND, L. et al. (2016): A self-inducible heterologous protein expression system in Escherichia coli. Scientific Reports 6: 33037.

CODY, C. W. et al. (1993): Chemical structure of the hexapeptide chromophore of the Aequorea green-fluorescent protein. Biochemistry 32(5): 1212-8.

CORMACK, B. P.; VALDIVIA, R. H.; FALKOW, S. (1996): FACS-optimized mutants of the green fluorescent protein (GFP). Gene 173: 33-38.

GATTI-LAFRANCONI, P. et al. (2013): A single mutation in the core domain of the lac repressor reduces leakiness. Microbial Cell Factories 12: 67.

GLASCOCK, C.; WEICKERT, M. (1998): Using chromosomal lacIQ1 to control expression of genes on high-copy-number plasmids in Escherichia coli. Gene 223: 221-31.

GRAUMANN, K.; PREMSTALLER, A. (2006): Manufacturing of recombinant therapeutic proteins in microbial systems. Biotechnology Journal 1: 164-186.

ITO, K.; INABA, K. (2008): The disulfide bond formation (Dsb) system. Current Opinion in Structural Biology 18: 450-458.

JISHAGE, M.; ISHIHAMA, A. (1995): Regulation of RNA polymerase sigma subunit synthesis in Escherichia coli: intracellular levels of sigma 70 and sigma 38. Journal of bacteriology 177: 6832-6835.

JISHAGE, M. et al. (2002): Regulation of sigma factor competition by the alarmone ppGpp. Genes & development 16: 1260-1270.

KASLI, I. M.; THOMAS, O. R. T.; OVERTON, T. W. (2019): Use of a design of experiments approach to optimise production of a recombinant antibody fragment in the periplasm of Escherichia coli: selection of signal peptide and optimal growth conditions. AMB Express 9: 5.

KENSY, F. et al. (2009): Validation of a high-throughput fermentation system based on online monitoring of biomass and fluorescence in continuously shaken microtiter plates. Microbial Cell Factories 8: 31.

LAGASSÉ, H. A. D. et al. (2017): Recent advances in (therapeutic protein) drug development. F1000Research 6: 113-113.

LEE, H. et al. (2012): Rate and molecular spectrum of spontaneous mutations in the bacterium Escherichia coli as determined by whole-genome sequencing. Proceedings of the National Academy of Sciences 109: E2774.

LIM, H.-K.; JUNG, K.-H. (1998): Improvement of Heterologous Protein Productivity by Controlling Postinduction Specific Growth Rate in Recombinant Escherichia coli under Control of the PL Promoter. Biotechnology Progress 14: 548-553.

LODISH , H. et al. (2000): Molecular Cell Biology. New York: W. H. Freeman.

BioLector, https://www.m2p-labs.com/bioreactors/products/biolector/#!/5 (21.06.2020)

Flower Plates, https://www.m2p-labs.com/bioreactors/microtiter-plates/flowerplate/ (21.06.2020)

MAHARJAN, R. P.; FERENCI, T. (2017): A shifting mutational landscape in 6 nutritional states: Stress-induced mutagenesis as a series of distinct stress input–mutation output relationships. PLOS Biology 15: e2001477.

MAHARJAN, R. P.; FERENCI, T. (2018a): Escherichia coli mutation rates and spectra with combinations of environmental limitations. Microbiology 164: 1491-1502.

MAHARJAN, R. P.; FERENCI, T. (2018b): The impact of growth rate and environmental factors on mutation rates and spectra in Escherichia coli. Environmental Microbiology Reports 10: 626-633.
MARISCH, K. et al. (2013a): Evaluation of three industrial Escherichia coli strains in fed-batch cultivations during high-level SOD protein production. Microbial Cell Factories 12: 58.
MARISCH, K. et al. (2013b): A comparative analysis of industrial Escherichia coli K-12 and B strains in high-glucose batch cultivations on process-, transcriptome- and proteome level. PloS one 8: e70516-e70516.
MATTEAU, D. et al. (2015): A Small-Volume, Low-Cost, and Versatile Continuous Culture Device. PLOS ONE 10: e0133384.
MONK, J. M. et al. (2016): Multi-omics Quantification of Species Variation of Escherichia coli Links Molecular Features with Strain Phenotypes. Cell systems 3: 238-251.e12.
NAVILLE, M. et al. (2011): ARNold: A web tool for the prediction of Rho-independent transcription terminators. RNA Biology 8: 11-13.
NELSON, A. L. (2010): Antibody fragments: hope and hype. mAbs 2: 77-83.
NELSON, A. L.; REICHERT, J. M. (2009): Development trends for therapeutic antibody fragments. Nature Biotechnology 27: 331-337.
PEEBO, K.; NEUBAUER, P. (2018): Application of Continuous Culture Methods to Recombinant Protein Production in Microorganisms. Microorganisms 6: 56.
PEUBEZ, I. et al. (2010): Antibiotic-free selection in E. coli: new considerations for optimal design and improved production. Microbial cell factories 9: 65-65.
QUICK, M.; WRIGHT, E. M. (2002): Employing Escherichia coli to functionally express, purify, and characterize a human transporter. Proceedings of the National Academy of Sciences of the United States of America 99: 8597-8601.
ROSANO, G. L.; CECCARELLI, E. A. (2014): Recombinant protein expression in Escherichia coli: advances and challenges. Frontiers in Microbiology 5.
RUGBJERG, P.; SOMMER, M. O. A. (2019): Overcoming genetic heterogeneity in industrial fermentations. Nature Biotechnology 37: 869-876.
SAIDA, F. et al. (2006): Expression of Highly Toxic Genes in E. coli: Special Strategies and Genetic Tools. Current Protein and Peptide Science 7: 47-56.
SANTINI, C.-L. et al. (2001): Translocation of Jellyfish Green Fluorescent Protein via the Tat System of Escherichia coli and Change of Its Periplasmic Localization in Response to Osmotic Up-shock. Journal of Biological Chemistry 276: 8159-8164.
SCHIMEK, C. et al. (2020): Extraction of recombinant periplasmic proteins under industrially relevant process conditions: Selectivity and yield strongly depend on protein titer and methodology. Biotechnology Progress n/a: e2999.
SCHMIDEDER, A.; WEUSTER-BOTZ, D. (2017): High-performance recombinant protein production with Escherichia coli in continuously operated cascades of stirred-tank reactors. Journal of Industrial Microbiology & Biotechnology 44: 1021-1029.
SCHULLER, A. et al. (2020): Escherichia coli σ70 promoters allow expression rate control at the cellular level in genome-integrated expression systems. Microbial Cell Factories 19: 58.
SHARAN, S. K. et al. (2009): Recombineering: a homologous recombination-based method of genetic engineering. Nature Protocols 4: 206.
SHIMOMURA, O. (2005): The discovery of aequorin and green fluorescent protein. Journal of Microscopy 217: 3-15.
SIEBEN, M. et al. (2016): Testing plasmid stability of Escherichia coli using the Continuously Operated Shaken BIOreactor System. Biotechnology Progress 32: 1418-1425.

STRIEDNER, G. et al. (2003): Tuning the Transcription Rate of Recombinant Protein in Strong Escherichiacoli Expression Systems through Repressor Titration. Biotechnology Progress 19: 1427-1432.

STRIEDNER, G. et al. (2010): Plasmid-free T7-based Escherichia coli expression systems. Biotechnology and Bioengineering 105: 786-794.

STUDIER, F. W. et al. (2009): Understanding the differences between genome sequences of Escherichia coli B strains REL606 and BL21(DE3) and comparison of the E. coli B and K-12 genomes. Journal of molecular biology 394(4): 644–652.

TERPE, K. (2006): Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Applied Microbiology and Biotechnology 72: 211.

TOEROEK, C. et al. (2015): Fed-batch like cultivation in a micro-bioreactor: screening conditions relevant for Escherichia coli based production processes. SpringerPlus 4: 490.

TOMATIS, P. E. et al. (2019): Mutations in sigma 70 transcription factor improves expression of functional eukaryotic membrane proteins in Escherichia coli. Scientific Reports 9: 2483.

TSIEN, R. Y. (1998): The green fluorescent protein. Annual Review of Biochemistry 67: 509-44.

TSIRIGOTAKI, A. et al. (2017): Protein export through the bacterial Sec pathway. Nature Reviews 15.

UNIPROTKB (1986): UniProtKB - P00579 (RPOD_ECOLI). July 21, 1986 ed.

UNIPROTKB (1995): UniProtKB - P42212 (GFP-AEQVI). Nov. 1, 1995 ed.

VELUR SELVAMANI, R. S.; FRIEHS, K.; FLASCHEL, E. (2014): Extracellular recombinant protein production under continuous culture conditions with Escherichia coli using an alternative plasmid selection mechanism. Bioprocess and Biosystems Engineering 37: 401-413.

VETHANAYAGAM, J. G. G.; FLOWER, A. M. (2005): Decreased gene expression from T7 promoters may be due to impaired production of active T7 RNA polymerase. Microbial Cell Factories 4: 3.

WAEGEMAN, H.; SOETAERT, W. (2011): Increasing recombinant protein production in Escherichia coli through metabolic and genetic engineering. Journal of Industrial Microbiology & Biotechnology 38: 1891-1910.

WALSH, C. (2000): Molecular mechanisms that confer antibacterial drug resistance. Nature 406: 775.

WALTHER, J. et al. (2015): The business impact of an integrated continuous biomanufacturing platform for recombinant protein production. Journal of Biotechnology 213: 3-12.

WARSI, O. M.; DYKHUIZEN, D. E. (2017): Evolutionary implications of Liebig's law of the minimum: Selection under low concentrations of two nonsubstitutable nutrients. Ecology and Evolution 7: 5296-5309.

YAO, T.; ASAYAMA, Y. (2017): Animal-cell culture media: History, characteristics, and current issues. Reproductive medicine and biology 16: 99-117.

8 Abbreviations

1lacO-A1 ..... *one lac operator - A1 promotor system*
CDM ..... *Cell Dry Mass*
CSTR ..... *Continuously Stirred Tank Reactor*
dH_2O ..... *deionized water*
DO ..... *dissolved oxygen*
E. coli ..... *Escherichia coli*
EDTA ..... ethylenediaminetetraacetic acid
ELISA ..... *Enzyme Linked Immunosorbent Assay*
Fab ..... *fragment antigen binding*
FACS ..... *Fluorescence Activated Cell Sorting*
GFP ..... *Green Fluourescent Protein*
IPTG ..... Isopropyl-β-D-thiogalactopyranoside
mAB ..... *monoclonal antibody*
mut. ..... *mutant*
OD_{600} ..... *optical density at 600 nm*
RNAP ..... *RNA polymerase*
rpm ..... *revolutions per minute*

9 List of tables

10 Table of figures

11 Appendix

11.1 Genome integration cassettes

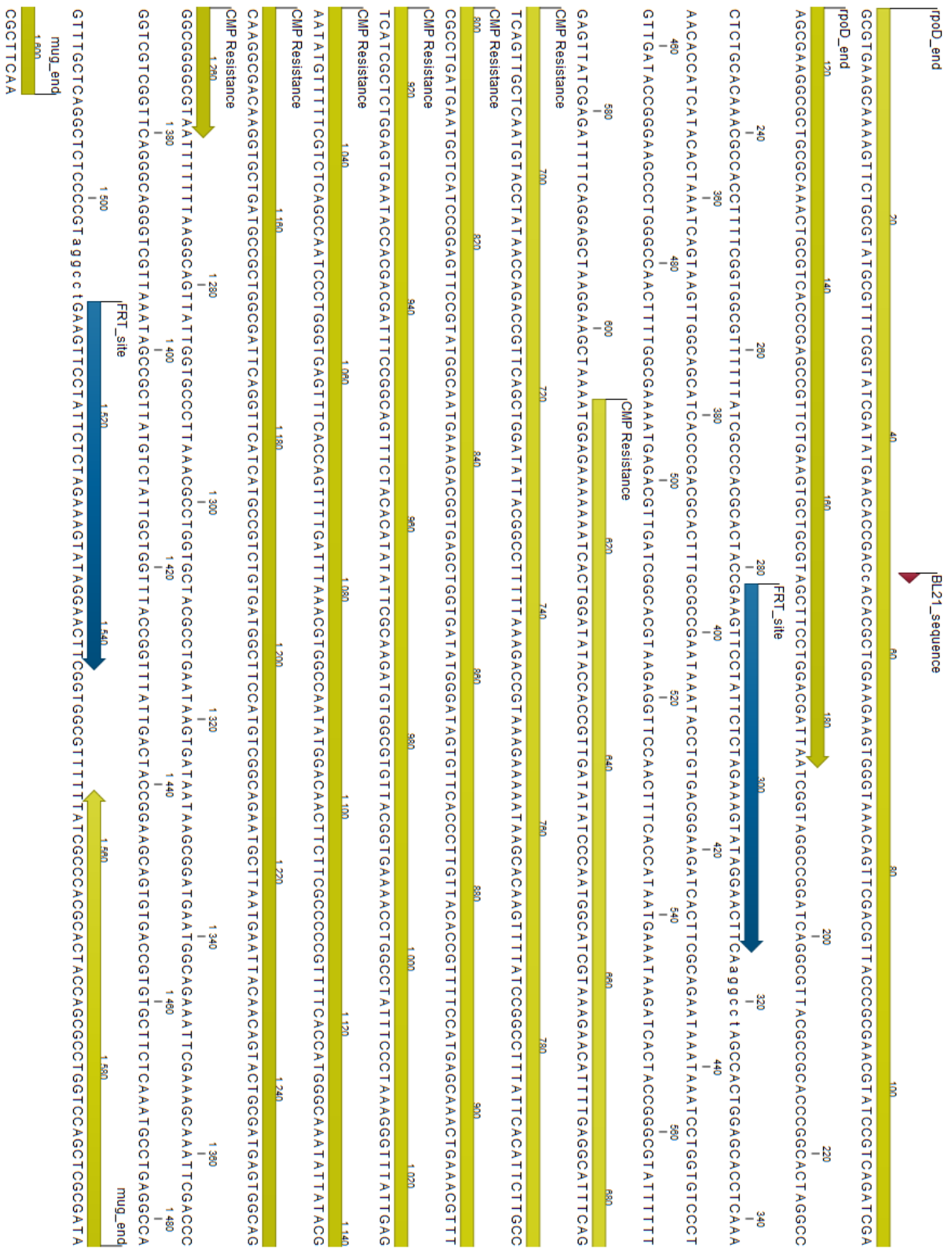

Figure 22: Genome integration cassette rpoD_opt

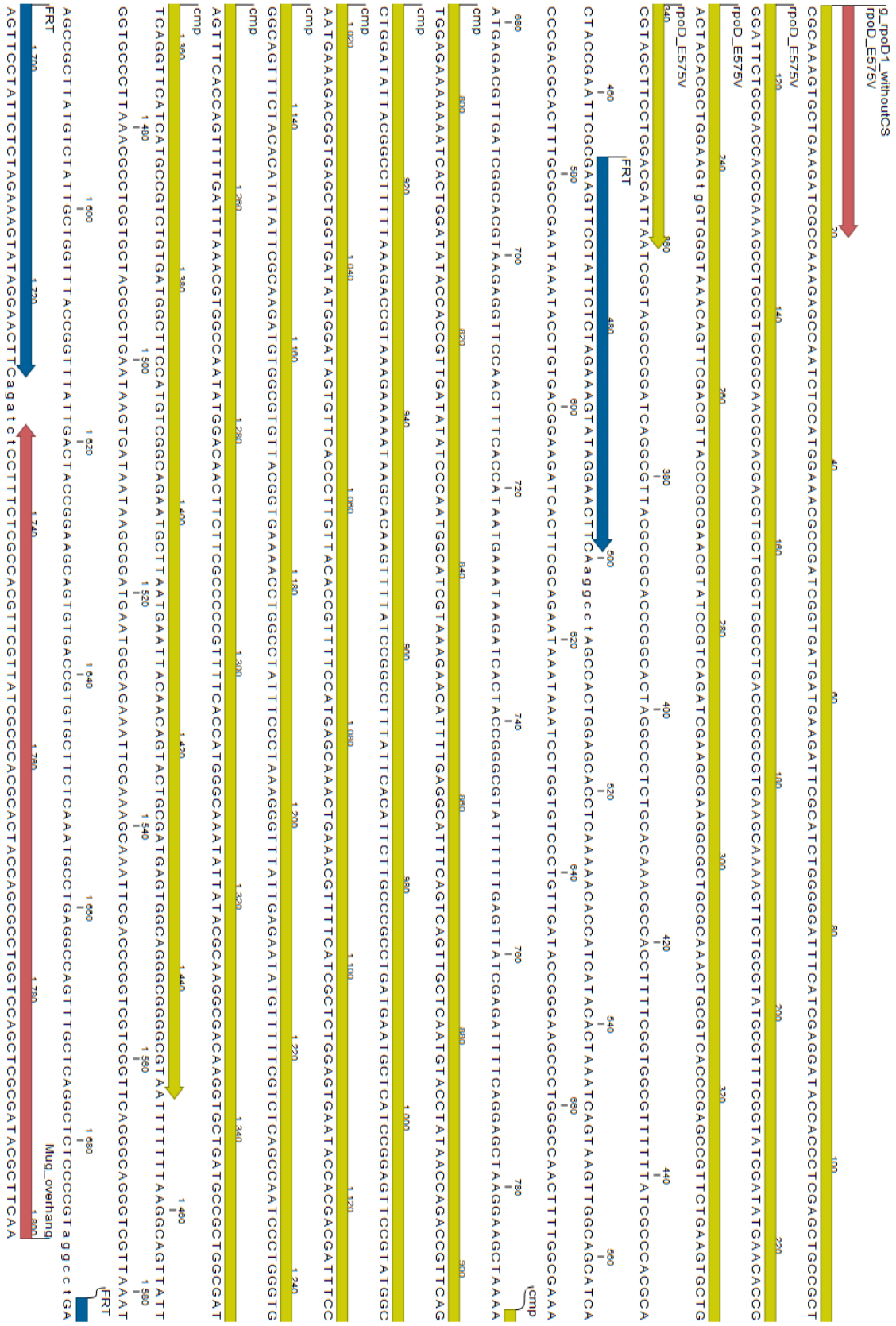

Figure 23: Genome integration cassette rpoD_E575V

Figure 24: Sequence of the 1lacO-A1 genome integration cassette with GFP (for full length overhang primers see Table 15)

11.2 Gene maps

Figure 25: Alignment of rpoD in BL21, HMS174 and RV308

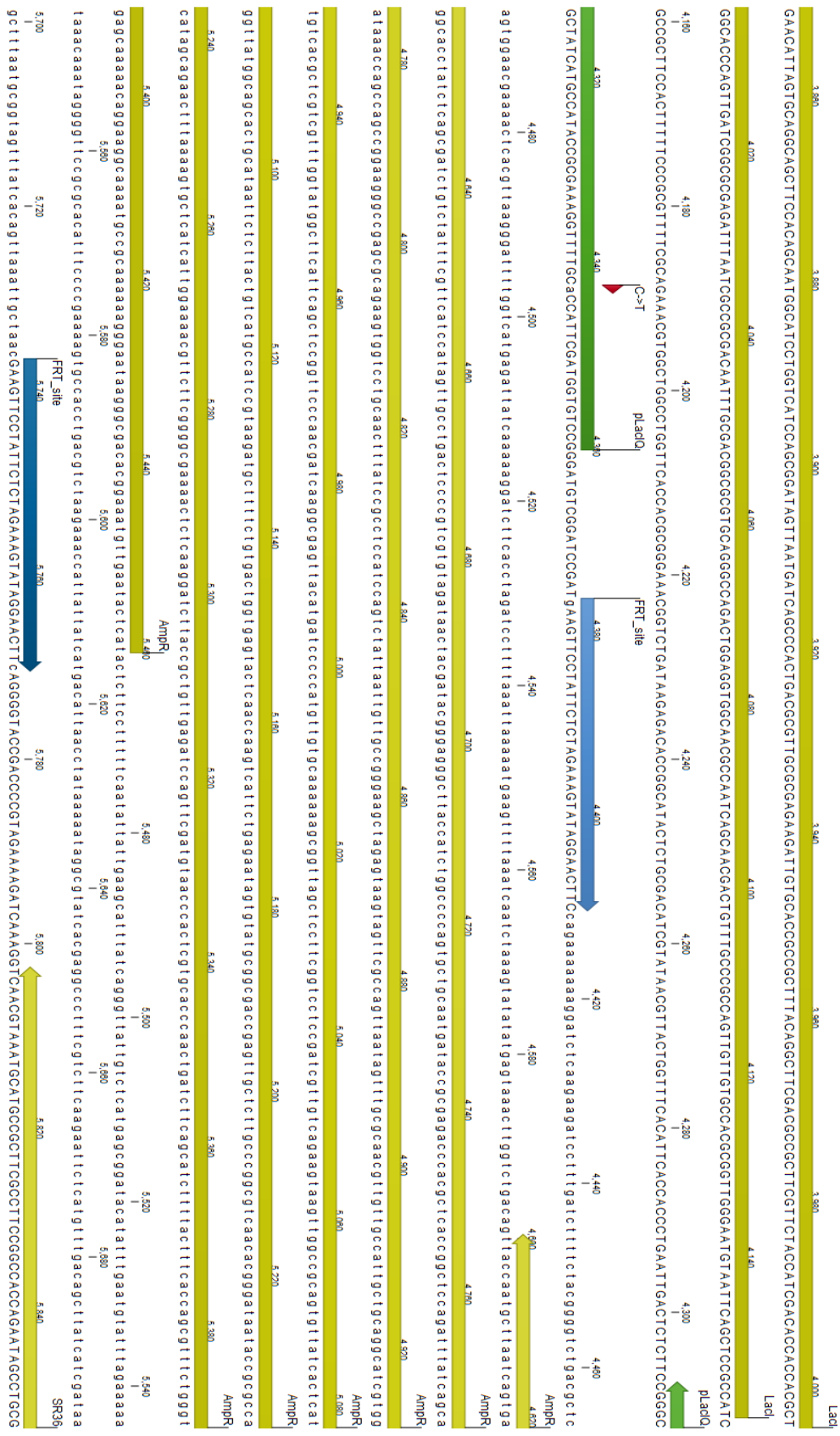

Figure 26: Sequence of the lacI^{Q} promotor at the integrated site (changed bases indicated in small letters)

11.3 Plasmid maps

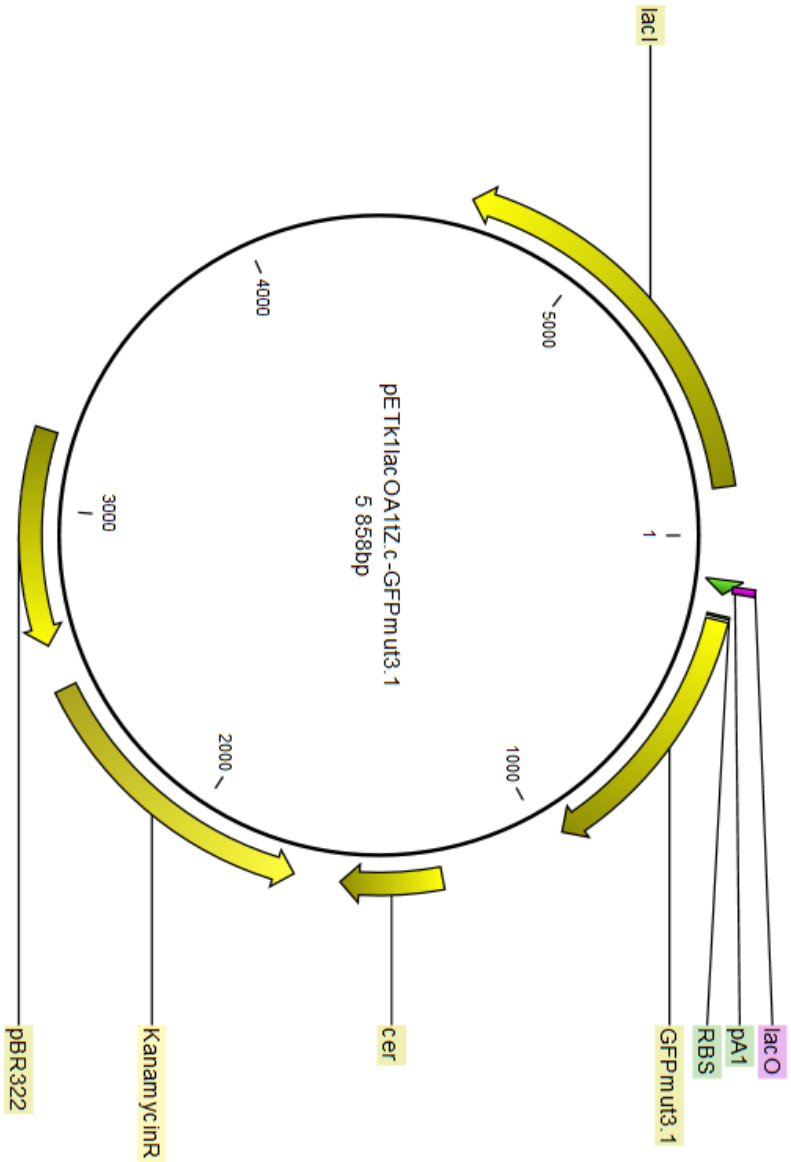

Figure 27: Structure of pETk1lacOA1tZ.c-GFPmut3.1

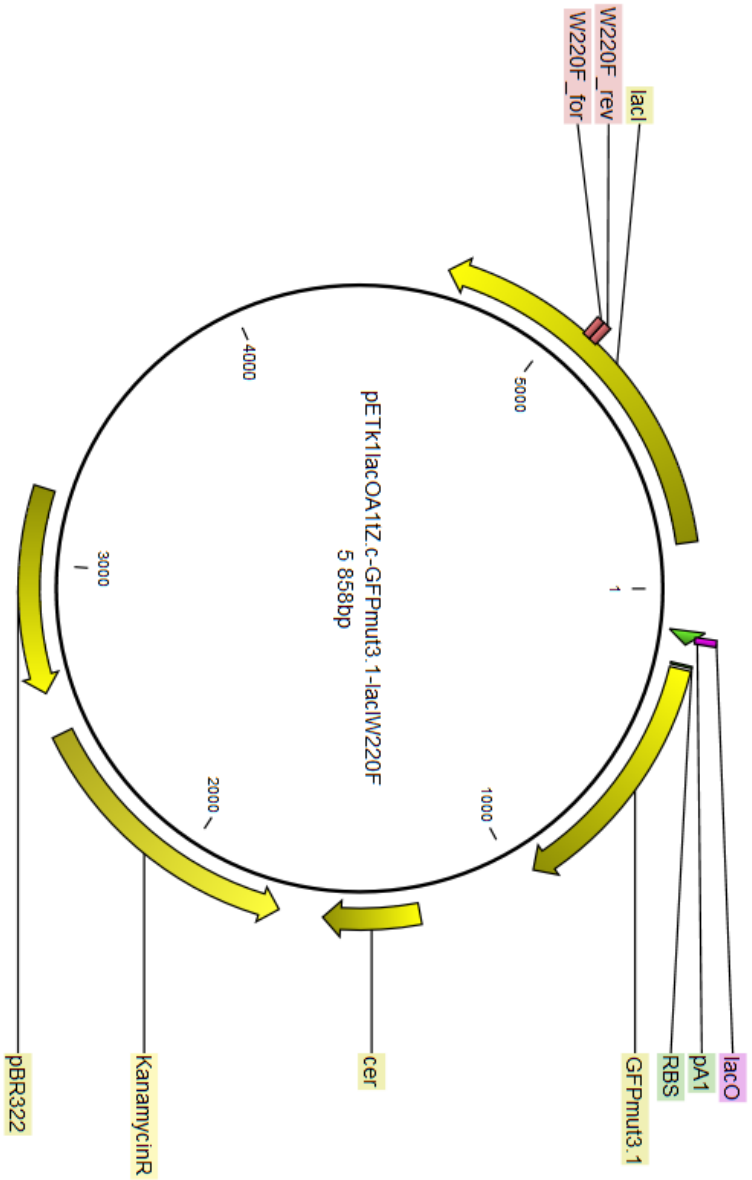

Figure 28:Structure of pETk1lacOA1tZ.c-GFPmut3.1-lacIW220F

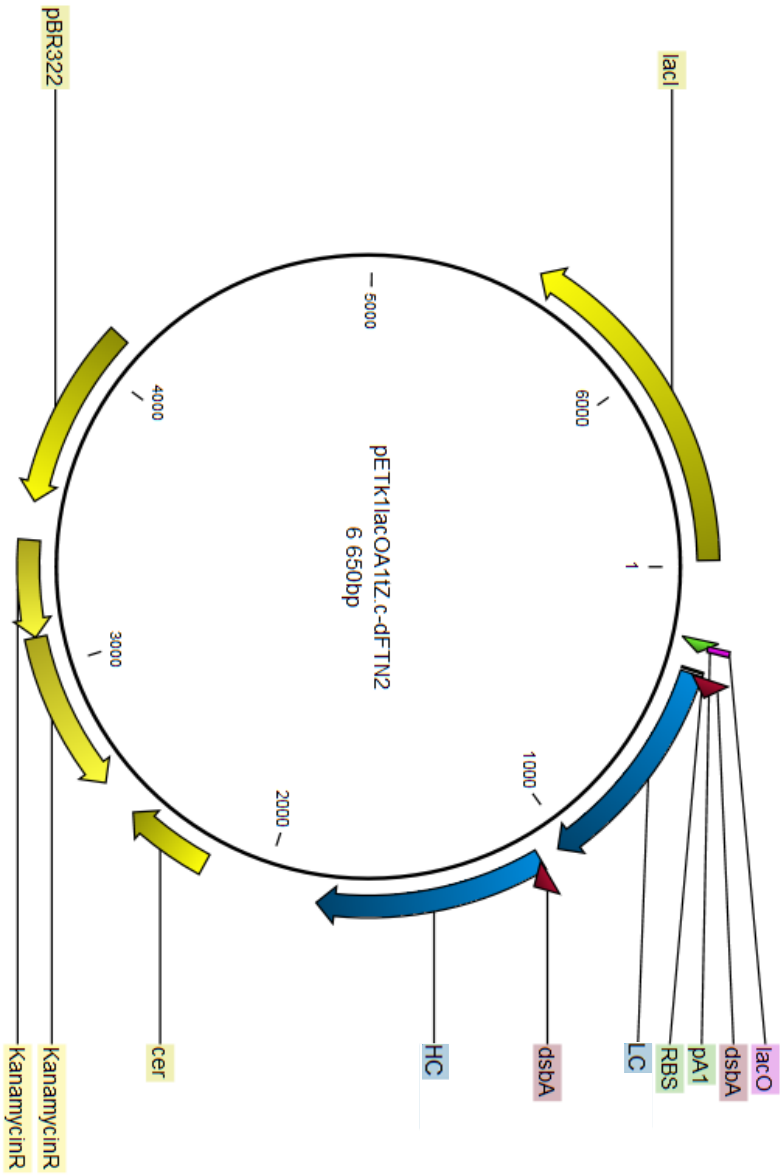

Figure 29: Structure of pETk11lacOA1tZ.c-dFTN2

11.4 Primer

Table 15: Sequence and purpose of primers

Primer name	Sequence (5' – 3')	Purpose
W220F_for	GAACGGGAAGGCGACTTCAGTG	Insertion of W220F into lacI
W220F_rev	CGCTATCGGCTGAATTTGATTGCG	Insertion of W220F into lacI
TN7_1_pET30aw/oKanR_for	AGATGACGGTTTGTCACATGGAGT TGGCAGGATGTTTGATTAAAAACATA GTAGTAGGTTGAGGCCGTTG	Overhang primer for generation of genome integration cassette <1lacO-A1-GFP-t.Z>
TN7_2_pET30a_for	CAGCCGCGTAACCTGGCAAAATCGG TTACGGTTGAGTAATAAATGGATGC GAAGATCCTTTGATCTTTTCTACG	Overhang primer for generation of genome integration cassette <1lacO-A1-GFP-t.Z>
rpoD_ext1	GCGTGAAGCAAAAGTTCTGC	Amplification of rpoD_opt genome integration cassette
rpoD_ext2	AACTTCGCATAGCGCTCGAC	Amplification of rpoD_opt genome integration cassette
rpoD_seq1	GATTCTGCGACCACCGAAAG	Screening for rpoD_opt recombinants
rpoD_seq2	CAAATTCAGCGCAAAGTGAC	Screening for rpoD_opt recombinants
rpoD_check	GGTGATGCTGCCAACTTAC	sequencing of gBlock
rpoD_overhang	GGGATTTCATCGAGGATACCACCCTCGAGC TGCCGCTGGATTCTGCGACCACCGAAAGCC TGCGTGCGGCAACGCACGACGTGCTGGCT GGCCTGACCGCGCGTGAAGCAAAAGTTCTGC	amplification of rpoD_opt with overhang further into the rpoD gene
proc1	TACCG**AATT**CGCGAAGTTCCTATTCTCT	Construction of rpoD_E575V integration cassette (plasmid amplification), cutting site bold
proc2	GCGG**CATG**CATTTGACAGGCACATTATGC	Construction of rpoD_E575V integration cassette (plasmid amplification), cutting site bold
g_rpoD1	AATG**CATG**CCGCAAAGTGCTGAAGATCGC	rpoD gene, cutting site bold
g_rpoD2	CGCG**AATT**CGGTAGTGCGTGGGCGATAA	rpoD gene, cutting site bold
E575V_for	GCTGGAAGtgGTGGGTAAAC	Intoduction of E575V mutation
E575V_ref	GTGTAGTCGGTGTTCATATCG	Intoduction of E575V mutation
rpoD_seq3	GAAGAACTGGCTGAACGTATGC	Sequencing of rpoD genome integration site
g_rpoD1_withoutCS	CGCAAAGTGCTGAAGATCGC	Construction of rpoD_E575V integration cassette
Mug_overhang	TTGAAGCGTATCGCGAGCTGGACCAGGCGC TGGTAGTGCGTGGGCGATAACGAACGTGGC GAGAAAGGAG	Construction of rpoD_E575V integration cassette

11.5 BioLector® medium

Table 16: Components of BioLector medium

Salt MIX

Calculated Weight [g]	Chemicals	Real Weight	Company	Charge	Lotnumber
43,65	MOPS		Sigma		121K5405
10,42	$(NH_4)_2SO_4$		Roth	315232553	
3,13	K_2HPO_4		Roth	264215281	
3,13	$Na_3Citrate*2H_2O$		Roth	236235851	
2,08	Na_2SO_4		Merck	1.06649.1000	A0252349 125
1,04	NH_4Cl		Applichem	A0984.1000	0G008925
				pH auf 7.4 mit $NaOH/H_2SO_4$	☑
				deionized water fill up to [ml]: 500	☑
				sterile filtered	☑

Trace Element Solution

Calculated Weight [g]	Chemicals	Real Weight	Company	Charge	Lotnumber
0,011	$ZnSO_4*7H_2O$		Roth	-	
0,010	$CuSO_4*5H_2O$		Merck	1.02787.1000	A110987846
0,006	$MnSO_4*H_2O$		Merck	1.05941.0250	F1209241
0,835	$FeCl_3*6H_2O$		Roth	421166191	
0,668	Titriplex III		Merck	1.08421.1000	8421B046912
0,011	$CoCl_2*6H_2O$		Fluka	326994/1 1294	
0,040	$CaCl_2*2H_2O$		Applichem	A0775.5000	0L008540
				fill up to 20mL with deionized water	☑
				sterile filtered	☑

Thiamin Solution/Vitamin Solution

Calculated Weight [g]	Chemicals	Real Weight	Company	Charge	Lotnumber
0,02	Thiamin*HCl		Calbiochem		B73434
				fill up to 20mL with deionized water	☑
				sterile filtered	☑

Magnesium Solution

Calculated Weight [g]	Chemicals	Real Weight	Company	Charge	Lotnumber
1,00	$MgSO_4*7H_2O$		Roth	304217199	
				fill up to 20mL with deionized water	☑
				sterile filtered	☑

Table 17: Preparation of final BioLector medium

2xPS Final Mix of Stock Solutions	
Volume [mL]	Chemicals
225.0	Polysaccharide solution
180.0	Salt Mix
3.8	Magnesium solution
3.8	Thiamin Solution/Vitamin Solution
0.4	Trace elements solution
93.4	Deionized water

11.6 Media for continuous fermentations

11.6.1 Batch medium

Table 18: Batch medium preparation

Batch medium	R1 - A1 ref		R2 - T7 ref		R3 - T7 mut					
	Calculated Weight	Real Weight	Calculated Weight	Real Weight	Calculated Weight	Real Weight	company	Charge	Calculated Weight	Real Weight
Kaliumdihydrogenphosphat [g]	1.98		1.98		1.98		ROTH	437264392	6.13	6.14
85% H_3PO_4-Sol.[g]	0.67		0.67		0.67		ROTH	486249251	2.07	2.06
Yeast Extract [g]	0.60		0.60		0.60		Merck	UM679126504	1.86	1.86
Na-Citratdihydrat [g]	0.86		0.86		0.86		ROTH	448276477	2.68	2.67
Mg-Chlorid* 6 H_2O [g]	0.18		0.18		0.18		ROTH	438270383	0.57	0.58
Ca-Chlorid* 2 H_2O [g]	0.08		0.08		0.08		Merck	K93147600225	0.25	0.25
Trace Element Solution [ml]	0.20		0.20		0.20		-	14.3.19 JLO	0.62	0.62
Ammonsulfat [g]	0.95		0.95		0.95		ROTH	17249132	2.95	2.96
fill up to with dH_2O [g]	300	300.14	300	300.02	300	300.55				
Glucose-Monohydrat [g]	13.20	13.20	13.20	13.20	13.20	13.21	ROTH	149282386		
fill up to with dH_2O [g]	100.00	100.06	100.00	100.03	100.00	100.24				

11.6.2 Fed-batch medium

Table 19: Fed-batch medium preparation

Fed-Batch medium	R1 - A1 ref		R2 - T7 ref		R3 - T7 mut		spare medium					
	Calculated Weight	Real Weight	Calculated Weight	Real Weight	Calculated Weight	Real Weight	Calculated Weight	Real Weight	Company	Charge	Calculated Weight	Real Weight
Mg-Chlorid* 6 H_2O [g]	0.80		0.80		0.80		0.80		ROTH	438270383	3.27	3.27
Ca-Chlorid* 2 H_2O [g]	0.35		0.35		0.35		0.35		Merck	K93147600225	1.44	1.43
Trace Element Solution [ml]	0.87		0.87		0.87		0.87		-	14.3.19 JLO	3.55	3.55
solve salts in dH_2O [g]	115.00	115.09	115.00	115.10	115.00	115	115.00	115.02			471.50	471.72
solve glucose in dH_2O [g]	200.00	200.10	200.00	200.01	200.00	200.57	200.00	200.02				
Glucose-Monohydrat [g]	57.22	57.23	57.22	57.23	57.22	57.24	57.22	57.21	ROTH	149282386		

Table 20: Continuous medium preparation

Continuous medium	Components				1	2	3
	Calculated Weight (1 L)	Company	Charge	Calculated Weight (20 L)	Real Weight	Real Weight	Real Weight
Kaliumdihydrogenphosphat [g]	2.85	ROTH	437264392	57.05	57.06	57.05	57.05
Di-Kaliumhydrogenphosphat [g]	4.35	ROTH	264215281	87.09	87.08	87.09	87.09
Na-Citratdihydrat [g]	7.50	ROTH	89276477	150.00	150.02	150.01	150.02
Mg-Sulfat* 7 H2O [g]	3.00	ROTH	304217199	60.00	60.02	60.02	60.01
Ca-Chlorid* 2 H2O [g]	0.60	Merck	K93147600225	12.00	12.01	11.99	11.99
Trace Element Solution [mL]	1.50	-	14.3.19 JLO	30.00	30.00	30.00	30.00
Glucose-Monohydrat [g]	99.00	ROTH	149282386	1980.00	1980.04	1980.08	1980.01
fill up to with dH2O [g]	1000			20000	20000	20000	20000
IPTG	0.14			2.86	2.87	2.87	2.86

11.6.4 Trace element solution

Table 21: Composition of trace element solution

Sequence	Substance	Amount	
1	$FeSO_4*7H_2O$	40.0	g/L
2	$MnSO_4*H_2O$	10.0	g/L
3	$AlCl_3*6H_2O$	10.0	g/L
4	$CoCl_2*6H_2O$	7.3	g/L
5	$ZnSO_4*7H_2O$	2.0	g/L
6	$Na_2MoO_4*2H_2O$	2.0	g/L
7	$CuCl_2*2H_2O$	1.0	g/L
8	H_3BO_3	0.5	g/L

11.7 List of additional materials

Table 22: List of additional materials (as not already specified in section Materials and methods

Orbital shaker	Inforce HT multitron
Heated orbital shaker	Sartorius, Certomat® IS
Electroporator	BTX Harvard Apparatus: ECM630 Electro Cell Manipulator
Electroporation-stand	BTX Harvard Apparatus: BTX safety stand 630B
Laminar flow cabinet	Holton LaminAir HBB2448
PCR Thermocycler	peQLab peQSTAR
Spectrometer (DNA)	Nanodrop ND-1000 Spectrophotometer
Nanodrop software	ND-1000 V3.8.1
Centrifuge	Eppendorf MiniSpin rotor: F45-12-11
Centrifuge	Eppendorf Centrifuge 5415R, rotor: F45-24-11
Centrifuge	Eppendorf Centrifuge 5804R, rotor: F34-6-38 / A-2-DWP
Incubator (37 °C/30 °C)	Heraeus Instruments B6120
Laboratory scale	Precisa XB4200C
Thermomix	Eppendorf Thermomix Comfort
Microscope	Olympus Model CHT, CH-2
Spectrometer (OD_{600})	Amersham Biosciences, Ultraspec 500 pro
Spectrometer cuvettes	Brand, PMMA 1.5 mL Cat. No. 7591 15
Electroporation cuvettes	Biozym scientific GmbH, 1mm Polycarbonate Art. No. 748010
Piston-stroke pipette	Gilson Pipetman 1000 µL, X67941D
Piston-stroke pipette	Gilson Pipetman 200 µL, W59835G
Piston-stroke pipette	Gilson Pipetman 100 µL, Y62055K
Piston-stroke pipette	Gilson Pipetman 20 µL, T63281H
Piston-stroke pipette	Gilson Pipetman 10 µL, CG66690

Multichannel pipette	Finnpipette™ F2 30-300 µL, Thermo Scientific
Pipette tip 1200 µL	Gilson, D1200 Diamond® ECO·PACK™
Pipette tip 200 µL	Gilson, D200 Diamond® ECO·PACK™
Pipette tip 10 µL	Gilson, D10 Diamond® ECO·PACK™
Culture dishes	Gosselin™ petri dishes
Electrophoresis power supply	Bio-Rad, PowerPac™ Basic Power Supply
Agarose electrophoresis chamber	Bio-Rad, Wide Mini-Sub Cell GT
UV imaging system	Bio-Rad, ChemiDoc™ XRS+
Water-bath/Shaker	GFL, Typ 1o83, No.: 1c9219899H
Autoclave	Certoclav-Tisch-Autoclav CV-EL 12L/18L
Autoclave	Varioklav, H+P Labortechnik GmbH, Type EH
Precision scale	Mettler, Type PM4600
Magnetic stirrer	Ikamag RCT
Agarose electrophoresis power supply	PowerPac™ Basic Power Supply, Bio-Rad
Agarose electrophoresis chamber	Sub-Cell GT Cell, Bio-Rad
SDS Page power supply	EPS 301 POWER SUPPLY,GE Healthcare
SDS Page chamber	Invitrogen™ Novex™ XCellSureLock™