Unterschrift des Betreuers

TECHNISCHE UNIVERSITÄT WIEN

DIPLOMARBEIT

Expression tuning in *Escherichia coli* via a glucose/lactose mixed feed system

ausgeführt am Institut für
Verfahrenstechnik, Umwelttechnik und technische Biowissenschaften
der Technischen Universität Wien

unter der Anleitung von Dipl.-Ing. Dr.nat.techn. Oliver Spadiut und
Dipl.-Ing. David Johannes Wurm

durch

Julian Quehenberger

Donaufelder Straße 101/6/9

1210 Wien

Wien, 1. Februar 2016

Julian Quehenberger

Abstract

Short production phases due to high metabolic burden and formation of insoluble inclusion bodies are known limitations for recombinant protein production in *Escherichia coli*. Up to now, these issues are mainly tackled by genetic engineering.

In this work the following issues are studied: (1) the transferability of the relationship between the specific uptake rates of glucose and lactose during their concomitant uptake is investigated. Therefore specific substrate uptake rates of two *E. coli* BL21(DE3) strains producing different target proteins under the control of the T7*lac* promotor system are compared during limiting glucose concentrations, while lactose was available in excess; (2) the tunability of the recombinant protein production rate and the ratio of soluble protein (SP) to inclusion bodies (IBs) are assessed; (3) the degree of induction during lactose feeding is investigated on the cellular level.

The results of this Thesis show that the maximum specific lactose uptake rate is strongly influenced by the expressed target protein. Furthermore it can be concluded that lactose induction via a limiting mixed feeding strategy can increase the recombinant protein production rate and the ratio of SP to IBs while differently induced subpopulation can be avoided. This meets the demands of pharmaceutical bioprocesses to efficiently yield high amounts of active target protein and to extend the production phase by reducing metabolic burden and enhancing cell fitness.

Zusammenfassung

Kurze Produktionsphase aufgrund von hoher Stoffwechselbelastung und Bildung von unlöslichen Einschlusskörperchen sind bekannte Einschränkungen für die rekombinante Proteinproduktion in *Escherichia coli*. Zum heutigen Zeitpunkt wird die Lösung dieser Problemen vor allem mittels Methoden der Gentechnik in Angriff genommen.

In der vorliegenden Arbeit werden die folgenden Fragestellungen behandelt: (1) die Übertragbarkeit der Beziehung zwischen der Glucose- und Lactoseaufnahme bei Glucoselimitierung und gleichzeitigem Lactoseüberschuss wird geprüft. Dazu werden die spezifischen Substrataufnahmeraten zweier *E. coli* Stämme BL21(DE3), welche unterschiedliche Zielproteine unter der Kontrolle des T7*lac* Promoters exprimieren, verglichen; (2) die Regelung der Rate der rekombinanten Proteinproduktion und das Verhältnis von gelöstem Protein zu Protein in Einschlusskörperchen wird untersucht; (3) die Stärke der Induktion bei Lactosefütterung wird auf zellulärer Ebene untersucht.

Die Ergebnisse dieser Arbeit zeigen, dass die maximale spezifische Lactoseaufnahmerate stark von der Expression des Zielproteins beeinflusst wird. Weiters kann aus den Resultaten geschlossen werden, dass eine Induktion mit Laktose durch eine limitierende Fütterungsstrategie mit Glucose und Lactose die Rate der rekombinanten Proteinproduktion und das Verhältnis von gelöstem Protein zu Protein in Einschlusskörperchen steigern kann. Gleichzeitig ist es damit möglich die Ausbildung von unterschiedlich stark induzierten Subpopulationen zu vermeiden. Diese Ergebnisse werden den Forderungen der pharmazeutischen Industrie gerecht, effizient hohe Erträge an aktivem Zielprotein zu erreichen und dabei potentiell die Produktionsphase durch Reduktion der Stoffwechselbelastung der produzierenden Zellen auszuweiten.

Danksagung

An dieser Stelle finde ich Platz mich bei allen zu bedanken, die das Entstehen dieser Arbeit ermöglicht haben:

Vielen Dank an Prof. Christoph Herwig für die Aufnahme in den Forschungsbereich BioVT.

An meinen Betreuer Dr. Oliver Spadiut richte ich meinen herzlichen Dank für entscheidenden Input zu Design und Durchführung dieser Arbeit.

Ein ebenso herzliches Dankeschön ergeht an Dipl.-Ing. David Wurm für die hands-on Betreuung im Labor sowie bei der Auswertung und Interpretation der Experimente.

Weiters möchte ich mich bei alle Kollegen im Labor für deren Unterstützung danken.

Zu guter Letzt ergeht der größte Dank an meine Freundin Birgit und meine Familie für Motivation und Unterstützung, die weit über diese Arbeit und das Studium hinausgehen!

Table of contents

Introduction

Benchmarks for the recombinant protein production

The market for recombinant proteins is strongly dominated by products for the medical sector. Human therapeutics and human diagnostics are estimated to account for more than 80% of the US market for recombinant proteins while the rest is distributed between the agricultural and industrial sector and others (Palomares et al. 2002).

Although by 2014 a total of 246 biopharmaceuticals have been approved in the US and EU, more than two thirds of the market are covered by seven proteins: erythropoietin, alpha interferon, hepatitis-B vaccine, granulocyte colony stimulating factor, insulin, human growth hormone and tissue plasminogen activator (Palomares et al. 2002; Maheshwari 2011). The market value of biopharmaceuticals is increasing continuously, reaching a total of $140 billion for the year of 2013 (Walsh 2014), while the 10 top selling drugs generate about 50% of the revenue.

While the biotech industry is growing exponentially, the number of approvals has been fairly constant with about 56 new approvals every 5 years in the US and EU. Recent approvals were mainly monoclonal antibodies, hormones and blood-related proteins. Thereby the main focus for application lies in the field of cancer treatment (Walsh 2014).

Although mammalian expression systems enjoy increasing popularity, in 2010 still 68% (about 18.000 kg) of total pure protein as active pharmaceutical ingredients were produced in microbial systems. Within microbial systems *E. coli* is by far the most common expression host. Regarding bacterial expression systems, only two other species (*Vibrio cholera* and *Bordetella pertussis*), each one only producing a single product, are in use for the production of biopharmaceuticals. Most approvals for biopharmaceuticals in *E. coli* are for insulin, growth factors and interferon (Walsh 2014).

Escherichia coli

Escherichia coli is the most widely used host organism for the production of recombinant proteins (Rosano and Ceccarelli 2014). Since its discovery by the German-Austrian pediatric Theodor Escherich (1886) *E. coli* soon received great

attention as fecal indicator (Smith 1892) and in the early twentieth century started to become the most important model organism for molecular biology (Zimmer 2009). The strain *E. coli* K12 was among the first organisms that had their genome sequenced (Blattner et al. 1997). Today the genetics of *E. coli* are best understood out of all organisms.

The development of tools in *E. coli* for genetic manipulation, namely the cloning of genes via restriction enzymes and introduction via plasmids by Cohen et al. (1973), paved the way for the rise of the biotechnological industry. The first commercial biotechnological product was a human hormone called somatostatin produced in *E. coli*. Soon the production of insulin and growth hormones followed (Russo 2003).

E. coli exhibits extremely fast growth kinetics and can be grown in high density cultures. Cultivation is possible on inexpensive and synthetic media (Lee 1996). The large variety of potent transformation systems and ease of genome modification are further key advantages of *E. coli* as production host.

Disadvantages of *E. coli*, especially for the production of pharmaceutical proteins, are the lack of glycosylation, incomplete posttranslational modifications, like the incorporation of cofactors or formation of disulfide bonds in the cytoplasm, limitations regarding the production of very large proteins, the formation of endotoxins and the tendency to form inactive protein aggregates, so called inclusion bodies (Choi et al. 2006).

Nevertheless, due to its superior cost efficiency, much effort is put in the development of techniques and strategies to improve *E. coli* as host for the production of pharmaceutical proteins.

In practice the recombinant protein production in *E. coli* is controlled by inducible expression systems. For this purpose the pET expression system is the most popular of its kind and therefore this system was chosen to be subject of this Thesis.

The pET expression system

The pET expression system was commercialized by Novagen and is the most popular system for recombinant protein production in research. In 2003 more than 90% of the entries in the protein data base (PDB; Berman et al. 2000) originated from a pET expression system (Sørensen and Mortensen 2005a). The pET system was

first described by (Studier et al. 1990). More than 40 different variants of the pET plasmid are commercially available. The plasmids differ in available fusion tags, reading frames, selectable markers and restriction sites. pET plasmids have to be used in combination with an *E. coli* host that contains the gene encoding for the T7 RNA polymerase. This gene is under the control of the *lacUV5* promoter which is less sensitive for regulation by cAMP-CAP complex compared to the wild type *lac* promoter. Binding of the inducer molecules isopropyl β-D-1-thiogalactopyranoside (IPTG), thiomethyl-β-D-galactoside (TMG) or allolactose, the natural inducer, prevents the tetrameric repressor complex LacI from binding the *lac* operator sequence. *lacI* is present in the genome of BL21(DE3) as well as on the pET plasmid. With LacI no longer blocking the *lacUV5* promoter the gene for the T7 polymerase is transcribed by *E. coli*´s native RNA polymerase. Ultimately T7 polymerase transcribes the target gene on the pET plasmid which is under control of a T7 promoter and a *lac* operator sequence.

Advantages of this system are the high specificity of the T7 promoter and high expression rate of the T7 polymerase. During induction the T7 polymerase outcompetes the *E. coli* RNA polymerase by approximately the factor 5 (230 compared to 50 nucleotides per second) (Sørensen and Mortensen 2005a). Nevertheless expression under the pET system is known to be leaky and basal levels of target protein are produced without addition of inducer. This is because the LacI repressor does not completely shut down the *lacUV5* promoter (and likewise the T7 promoter on the plasmid) and small amounts of T7 polymerase (and consequently target protein) are produced. Minimizing this background expression is especially important when the target protein is toxic for *E. coli*. This is achieved by co-expression of T7 lysozyme which is encoded on the pLysS or pLysE plasmids. The limited amount of T7 lysozyme is enough to degrade basal amounts of expressed T7 polymerase but does not diminish the expression of the target protein under induced conditions.

As mentioned above, the pET system can be induced with IPTG, TMG or allolactose. IPTG and TMG are synthetic molecules and are referred to as “gratuitous inducers” because they are capable of inducing the lac operon while not serving as substrates for the induced enzymes. These gratuitous inducers can diffuse into the cell without the help of a permease. In addition to this passive uptake to some extent IPTG and

TMG are also taken up via the lactose permease (LacY) (Fernández-Castané et al. 2012). IPTG and TMG are not degraded but inactivation via acetylation by the galactoside O-acetyltransferase (LacA) is possible (Marbach and Bettenbrock 2012). Lactose in contrast is only taken up via LacY and converted into allolactose by beta-galactosidase (LacZ). Degradation of lactose and allolactose is also carried out by LacZ. LacY, LacZ and LacA are expressed via the *lac* operon, which in turn is also induced with IPTG, TMG or allolactose.

The strain *E. coli* BL21(DE3) is a suitable partner for the pET expression system because the strain harbors the necessary copy of the gene for the T7 RNA polymerase in its genome.

The strain *E. coli* BL21(DE3)

The strain BL21(DE3) was created by Studier and Moffatt (1986) by introducing the gene for the strong RNA polymerase of the bacteriophage T7 into the strain BL21. To introduce the gene, the cloning vector DE3, a derivative of the lambda phage vector, was used. The gene for the T7 RNA polymerase was put under the control of the *lacUV5* promoter inducible by lactose or its artificial analogs.

The genotype of BL21(DE3) is described as F^- *ompT gal dcm lon hsdS*$_B$(r_B^- m_B^-) (DE3). A detailed explanation of the genotype is given in Table 1.

Table 1: Description of the genotype of *E. coli* BL21(DE3). Adapted from Novagen, Competent Cells, User Protocol TB009 Rev. H 0211JN.

Genotype	Description
F^-	The F plasmid is missing; therefore the bacterium cannot act as donor in horizontal gene transfer.
ompT	Mutation in the outer membrane protein protease VII to reduce proteolysis of expressed proteins especially after cell lysis.
gal	Galactose is not utilized due to missing enzymes in the Leloir pathway (galactokinase 2.7.1.6; galactose-1-phosphate uridylyltransferase 2.7.7.12; UDP-glucose 4-epimerase 5.1.3.2)
dcm	No cytosine methylation in the sequence CCWGG.
lon	Deletion of the Lon protease for enhanced protein stability.
***hsdS*$_B$(r_B^- m_B^-)**	Mutations in the restriction and methylation systems that allow *E. coli* to recognize foreign DNA. This genotype allows efficient transformation of DNA generated from PCR reactions and reduces plasmid loss.
(DE3)	The strain contains the T7 gene 1 under the control of the *lacUV5* promoter.

E. coli BL21(DE3) is the strain most commonly used for recombinant protein production due to its outstandingly high yields of target protein in combination with the T7*lac* promoter system (Baneyx 1999; Terpe 2006; Tegel et al. 2011).

Lactose uptake in *E. coli*

The uptake of lactose in *E. coli* is regulated via the *lac* operon. Since its first description by Jacob and Monod (1961) the *lac* operon became the standard model of gene regulation. The operon consists of three structural genes *lacY*, *lacZ* and *lacA*, three operator sequences (O_1, O_2 and O_3) and the promoter and terminator sequence (Figure 1). The transcription of the *lac* operon is influenced in three ways: negative regulation, positive regulation, and inducer exclusion.

Negative regulation is achieved via the constitutively expressed LacI repressor. This homotetramer binds to O_1 in close proximity to the promoter sequence and additionally to either O_2 (downstream of O_1) or O_3 (upstream of O_1). Binding of LacI to O_1 prevents the RNA polymerase from initiation at the promoter site. This negative regulatory effect is greatly enhanced by DNA looping due to the additional binding of O_2 or O_3 (Oehler et al. 1990). Since O3 is located adjacent to *lacI*, LacI auto-regulates its expression by hindering the complete transcription of its own gene upon binding O_3 (Abo et al. 2000). Figure 2 illustrates the binding of LacI to the promoter region of the *lac* operon.

Positive regulation is achieved via the catabolite activator protein (CAP). CAP is activated upon binding of cyclic adenosine monophosphate (cAMP). The CAP-cAMP complex binds to the DNA in proximity of the RNA polymerase binding site. Interaction with the RNA polymerase enhances the affinity of the polymerase to the promoter and therefore enhances the expression of the *lac* operon. Since the activity of the CAP-cAMP complex is depending on the cAMP concentration, cAMP concentrations varying with the available carbon source were thought to influence the expression of the *lac* operon (Pastan and Adhya 1976). Although stated otherwise in most text books, according to more recent findings the cAMP concentration does not change significantly regardless if glucose or lactose is utilized. Therefore this positive effect rather generally enhances the promoter strength but does not regulate it (Inada et al. 1996; Crasnier-Mednansky 2008). Also these findings are challenged (Görke

and Stülke 2008a) and the debate about the role of the CAP-cAMP complex in the regulation of the *lac* operon goes on (Görke and Stülke 2008b).

The third regulatory mechanism is inducer exclusion. It prevents the uptake of lactose independent of interaction with the *lac* operon. In *E. coli* glucose is mainly taken up via the phosphoenolpyruvate-phosphotransferase system (PEP-PTS). Via a cascade of enzymes the phosphoryl group of PEP is transferred to the internalized glucose. High glucose concentrations and therefore high activity of this transferase system drain the phosphate groups from the proteins of the PTS cascade. One enzyme of the cascade, EIIA, in its dephosphorylated form then directly inhibits LacY and leads to inducer exclusion (Deutscher et al. 2006).

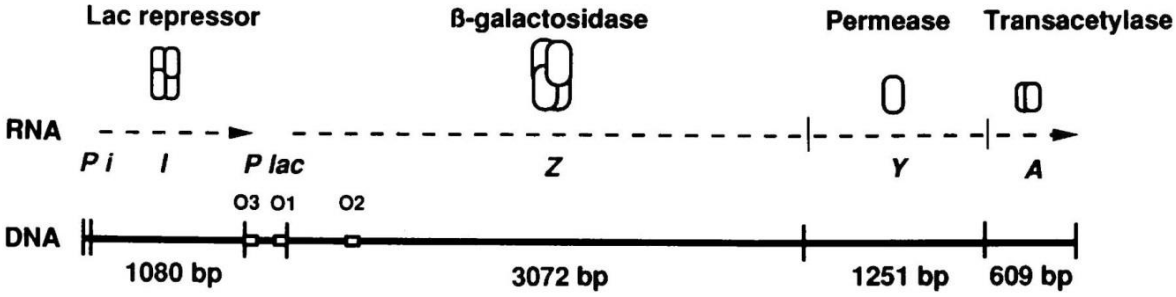


Figure 1 (top): Organization of the *lac* operon and the *lacI* gene. *LacI* and *LacZ* are homotetramers, while *LacY* is a monomer and *LacA* a dimer. *lacI* has its own promoter and terminator sequence, while transcription of the *lac* operon yields polycistronic mRNA. The operator sequence O3 overlaps with *lacI* and O2 is situated inside *lacZ*. *Pi*: *lacI* promoter. *Plac*: *lac* operon promoter. Figure from (Oehler et al. 1994).

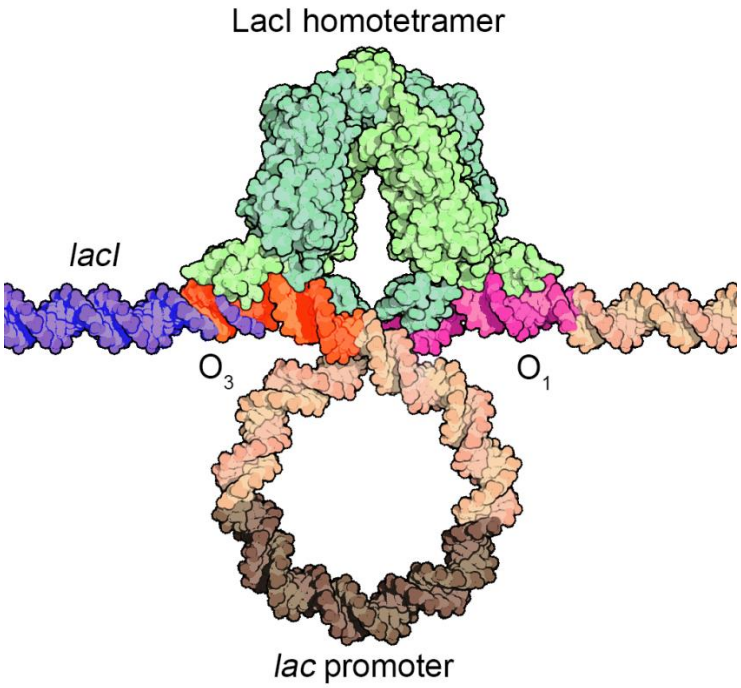


Figure 2 (left): Model of the LacI repressor bound to the promoter region of the lac operon. Upon binding the operator sites O_1 (pink) and O_3 (red) the LacI homotetramer (green) bends the DNA in a way that the promoter site (brown) is no longer accessible for the RNA polymerase. Overlapping of O_3 with *lacI* leads to repression of the transcription if LacI binds to O_3. Figure modified from the protein data bank (PDB; Berman et al. 2000; rcsb.org, retrieved 24 January 2016).

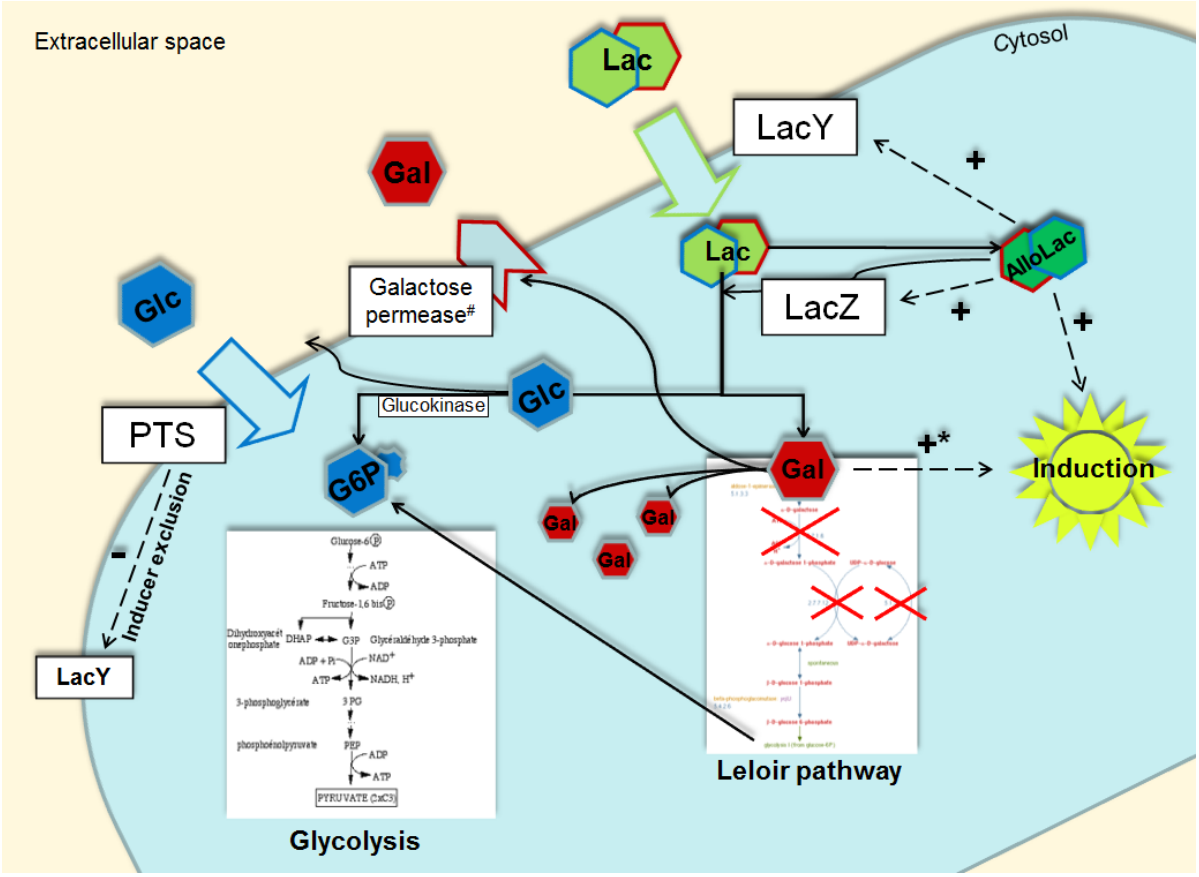


Figure 3: Schematic of the sugar uptake and metabolism in *E. coli* BL21(DE3). The steps of lactose utilization are explained in the running text above the figure. Regulatory effects are indicated with dashed lines. The quality of regulation is indicated with + or -. Induction, symbolized with a greenish yellow star, is achieved via inactivation of the LacI repressor and refers to the induction of an engineered *lac* based expression system. #Huber et al. (1980) mention proton-linked permeases as way of export of internal galactose. *Mattanovich et al. (1998) and Xu et al. (2012) report induction of *lac* based expression systems with galactose.

Provided that the cells metabolic capacity is not completely occupied by the favored carbon source glucose (the effect of inductor exclusion is reduced) available lactose can be taken up by basal levels of expressed lactose permease (LacY). Subsequently lactose is isomerized to allolactose by beta-galactosidase (LacZ). Allolactose then binds to LacI and allosterically causes a conformational change, rendering LacI unable to bind the operator sequences. Now the structure genes of the *lac* operon can be transcribed freely by the RNA polymerase. As a consequence the increasing number of gene products LacY and LacZ further amplifies the uptake and isomerization of lactose and thereby the *lac* operon is induced autocatalytically via a positive feedback loop.

In alternative reactions to the formation of allolactose LacZ is also responsible for the cleavage of lactose and allolactose into glucose and galactose (Juers et al. 2012).

Intracellular glucose is phosphorylated by glucokinase (Meyer et al. 1997). Internal glucose that exceeds the capacity of the glucokinase is excreted and reenters the cell via the PTS pathway as glucose-6-phosphate (Huber et al. 1980; Hogema et al. 1999). Eventually glucose-6-phosphate enters the pathway of glycolysis. Galactose on the other hand is substrate of the Leloir pathway. In the *E. coli* strain BL21(DE3) three key enzymes of this pathway are missing (galactokinase, galactose-1-phosphate uridylyltransferase and UDP-glucose-4-epimerase). Therefore galactose is not converted to glucose-6-phosphate and cannot enter glycolysis. As a result lactose metabolism in *E. coli* BL21(DE3) ultimately leads to galactose accumulation inside the cell and in the growth medium. Huber et al. (1980) mention proton-linked permeases as way of export of internal galactose. A graphic interpretation of of the sugar uptake and metabolism in *E. coli* BL21(DE3) is outlined in Figure 3.

Recombinant protein production in *E. coli*

Today fermentations with *E. coli* are generally performed in high cell density cultures (HCDC). Culture volume, equipment costs, wastewater and downstream volumina all can be reduced while achieving high product titers (Shiloach and Fass 2005).

Described drawbacks of HCDC are high oxygen demand, reduced mixing efficiency, inhibitory CO_2 levels and high heat generation (Lee 1996).

Although these problems can be addressed by switching host strains or promoters, use of different substrates and gassing with oxygen enriched air or pure oxygen, heat and mass transfer remain an issue when upscaling a fermentation process.

Generally, to reach high cell densities after a batch phase a feeding strategy is applied.

Figure 4: 15 L stainless steel Biostat Cplus bioreactor (Satorius, Göttingen, Germany). Fed batch experiments of this Thesis were conducted in this Fermenter. In this picture an induced *E. coli* culture produces the marker protein enhanced green fluorescent protein.

Exponential feeding is used to achieve constant specific growth rates by providing a limiting amount of carbon source. Strategies with feedback control adjust the feed on basis of online process data like pH, or CO_2 evolution and oxygen consumption thereby avoiding over- or underfeeding (Lee et al. 1999).

Crucial for reaching high cell densities is the separation of growth and production phase due to the growth inhibiting effect of the overexpression of recombinant proteins (see next chapter).

Metabolic burden and stress response caused by overexpression of heterologous genes

Metabolic burden can be defined as the amount of resources (raw material and energy) that is withdrawn from the host's metabolism for maintenance and expression of the foreign DNA (Glick 1995).

While the metabolic burden resulting from plasmid maintenance is relatively low (Da Silva and Bailey 1986; Andersson et al. 1996), expression of genes encoded on the plasmid has a significant impact on the host metabolism. The cell division rate decreases upon the burden of high induction levels. The still observable increase of biomass was found to be partly due to increasing cell size rather than cell proliferation as shown in flow cytometry studies (Borth et al. 1998; Soriano et al. 1999). Additionally, impaired glucose uptake has been observed (Neubauer et al. 2003). Since protein production is the most resource-intensive cellular process, competition with the synthesis of housekeeping proteins and maintenance metabolism arises for energy and capacity of the cellular machinery (Neubauer et al. 2003).

Even more severe than simple competition for energy and protein synthesis capacity are induced changes in the host metabolism and triggered stress responses. Substrate level phosphorylation and acetate production, common phenomena of overflow metabolism during batch cultures are also observed during induced fedbatch phases with too high growth rates (Seeger et al. 1995; Sandén et al. 2003).

Depletion of aminoacyl-tRNA induces the so called stringent response resulting in the arrest of tRNA and rRNA synthesis and formation of guanosine 3′,5′-tetraphosphate (ppGpp) (Gallant 1979; Cashel et al. 1996). A shortage of amino acids can be

caused by the overproduction of recombinant protein and even more so if the amino acid pattern of the product differs considerably from the average host proteins. This shows that not only the amount of synthesized recombinant protein adds to the metabolic burden but also the protein´s specific properties can induce stress response.

Accumulation of misfolded proteins can induce a heat-shock-like response (Goff and Goldberg 1985). Heat shock proteins serve to maintain proper protein solubility by assisting the folding process as chaperons and by degrading proteins that are stuck in a misfolded conformation (Schlesinger 1990). While many reports state increased transcription, synthesis rate and accumulation of heat shock chaperons the strength of the response varies considerably among different expression systems and products (Hoffmann and Rinas 2004).

SOS-response is induced by the appearance of single stranded DNA which in turn is a result of the replication of damaged DNA (Little and Mount 1982). SOS-response results in arrested cell growth and mutation prone repair of damaged DNA. SOS-response and appearance of SOS-inducing signals have been reported during heat induction and IPTG induction of recombinant protein production (Lin et al. 2001; Lee et al. 2002). SOS-response in recombinant protein production is thought to be triggered indirectly via presence of single stranded DNA in the course of transcription and plasmid replication (Lin et al. 2001; Hoffmann and Rinas 2004).

These classical mechanisms of stress response involve increased synthesis of stress response proteins combined with accelerated protein degradation (Goff and Goldberg 1985; Harcum and Bentley 1993). Reorganization of the cellular machinery and energy demand for de novo synthesis of stress response proteins thereby contribute to the metabolic burden associated with recombinant protein production (Hoffmann and Rinas 2001).

Impaired growth and substrate uptake as well as reduced cell viability lead to an overall reduction of product yield. Gradual addition of the inducer IPTG via the feed medium has been observed to improve heterologous protein yields (Ramirez and Bentley 1995). Striedner et al. (2003) continuously added inducer (IPTG or lactose) during fed batch phase and achieved a 3.5- (IPTG) or 2-fold (lactose) increase of yield of target protein, compared to a process induced with a single IPTG pulse. Also

Grabherr et al. (2002) by continuously feeding a controlled amount of IPTG managed to keep levels of the stress indicator ppGpp low, while doubling product titer. By utilizing runaway replication Trepod and Mott (2002) managed to gradually increase gene dosage and reported a stable production process for more than 28 h. Induction with weaker inducers like galactose induction of the *lac* promoter has been reported to result in increased final biomass while maintaining the level of specific product concentration (Mattanovich et al. 1998; Xu et al. 2012).

Protein folding and inclusion bodies

Correct protein folding is one of the major challenges during heterologous gene expression in *E. coli*. Different chemical and enzymatic environment between the *E. coli* cytosol and the original host frequently lead to slow, incomplete and/or incorrect folding of the heterologous protein. Differences in pH, osmolarity and temperature as well as increased protein concentration lead to reduced protein solubility. Large proteins in particular often need the help of a complex folding machinery that is not present in *E. coli*. Disulfide bridges cannot be formed in the cytoplasm due to the reducing activity of the thioredoxin-thioredoxin reductase system and the glutaredoxin-glutaredoxin reductase system (Kadokura et al. 2003). Therefore proteins containing disulfide bonds have to be excreted into the periplasm where the disulfide bond formation system (Dsb) promotes the formation of disulfide bonds (Ito and Inaba 2008).

Accelerated translation rates due to strong promoters and high inducer concentrations do not leave enough time for the slower folding kinetics and limited amounts of folding modulators to correctly maturate the heterologous protein. Consequences of incorrect folding are loss of activity, hindered protein transport (e.g. protein transport into the periplasm), accelerated proteolytic degradation and reduced solubility and therefore aggregation as inclusion bodies (IBs).

Some production strategies aim to yield high amounts of IBs to exploit their characteristics for facilitated downstream processes. Producing a protein as IBs is favorable when the active product is toxic for the host or sensitive to proteolytic degradation. Due to the insoluble nature of IBs they can be readily separated via centrifugation from a huge portion of the soluble host cell proteins. IBs have been reported to contain 80% to 95% of the target protein with impurities of hydrophobic

membrane proteins, ribosomal compounds, phospholipids as well as nucleic acids (Valax and Georgiou 1993; Rattenholl et al. 2001).

IBs then have to be solubilized and consecutively refolded to assume their native quaternary structure and full activity. In general solubilization is achieved by resuspending the IBs in alkaline solutions of chaotropic reagents like urea or guanidinium hydrochloride containing reducing agents like 2-mercaptoethanol or dithiothreitol for the cleavage of non-native disulfide bridges (Singh and Panda 2005). To facilitate correct folding of the solubilized protein diafiltration and dilution steps have to be performed. It is also possible to perform on-column refolding.

Solubilization and refolding of IBs is laborious and frequently the procedure suffers from unsatisfactory low yields. Therefore much effort is put into the development of measures to prevent the formation of IBs. Strategies are either based on genetic alterations of the host organism, engineering of the target protein or process technological means.

●) Genetic alteration of the host organism: A number of strains with altered genetics for increased protein solubility are commercially available: strains for improved protein processing at low temperatures (Arctic Express strain, Agilent Technologies); strains for improved disulfide bond formation containing mutations in the glutathione reductase (*gor*) and thioredoxin reductase (*trxB*) genes (Origami strains from Novagen) or strains expressing additional chaperons to help increase the protein solubility by actively assisting the folding process, unfolding incorrectly folded proteins as well as accelerating the degradation of proteins that are stuck in a none-native configuration. Examples of chaperons in *E. coli* are trigger factor, DnaK, GroEL, heat shock proteins (Hsp70 and Hsp60) and ClpB (Costa et al. 2014).

●) Target protein engineering: Another strategy is to engineer target proteins to contain additional solubility tags. Solubility tags are peptides varying in length from a few amino acids to up to several hundred. Mostly tags are fused to the N-terminus of the target protein due to easier tag removal and the possibility to promote translation initiation (Waugh 2005; Malhotra 2009). After purification of the target protein, the solubility tag is removed by enzymatic or chemical cleavage. Besides the necessity for empirical evaluation of new target protein/solubility tag combinations, unspecific tag cleavage, low protein yields during tag removal, costs for proteases and

requirement of additional purification steps after tag removal are disadvantages of the fusion tag technology.

●) Thirdly the production of soluble proteins can be enhanced via process engineering: Reducing the temperature and thereby reducing the rate of protein synthesis is the simplest way to enhance protein solubility (Ferrer et al. 2003; Schumann and Ferreira 2004; Sørensen and Mortensen 2005b). Drawback is the concurrent decrease of growth rate and therefore decreasing overall productivity. Another possibility to increase the amount of soluble protein is induction with weaker inducers like lactose instead of IPTG or TMG (Hoffman et al. 1995; Monteiro et al. 2000).

Another strategy for increasing the protein solubility is tuning the recombinant protein expression to the specific needs.

Expression tuning in *E. coli*

In order to increase protein solubility and reduce the metabolic burden to prolong the production phase the need for tuning of the protein expression arises.

The term “expression tuning” is not clearly defined in the literature and is used for a variety of very different strategies.

In general all these approaches fit in one of the following categories: to influence the expression rate either the promoter strength or the gene dosage can be lowered or the supply of inducer molecules can be limited.

Tuning the expression via the promoter strength is tempting, but a rather crude approach with no possibility of fine tuning and generally comes along with numerous requirements regarding the genetic background of the host, inducer and induction conditions. Nevertheless at the beginning of the development of a bioprocess a suitable promoter has to be chosen prudently and it is possible to select a promoter that roughly meets the required strength for the pursued production strategy. Besides the most frequently used T7 promoter (pET expression system) which can induce expression levels of up to 50% of the total protein as target protein (Baneyx 1999), a variety of other promoters are available. Although focus in the development of many promoter systems was to circumvent basal expression like in the pET

plasmid, all these promoters also vary in their strength. As shown by Tegel et al. (2011) the actual expression rate of certain promoter is greatly dependent on the produced protein with significant influence of the amount of rare codons present. Therefore data concerning the promoter strength can only give a rough overview as presented for a selection of popular promoter systems in Table 2, adapted from Terpe (2006).

Table 2: Selection of popular promoter systems used for the heterologous protein expression in *E. coli*. Adapted from Terpe (2006).

Promoter system	Induction	Level of expression	Key features	Original reference
***lac* promoter**	IPTG, TMG, lactose	Low level up to middle	Suitable for gene products at very low intracellular level	Gronenborn (1976)
***trc* and *tac* promoter**	IPTG, TMG, lactose	Moderately high	High basal expression	Brosius et al. (1985)
T7*lac* promoter	IPTG, TMG, lactose	Very high	Highly specific Basal expression can be reduced with suitable strains (pLys)	Studier and Moffatt (1986)
Phage promoter p_L	Temperature shift from 30 to 42 °C	Moderately high	Temperature-sensitive host required Basal expression No chemical inducer	Elvin et al. (1990)
***tetA* promoter/ operator**	Anhydro-tetracycline	Variable from middle to high	Tight regulation Independent on *E. coli* strain Low basal level	Skerra (1994)
***araBAD* system**	L-arabinose	Variable from low to high	Tight regulation possible Low basal level	Guzman et al. (1995)
***rhaBAD* system**	L-rhamnose	Variable from low to high	Tight regulation Low basal activity	Haldimann et al. (1998)

Another option for expression tuning is to vary the gene dosage via the plasmid copy number. The copy number depends on the genetic background of the host (*endA*$^-$ strains, deficient of the endonuclease A which is responsible for DNA and plasmid degradation, yield higher copy numbers) and most importantly on the origin of replication (*ori*) encoded on the plasmid.

Besides the uncontrolled strategy of runaway replication of the plasmid (Nordström and Uhlin 1992), both the promoter strength and the gene dosage, once altered by genetic engineering, cannot be changed during a fermentation process.

Therefore it may be more suitable to apply a stricter definition for expression tuning and restrict it to strategies utilizing process tuning which is a much more flexible approach. Tuning the expression by lowering the fermentation temperature is a well described strategy (Vera et al. 2007; Pinsach et al. 2008; Rodríguez-Carmona et al. 2012). Unfortunately just like tuning the expression by reducing the available carbon, this leads to decreased growth rates and therefore reduced overall yields are the consequence.

Limiting the inducer concentration can be carried out without the need for genetic alteration of the host and moreover allows for adaption of expression during the course of a fermentation. This makes an inducer based strategy the most promising approach for expression tuning and harbors the potential of being established as a product independent platform technology.

It was shown that inducer based tuning of the expression rate can be achieved by one-point addition of IPTG (Turner et al. 2005; Hartinger et al. 2010). Therefore expression systems that exhibit a defined ratio of supplied inducer and gene expression are required. This is an important aspect that has to be kept in mind when tuning is carried out with inducers that are not gradually taken up, but amplify their uptake via induction of their own permeases like in the case of the T7*lac* system. To circumvent this issue so called tuner strains (Novagen) have been developed. Since the uptake of IPTG is accelerated by the LacY permease (Fernández-Castané et al. 2012), these strains lack the *lacY* gene to uncouple the uptake rates of inducer from the induction level of the cell. Nonetheless systems based on the T7*lac* promoter have been utilized for expression tuning in non-tuner strains also (Hoffman et al. 1995; Striedner et al. 2003; Pei et al. 2011). In this case the inducer has to be supplied in limiting amounts via a feeding strategy so the internal inducer concentration is independent of the amount of permeases and also independent of the amount of beta-galactosidase that catalyze the formation of the actual inducer allolactose.

Further examples of promoter systems that were reported to be tunable via the supplied amount of inducer are: *araBAD* system (Sagmeister et al. 2014), *rhaBAD* system (Giacalone et al. 2006), pPro system (Lee and Keasling 2005).

Unfortunately, these tuning strategies are based on scarcely deployed expression systems that often require expensive inducers. In contrast, this Thesis strives to achieve tunability using the by far most popular T7*lac* based pET expression system in combination with the cheap inducer lactose.

All-or-none induction

It is most desirable to tune the gene expression not only on the population level as average across all cells, but to reach uniform induction of every single cell. This tuning on a cellular level (i) permits the exploitation of the cell factory to the utmost level, since neither space nor substrates are wasted on nonproducing cells and (ii) is precondition to decrease the metabolic burden of producing cells in order to enhanced fidelity of post-translational processes, increase protein solubility and extend the production phase to achieve higher overall yields.

Inhomogeneous induction on the cellular level results in the development of distinct populations. This phenomenon is described in the literature as bistability or all-or-none induction. Novick and Weiner (1957) originally explained the formation of subpopulations to be caused by small random differences in the number of permeases in uninduced cell. Upon addition of inducer the autocatalytic nature of inducer uptake (explained in the chapter “Lactose uptake in *E. coli*”) then leads to distinct cell populations with significant differences in the internal inducer concentration and consequently also in induction levels. The development of bistability during the induction of the *E. coli lac* operon has been subject to extensive modelling efforts (Ozbudak et al. 2004; Van Hoek and Hogeweg 2006; Santillán et al. 2007; Dreisigmeyer et al. 2008; Afroz et al. 2014). All these studies report that the occurrence of bistability depends on the inducer type: The gratuitous inducers IPTG and TMG promote all-or-none induction, while induction with lactose results in a graded response. Since the level of induction is dependent on the amount of inducer molecules inside the cell, it is influenced by two factors: the rate of inducer uptake and the rate of inducer degradation/inactivation/excretion. According to Afroz et al. (2014) high catabolic activity and low inducer uptake promote graded response (lactose case) while low catabolic activity paired with high inducer uptake (like in the case of gratuitous inducers) promotes all-or-none induction.

The marker protein green fluorescent protein (GFP)

GFP is probably one of the most well-known proteins both to the public and in the field of life sciences. The public considers fluorescent mice to be standard laboratory equipment, GFP-rabbits (Kac 2003) found echo in mainstream media (Der Spiegel, 2003, p. 195; Der Standard, 22 Sptember 2000), a bioluminescent eco system was established as art project (Frankfurter Rundschau, 22 August 2001, p. 19), fluorescent trees were contemplated as city lights (Estévez 2007) and *Danio rerio* (zebrafish) engineered to express GFP and other fluorescent proteins are commercially available since 2003 (Yorktown Technologies, Austin, TX, USA).

A PubMed search generates 16,500 hits, containing “green fluorescent protein” in the title or abstract, for the time until 2008 when Osamu Shimomura, Martin Chalfie and Roger Tsien were awarded the Nobel Prize in Chemistry in 2008 for the “discovery and development of GFP”. Until the end of 2015 the search delivers 27,800 hits.

Discovery and development of GFP

GFP was first described by Shimomura et al. (1962). It was discovered together with aequorin, a chemiluminescent protein both produced by the jellyfish *Aequorea vctoria*. While aequorin requires Ca^{2+} ions as cofactor and emits blue light, the wild-type (wt) GFP *in vivo* is exited via radiationless energy transfer from aequorin, eventually resulting in green fluorescence of the jellyfish (Hastings and Morin 1969).

Further milestones were the purification and crystallization of wtGFP and the measurement of the absorption spectrum (Morise et al). Prendergast and Mann were the first to estimate the monomeric molecular mass (Prendergast and Mann 1978) and Shimomura correctly identified the functional core of the chromophore by proteolysis and subsequent analysis of the resulting fluorescent peptide (Shimomura 1979). Cody et al. (1993) solved the correct structure of the whole chromophore, 4-(p-hydroxybenzylidene)imidazolidin-5-one.

Breakthrough for the nowadays widespread use of GFP and its derivatives in molecular biology was the cloning of its gene (Prasher et al. 1992) and successful protein expression in heterologous host species namely *E. coli* and the nematode *Caenorhabditis elegans* (Chalfie et al. 1994; Inouye and Tsuji 1994). The crystal structure was first solved independently by Ormö et al. (1996, PDB accession number 1EMA), and Yang (1997, PDB accession number 1GFL).

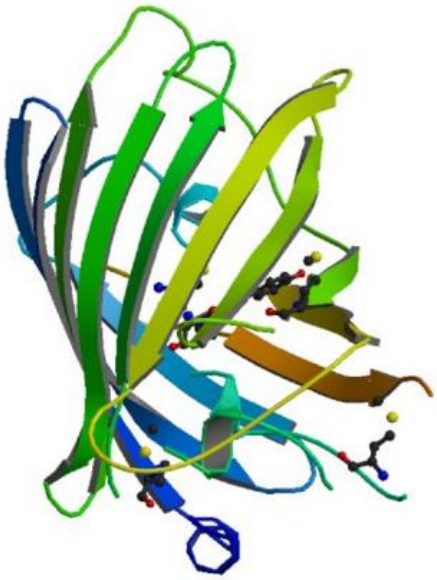


Figure 5: Ribbon diagram of GFP, PDB-entry 1EMA (Ormö et al. 1996). GFP is a small protein of 27 kDa consisting of 238 amino acids. The tertiary structure of the protein is an 11 stranded 2.4 by 4 nm beta-barrel with an alpha-helix, containing the chromophore, running through its center.

GFP is a small protein of 27 kDa consisting of 238 amino acids (Figure 5). The tertiary structure of the protein is an 11 stranded 2.4 by 4 nm beta-barrel with an alpha-helix running through its center (Ormö et al. 1996). This alpha-helix contains the tripeptide sequence Ser65-Tyr66-Gly67 which autocatalytically forms the chromophore after correct folding of the protein (Cubitt et al. 1995). Maturation of the chromophore is a multi-step process and occurs in the protected environment of the readily folded protein with various amino acid residues, most importantly Arg96, (Stepanenko et al. 2008) and trapped water molecules stabilizing the transition states. First a series of torsional adjustments relocate the carboxyl carbon of Ser65 in close proximity to the amino nitrogen of Gly67. Via the following nucleophilic attack which is accompanied by the loss of a water molecule the tripeptide cyclizes. The protein´s fluorescence originates from a π-electron system. To form this system a connection of the phenyl ring of Tyr66 with the newly formed imidazolin-5-one heterocyclic ring system has to be established. This is achieved by oxidizing the alpha-beta carbon bond of Tyr66 with molecular oxygen to yield a double bond (Cubitt et al. 1995; Reid and Flynn 1997). This oxidation results in the equimolar formation of H_2O_2. The mechanism of the chromophore formation is depicted in Figure 6.

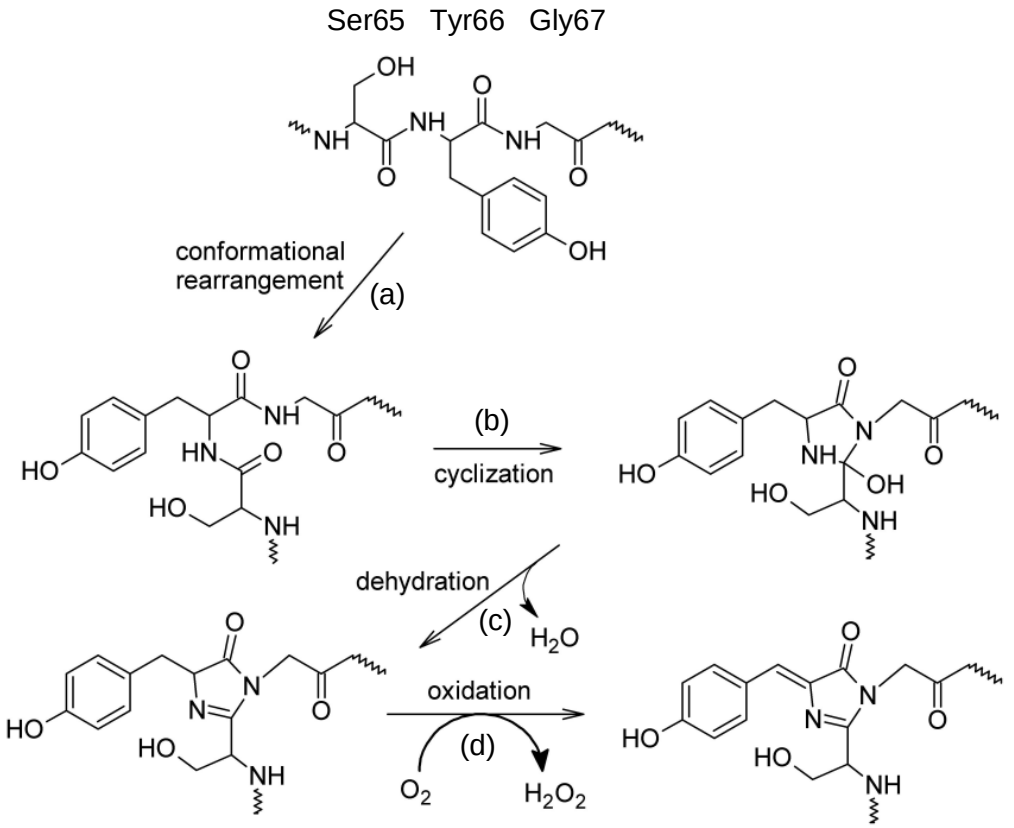


Figure 6: Schematic of the chromophore formation of wtGFP starting with the tripeptide Ser65-Tyr66-Gly67. First a series of conformational rearrangements relocate the carboxyl carbon of Ser65 in close proximity to the amino nitrogen of Gly67 (a). The following nucleophilic attack leads to cyclization of the tripeptide (b). This step is accompanied by the loss of a water molecule (c). Oxidation of the alpha-beta carbon bond of Tyr66 with molecular oxygen connects the p-electron system of the phenyl ring of Tyr66 with the newly formed imidazolin-5-one heterocyclic ring system via formation of a double bond (d). Figure adapted from Moeyaert (2010).

The half-time for the formation of the chromophore is estimated for initial folding to be 10 min, cyclisation and dehydration 3 min and oxidation 19-83 min (Tsien 1998). The time for chromophore oxidation varies greatly among different mutants and takes about four times longer for the wtGFP compared to the Ser65Thr mutant (Heim et al. 1995).

Absence of any cofactors and no need for additional enzymes to catalyze the maturation of GFP are the crucial features that make GFP so popular in molecular biology and promote the application of GFP as marker for protein localization and gene expression *in vivo* in virtually all tissues provided molecular oxygen is present.

Numerous variants of GFP are known (Table 3) which can be distinguished based on the amino acid sequence and structure of the chromophore (Figure 7). Leading drivers for the creation and improvement of GFP variants are the developments made in the Tsien laboratory (University of California, San Diego, CA, USA). Today

GFP variants cover a range of about 100 nm of the visible spectrum from blue to greenish yellow (emission wavelengths). The rest of the visible spectrum is covered by red fluorescent protein, derived from soft coral species *Discosoma* sp., and its derivatives (Shaner et al. 2004, see Figure 8).

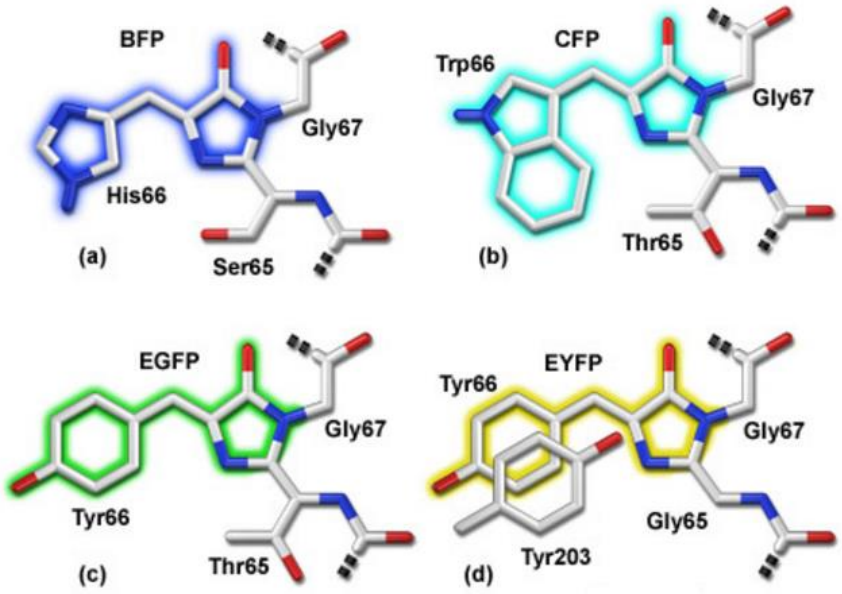


Figure 7: Chromophore structure of GFP variants. The π-electron system is outlined in the colors corresponding to the emitted light of the protein. Mutations of the amino acid positions 65 and 66 lead to changes in the electron density of the delocalized π-electron system, which are responsible for the altered fluorescence spectra. In the case of EYFP (d) not only the amino acids that are neighboring in the primary sequence, the stacking of the aromatic rings of Tyr66 and Tyr203 induced a bathochromic shift and thereby increases the emission wavelength (Day and Davidson 2009).

The presumably most popular variant is enhanced GFP or EGFP which, equipped with a 6xHis tag, is also used in this Thesis. The mutation Phe64Leu was introduced to increase folding efficiency at 37 °C (necessary due to the wild-type´s origin in the much colder North Pacific). Even more important, the mutation Ser65Thr results in a unimodal excitation spectrum (peak excitation wavelength (Ex) = 488 nm, peak emission wavelength (Em) = 507-509 nm) due to chromophore now preferably assuming its anionic state (as phenolate) in contrast to the bimodal spectrum of wtGFP (Ex_1 = 395 nm, Ex_2 = 470 nm, Em = 504 nm) where anionic and neutral state are both present (Zimmer 2002).

Table 3: Optical properties of selected GFP variants. The peak excitation (Ex) and emission (Em) wavelengths, molar extinction coefficient (EC), quantum yield (QY) and relative brightness, calculated as the product of EC times QY divided by the value for EGFP, are listed. This table was adapted from Day and Davidson (2009).

Protein (acronym)	Ex [nm]	Em [nm]	EC [$mM^{-1}cm^{-1}$]	QY	Brightness [% of EGFP]	Reference
Blue Fluorescent Proteins						
Sirius	355	424	15.0	0.24	11	(Tomosugi et al. 2009)
Azurite	384	450	26.2	0.55	43	(Mena et al. 2006)
EBFP	383	445	29.0	0.31	27	(Patterson et al. 1997)
EBFP2	383	448	32.0	0.56	53	(Ai et al. 2007)
Cyan Fluorescent Proteins						
ECFP	439	476	32.5	0.40	39	(Cubitt et al. 1995)
Cerulean	433	475	43.0	0.62	79	(Rizzo et al. 2004)
CyPet	435	477	35.0	0.51	53	(Nguyen and Daugherty 2005)
mTurquoise	434	474	34.0	0.84	85	(Goedhart et al. 2012)
SCFP	433	474	30.0	0.50	45	(Kremers et al. 2006)
Green Fluorescent Proteins						
EGFP	488	507	56.0	0.60	100	(Heim et al. 1995)
Emerald	487	509	57.5	0.68	116	(Cubitt et al. 1998)
Superfolder	485	510	83.3	0.65	160	(Pédelacq et al. 2006)
T-Sapphire	399	511	44.0	0.60	79	(Zapata-Hommer and Griesbeck 2003)
Yellow Fluorescent Proteins						
EYFP	514	527	83.4	0.61	151	(Miyawaki et al. 1999)
Citrine	516	529	77.0	0.76	174	(Griesbeck et al. 2001)
Venus	515	528	82.2	0.57	156	(Nagai et al. 2002)
Topaz	514	527	94.5	0.60	169	(Tsien 1998)
YPet	517	530	104.0	0.77	238	(Nguyen and Daugherty 2005)
SYFP	515	527	101.0	0.68	204	(Kremers et al. 2006)
mAmetrine	406	526	45.0	0.58	78	(Ai et al. 2008)

In addition to those two modifications numerous mutations have been conducted to generate proteins with altered excitation and emission spectra. Blue (BFP), cyan (CFP) and yellow (YFP) variants were developed. This enables the simultaneous use of several GFP variants and the combination with other fluorophores during an experiment. On the DNA level various codon optimized variants of GFP for mammals, plants, yeast and fungi were designed.

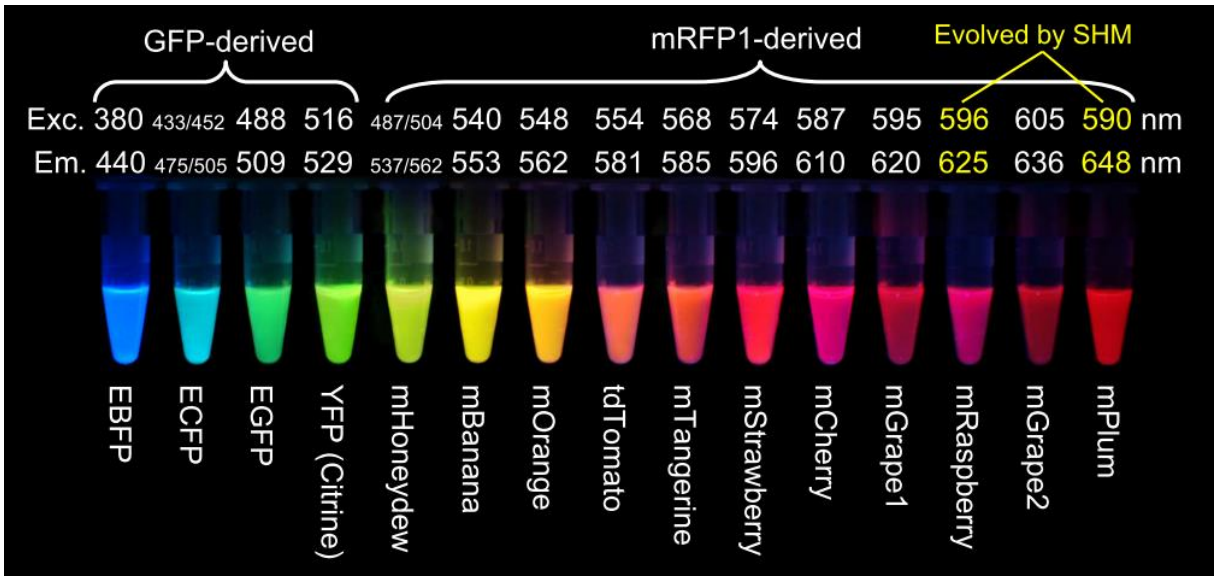


Figure 8: Palette of nonoligomerizing fluorescent proteins (Figure from Tsien (2009)). GFP-derived proteins originate from *Aequorea Victoria* and cover about 100 nm of the visible spectrum from blue to greenish yellow, while derivatives of the red fluorescent protein (mRFP1) originate from *Discosoma* sp. cover another 100 nm of the higher wavelength area. SHM: somatic hypermutation.

Applications of GFP and GFP derivatives

As already mentioned the autocatalytic formation of the chromophore without the need for a cofactor or assisting enzymes is one of the main reasons for GFP being such a popular marker protein. Additionally, GFP is nontoxic in most cases and due to its small size of 27 kDa and only a weak tendency for dimerization, fusion of GFP to a protein rarely affects mobility and activity of the fusion partner. The tendency for dimerization was successfully eliminated by replacing hydrophobic amino acids at crystallographic dimer interface with positively charged residues: Ala206Lys, Leu221Lsy, or Phe223Arg (Zacharias et al. 2002). GFP shows high stability towards heat, alkaline pH, detergents, chaotropic salts, photobleaching and many proteases (Ehrmann et al. 2001; Ward 2005).

Limitations of GFP are the time required for the chromophore formation, the need for molecular oxygen and low brightness when present in low concentrations compared to background fluorescence (Tsien 1998; Zimmer 2002). A drawback for quantification of overexpressed GFP is that reportedly the specific fluorescence of IBs is not only decreased - García-Fruitós et al. (2005) report a specific activity of IBs of 20% compared to SP - but also depends on process parameters like temperature and growth rate (Iafolla et al. 2008) and on the induction regime (Peternel et al. 2009).

Reporter for gene expression: The first application of GFP was as a reporter for gene expression (Chalfie et al. 1994). Up to date GFP has been extensively used for this purpose and is still commonly used as reporter gene. Promoters of interest and the corresponding genes can be investigated regarding their location and time of activity in living cells. Major hindrance when investigating weak promoters is the absence of signal amplification because no substrate is being turned over and one expressed protein results in only one fluorescent chromophore. Additionally the slow chromophore formation limits the possibility to track fast shifts in expression patterns.

Fusion tag: When GFP is fused to a protein of interest, protein location and cellular dynamics can be investigated. In general GFP is fused via its N- or C-terminus to the protein of interest, but also circular permutations which incorporate the protein of interest within the GFP sequence via newly created N- and C-termini are described (Baird et al. 1999). GFP not altering the activity, location and cell-to-cell motility of its fusion partner were perfect conditions to promote the use as fusion tag to become the most common application of GFP in molecular biology.

***In vivo* pH indicator:** EGFP as well as many other GFP variants exhibit a strong pH dependency in their fluorescence due to reversible protonation of the anionic state (Niwa et al. 1996; Elsliger et al. 1999). pH dependent changes of the fluorescence of GFP variants have been reported in the range of pH 5-12 (Ward and Bokman 1982; Kneen et al. 1998). The non-toxicity of GFP allows for *in vivo* pH indication of organelles and cell compartments. GFP-based pH markers have been successfully applied for the monitoring of cytosolic, mitochondrial and Golgi pH (Takahashi et al. 2001). While EGFP (pKa = 6) was used for investigation of the acidic secretory pathway and Golgi environment, EYFG (pKa = 7) is better suited for investigation of the cytosolic and mitochondrial pH. Many pH sensitive GFP variants have been developed by modification of key amino acids in the spatial vicinity of the chromophore. Miesenböck et al. (1998) coined the term pHluorins and grouped them in two classes: ratiometric and ecliptic pHluorins. In ratiometric pHluorins a change in pH induces a shift of excitation maximum (generally between about 395 nm to 475 nm) while ecliptic pHluorins lose their fluorescence in milieus of decreasing pH (Zimmer 2002).

***In vivo* FRET biosensor:** Foerster resonance energy transfer (FRET) is a quantum mechanical phenomenon resulting in radiationless energy transfer between two

chromophores. Prerequisites are an overlap of the emission spectrum of the donor and the excitation spectrum of the acceptor molecule as well as spatial proximity (< 10 nm), since FRET decreases with the distance to the power of six between the partners. CFP and YFP have overlapping spectra and are therefore commonly used as a FRET pair.

The first approach for GFP based FRET applications was the investigation of proteolytic activity (Heim and Tsien 1996). FRET partners are connected with a cleavable linker and fluorescence is decreased upon proteolysis of the linker. In a second approach FRET partners are fused to different proteins for reporting protein-protein interactions in real time. Thirdly fusion to opposite positions of a single protein or peptide enables studying conformational changes of the protein or peptide of interest. Conformational change is a result of ligand binding or other protein modifications and often a prerequisite for enzyme activity. In addition to changes in the distance, changes in the orientation of the FRET partners are sufficient to shift their fluorescence.

Among other applications FRET biosensors can be used to measure Ca^{2+} concentrations (Miyawaki et al. 1997), cyclic adenosine monophosphate (Zaccolo et al. 2000), phosphorylation activity (Ng et al. 1999), protease activity (Hwang et al. 2008) as well as signaling of small G proteins (Janetopoulos et al. 2001). Even FRET systems to study multiprotein complexes and protein trimerization, consisting of three FRET partners (CFP, YFP and monomeric red fluorescent protein mRFP), have been reported (Galperin et al. 2004; He et al. 2005).

Photobleaching: Photobleaching is used to investigate *in vivo* protein dynamics. A cellular region containing the protein of interest fused to GFP is bleached by laser illumination. In fluorescence recovery after photobleaching (FRAP) experiments the time for recovery of the lost fluorescence corresponds to the mobility of the protein of interest. Fluorescence loss in photobleaching (FLIP) is a complementary technique that additionally allows for the determination of flux between distinct cellular regions and for the determination of immobilized fusion constructs. Therefore photobleaching is conduced multiple times outside of the region of interest. If connected, mobile GFP tagged proteins move from the region of interest to the photobleaching area and are bleached leading to a loss of fluorescence, while immobile tagged proteins remain in the region of interest and maintain their fluorescent. Photobleaching experiments

have been extensively and very successfully conduced to investigate protein movement within and between cellular compartments (Ishikawa-Ankerhold et al. 2012).

Anion sensor: The YFP-His148Gln mutant can be used as indicator for halides and anions. Certain anions reduce the fluorescence of the YFP mutant due to protonation of the chromophore upon binding reversibly to the protein. YFP-His148Gln is reported to be sensitive to many anions with the indicated relative potencies: F^- ~ ClO_4^- > I^- > SCN^- > NO_3^- > Cl^- > Br^- > formate > acetate (Jayaraman et al. 2000).

Further applications: Gather and Yun (2011) reported the first successful realization of biological lasers based on GFP. They could show that fluorescent proteins are a viable gain medium for optical amplification. Therefore they successfully used both GFP solutions and living human embryonic kidney cells which stayed alive and were unharmed during the lasing.

Goal and scientific questions

Goal of this Thesis is the development of an induction strategy for tuning the recombinant protein production rate in *E. coli*, using glucose as primary carbon source and lactose as inducer and secondary carbon source. Therefore the model protein EGFP was produced in *E. coli* BL21(DE3) in a mixed feed environment. The scientific challenge of this Thesis can be summarized with the following three scientific questions.

1st Scientific question

Is the relationship between the concomitant specific substrate uptake rates of* E. coli *BL21(DE3) transferable between strains producing different target proteins?

The dependency of the maximum specific lactose uptake rate ($q_{Lac,\ MAX}$) on the specific glucose uptake rate (q_{Glc}) during lactose excess is investigated in a BL21(DE3) strain producing EGFP. Results are compared with investigations of a BL21(DE3) strain producing a scFv antibody fragment, conducted in a former work. The influence of the produced heterologous protein on the substrate uptake is investigated.

To assess the transferability between the two strains four fermentations at different specific glucose uptake rates (q_{Glc}) with a permanent excess of lactose were performed.

2nd Scientific question

Are the protein production rate and the ratio of SP to IBs tunable by adjusting the specific lactose uptake rate?

The influence of the specific lactose uptake rate on the recombinant protein production rate and the ratio of soluble protein to inclusion bodies were investigated.

Therefore four fermentations with different specific lactose uptake rates (q_{Lac}) were performed at a set value of q_{Glc}.

3rd Scientific question

Does our induction strategy promote the phenomenon of all-or-none induction or is it possible to tune the expression on a cellular level?

Time resolved measurements of the cellular induction levels allowed for the detection of possible subpopulations formed during the course of the induction phase.

To investigate the all-or-none induction phenomenon confocal laser scanning microscopy (CLSM) was applied.

I answer these three scientific questions in form of a paper manuscript. An elaborate discussion and outlook of this Thesis as well as data, methods and results not shown in the manuscript draft are presented and discussed thereafter in the Appendix.

The conceptual framework for this study was set by my supervisors Dipl.-Ing. David Wurm (D.W.) and Dipl.-Ing. Dr.nat.techn. Oliver Spadiut. D.W. and I jointly designed the experiments. I performed the experiments and laboratory work under supervision of D.W and partly with the help of Tadej Bosilj. I prepared the manuscript draft and together with D.W. jointly conducted data analysis and interpretation. Development of the used model (described by Wurm et al. Manuscript in preparation) was not part of my Thesis and was performed by Sophia Ulonska and D.W.

References

Abo, T., Inada, T., et al. (2000). *SsrA-mediated tagging and proteolysis of LacI and its role in the regulation of lac operon.* The EMBO Journal 19(14): 3762-3769.

Afroz, T., Biliouris, K., et al. (2014). *Bacterial sugar utilization gives rise to distinct single-cell behaviours.* Molecular microbiology 93(6): 1093-1103.

Ai, H.-w., Hazelwood, K. L., et al. (2008). *Fluorescent protein FRET pairs for ratiometric imaging of dual biosensors.* Nature methods 5(5): 401-403.

Ai, H.-w., Shaner, N. C., et al. (2007). *Exploration of new chromophore structures leads to the identification of improved blue fluorescent proteins.* Biochemistry 46(20): 5904-5910.

Andersson, L., Yang, S., et al. (1996). *Impact of plasmid presence and induction on cellular responses in fed batch cultures of Escherichia coli.* Journal of biotechnology 46(3): 255-263.

Baird, G. S., Zacharias, D. A., et al. (1999). *Circular permutation and receptor insertion within green fluorescent proteins.* Proceedings of the National Academy of Sciences 96(20): 11241-11246.

Baneyx, F. (1999). *Recombinant protein expression in Escherichia coli.* Current opinion in biotechnology 10(5): 411-421.

Berman, H. M., Westbrook, J., et al. (2000). *The protein data bank.* Nucleic acids research 28(1): 235-242.

Blattner, F. R., Plunkett, G., et al. (1997). *The complete genome sequence of Escherichia coli K-12.* Science 277(5331): 1453-1462.

Borth, N., Mitterbauer, R., et al. (1998). *Flow cytometric analysis of bacterial physiology during induction of foreign protein synthesis in recombinant Escherichia coli cells.* Cytometry 31(2): 125-129.

Brosius, J., Erfle, M., et al. (1985). *Spacing of the-10 and-35 regions in the tac promoter. Effect on its in vivo activity.* Journal of Biological Chemistry 260(6): 3539-3541.

Cashel, M., Gentry, D., et al. (1996). *The stringent response.* Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology. F. Neidhardt. Washington, ASM Press: 1458–1496.

Chalfie, M., Tu, Y., et al. (1994). *Green fluorescent protein as a marker for gene expression.* Science 263(5148): 802-805.

Choi, J. H., Keum, K. C., et al. (2006). *Production of recombinant proteins by high cell density culture of Escherichia coli.* Chemical Engineering Science 61(3): 876-885.

Cody, C. W., Prasher, D. C., et al. (1993). *Chemical structure of the hexapeptide chromophore of the Aequorea green-fluorescent protein.* Biochemistry 32(5): 1212-1218.

Cohen, S. N., Chang, A. C., et al. (1973). *Construction of biologically functional bacterial plasmids in vitro.* Proceedings of the National Academy of Sciences 70(11): 3240-3244.

Costa, S., Almeida, A., et al. (2014). *Fusion tags for protein solubility, purification and immunogenicity in Escherichia coli: the novel Fh8 system.* Frontiers in microbiology 5.

Crasnier-Mednansky, M. (2008). *Is there any role for cAMP–CRP in carbon catabolite repression of the Escherichia coli lac operon?* Nature Reviews Microbiology 6(12): 954-954.

Cubitt, A. B., Heim, R., et al. (1995). *Understanding, improving and using green fluorescent proteins.* Trends in biochemical sciences 20(11): 448-455.

Cubitt, A. B., Woollenweber, L. A., et al. (1998). *Understanding Structure—Function Relationships in the Aequorea victoria Green Fluorescent Protein.*

Da Silva, N. A. and Bailey, J. E. (1986). *Theoretical growth yield estimates for recombinant cells.* Biotechnology and bioengineering 28(5): 741-746.

Day, R. N. and Davidson, M. W. (2009). *The fluorescent protein palette: tools for cellular imaging.* Chemical Society Reviews 38(10): 2887-2921.

Deutscher, J., Francke, C., et al. (2006). *How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria.* Microbiology and Molecular Biology Reviews 70(4): 939-1031.
Dreisigmeyer, D. W., Stajic, J., et al. (2008). *Determinants of bistability in induction of the Escherichia coli lac operon.* IET systems biology 2(5): 293-303.
Ehrmann, M., Scheyhing, C., et al. (2001). *In vitro stability and expression of green fluorescent protein under high pressure conditions.* Letters in applied microbiology 32(4): 230-234.
Elsliger, M.-A., Wachter, R. M., et al. (1999). *Structural and spectral response of green fluorescent protein variants to changes in pH.* Biochemistry 38(17): 5296-5301.
Elvin, C. M., Thompson, P. R., et al. (1990). *Modified bacteriophage lambda promoter vectors for overproduction of proteins in Escherichia coli.* Gene 87(1): 123-126.
Escherich, T. (1886). *Die darmbakterien des säuglings und ihre beziehungen zur physiologie der Verdauung*, F. Enke.
Estévez, A. T. (2007). *The genetic creation of bioluminescent plants for urban and domestic use.* Leonardo 40(1): 18-18.
Fernández-Castané, A., Vine, C. E., et al. (2012). *Evidencing the role of lactose permease in IPTG uptake by Escherichia coli in fed-batch high cell density cultures.* Journal of biotechnology 157(3): 391-398.
Ferrer, M., Chernikova, T. N., et al. (2003). *Chaperonins govern growth of Escherichia coli at low temperatures.* Nature biotechnology 21(11): 1266-1267.
Gallant, J. A. (1979). *Stringent control in E. coli.* Annual review of genetics 13(1): 393-415.
Galperin, E., Verkhusha, V. V., et al. (2004). *Three-chromophore FRET microscopy to analyze multiprotein interactions in living cells.* Nature methods 1(3): 209-217.
García-Fruitós, E., González-Montalbán, N., et al. (2005). *Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins.* Microbial cell factories 4(1): 27.
Gather, M. C. and Yun, S. H. (2011). *Single-cell biological lasers.* Nature Photonics 5(7): 406-410.
Giacalone, M. J., Gentile, A. M., et al. (2006). *Toxic protein expression in Escherichia coli using a rhamnose-based tightly regulated and tunable promoter system.* Biotechniques 40(3): 355.
Glick, B. R. (1995). *Metabolic load and heterologous gene expression.* Biotechnology advances 13(2): 247-261.
Goedhart, J., von Stetten, D., et al. (2012). *Structure-guided evolution of cyan fluorescent proteins towards a quantum yield of 93%.* Nature communications 3: 751.
Goff, S. A. and Goldberg, A. L. (1985). *Production of abnormal proteins in E. coli stimulates transcription of Ion and other heat shock genes.* Cell 41(2): 587-595.
Görke, B. and Stülke, J. (2008a). *Carbon catabolite repression in bacteria: many ways to make the most out of nutrients.* Nature Reviews Microbiology 6(8): 613-624.
Görke, B. and Stülke, J. (2008b). *Is there any role for cAMP–CRP in carbon catabolite repression of the Escherichia coli lac operon? Reply from Görke and Stülke.* Nature Reviews Microbiology 6(12): 954-954.
Grabherr, R., Nilsson, E., et al. (2002). *Stabilizing plasmid copy number to improve recombinant protein production.* Biotechnology and bioengineering 77(2): 142-147.
Griesbeck, O., Baird, G. S., et al. (2001). *Reducing the environmental sensitivity of yellow fluorescent protein mechanism and applications.* Journal of Biological Chemistry 276(31): 29188-29194.
Gronenborn, B. (1976). *Overproduction of phage lambda repressor under control of the lac promotor of Escherichia coli.* Molecular and General Genetics MGG 148(3): 243-250.
Guzman, L.-M., Belin, D., et al. (1995). *Tight regulation, modulation, and high-level expression by vectors containing the arabinose PBAD promoter.* Journal of bacteriology 177(14): 4121-4130.

Haldimann, A., Daniels, L. L., et al. (1998). *Use of New Methods for Construction of Tightly Regulated Arabinose and Rhamnose Promoter Fusions in Studies of the Escherichia coli Phosphate Regulon.* Journal of bacteriology 180(5): 1277-1286.

Harcum, S. W. and Bentley, W. E. (1993). *Response dynamics of 26-, 34-, 39-, 54-, and 80-kDa proteases in induced cultures of recombinant Escherichia coli.* Biotechnology and bioengineering 42(6): 675-685.

Hartinger, D., Heinl, S., et al. (2010). *Enhancement of solubility in Escherichia coli and purification of an aminotransferase from Sphingopyxis sp. MTA 144 for deamination of hydrolyzed fumonisin B 1.* Microbial cell factories 9: 62-62.

Hastings, J. and Morin, J. (1969). Comparative Biochemistry of Calcium-Activated Photoproteins from Ctenophore, Mnemiopsis and Coelenterates Aequorea, Obelia, Pelagia and Renilla. Biological Bulletin, MARINE BIOLOGICAL LABORATORY 7 MBL ST, WOODS HOLE, MA 02543.

He, L., Wu, X., et al. (2005). *Determination of tumor necrosis factor receptor-associated factor trimerization in living cells by CFP→ YFP→ mRFP FRET detected by flow cytometry.* Nucleic acids research 33(6): e61-e61.

Heim, R., Cubitt, A. B., et al. (1995). *Improved green fluorescence.* Nature 373(6516): 663-664.

Heim, R. and Tsien, R. Y. (1996). *Engineering green fluorescent protein for improved brightness, longer wavelengths and fluorescence resonance energy transfer.* Current Biology 6(2): 178-182.

Hoffman, B. J., Broadwater, J. A., et al. (1995). *Lactose fed-batch overexpression of recombinant metalloproteins in Escherichia coli BL21 (DE3): process control yielding high levels of metal-incorporated, soluble protein.* Protein expression and purification 6(5): 646-654.

Hoffmann, F. and Rinas, U. (2001). *On-line estimation of the metabolic burden resulting from the synthesis of plasmid-encoded and heat-shock proteins by monitoring respiratory energy generation.* Biotechnology and bioengineering 76(4): 333-340.

Hoffmann, F. and Rinas, U. (2004). *Stress induced by recombinant protein production in Escherichia coli.* Physiological Stress Responses in Bioprocesses, Springer: 73-92.

Hogema, B. M., Arents, J. C., et al. (1999). *Autoregulation of lactose uptake through the LacYpermease by enzyme IIAGlc of the PTS in Escherichia coli K-12.* Molecular microbiology 31(6): 1825-1833.

Huber, R. E., Lytton, J., et al. (1980). *Efflux of beta-galactosidase products from Escherichia coli.* Journal of bacteriology 141(2): 528-533.

Hwang, Y.-C., Chu, J. J.-H., et al. (2008). *Rapid identification of inhibitors that interfere with poliovirus replication using a cell-based assay.* Antiviral research 77(3): 232-236.

Inada, T., Kimata, K., et al. (1996). *Mechanism responsible for glucose–lactose diauxie in Escherichia coli: challenge to the cAMP model.* Genes to Cells 1(3): 293-301.

Inouye, S. and Tsuji, F. I. (1994). *Aequorea green fluorescent protein: Expression of the gene and fluorescence characteristics of the recombinant protein.* Febs Letters 341(2): 277-280.

Ishikawa-Ankerhold, H. C., Ankerhold, R., et al. (2012). *Advanced fluorescence microscopy techniques—Frap, Flip, Flap, Fret and flim.* Molecules 17(4): 4047-4132.

Ito, K. and Inaba, K. (2008). *The disulfide bond formation (Dsb) system.* Current opinion in structural biology 18(4): 450-458.

Jacob, F. and Monod, J. (1961). *Genetic regulatory mechanisms in the synthesis of proteins.* Journal of molecular biology 3(3): 318-356.

Janetopoulos, C., Jin, T., et al. (2001). *Receptor-mediated activation of heterotrimeric G-proteins in living cells.* Science 291(5512): 2408-2411.

Jayaraman, S., Haggie, P., et al. (2000). *Mechanism and cellular applications of a green fluorescent protein-based halide sensor.* Journal of Biological Chemistry 275(9): 6047-6050.

Juers, D. H., Matthews, B. W., et al. (2012). *LacZ β-galactosidase: Structure and function of an enzyme of historical and molecular biological importance.* Protein Science 21(12): 1792-1807.
Kac, E. (2003). *GFP bunny.* Leonardo 36(2): 97-102.
Kadokura, H., Katzen, F., et al. (2003). *Protein disulfide bond formation in prokaryotes.* Annual review of biochemistry 72(1): 111-135.
Kneen, M., Farinas, J., et al. (1998). *Green fluorescent protein as a noninvasive intracellular pH indicator.* Biophysical journal 74(3): 1591-1599.
Kremers, G.-J., Goedhart, J., et al. (2006). *Cyan and yellow super fluorescent proteins with improved brightness, protein folding, and FRET Förster radius.* Biochemistry 45(21): 6570-6580.
Lee, J., Kim, H. C., et al. (2002). *Interplay of SOS induction, recombinant gene expression, and multimerization of plasmid vectors in Escherichia coli.* Biotechnology and bioengineering 80(1): 84-92.
Lee, J., Lee, S. Y., et al. (1999). *Control of fed-batch fermentations.* Biotechnology advances 17(1): 29-48.
Lee, S. K. and Keasling, J. D. (2005). *A propionate-inducible expression system for enteric bacteria.* Applied and environmental microbiology 71(11): 6856-6862.
Lee, S. Y. (1996). *High cell-density culture of Escherichia coli.* Trends in biotechnology 14(3): 98-105.
Lin, H. Y., Hanschke, R., et al. (2001). *Cellular response to strong overexpression of recombinant genes in Escherichia coli.* Recombinant protein production with prokaryotic and eukaryotic cells. A comparative view on host physiology. O.-W. Merten, D. Mattanovich, C. Lang et al. Netherlands, Springer: 55.
Little, J. W. and Mount, D. W. (1982). *The SOS regulatory system of Escherichia coli.* Cell 29(1): 11-22.
Maheshwari, S. (2011). "Global protein therapeutics market: beefing up towards futuristic growth." Retrieved December 6th 2015, from http://www.pharmaphorum.com/articles/global-protein-therapeutics-market-beefing-up-towards-futuristic-growth.
Malhotra, A. (2009). *Tagging for protein expression.* Methods in enzymology 463: 239-258.
Marbach, A. and Bettenbrock, K. (2012). *lac operon induction in Escherichia coli: Systematic comparison of IPTG and TMG induction and influence of the transacetylase LacA.* Journal of biotechnology 157(1): 82-88.
Mattanovich, D., Kramer, W., et al. (1998). *Rational design of an improved induction scheme for recombinant Escherichia coli.* Biotechnology and bioengineering 58(2-3): 296-298.
Mena, M. A., Treynor, T. P., et al. (2006). *Blue fluorescent proteins with enhanced brightness and photostability from a structurally targeted library.* Nature biotechnology 24(12): 1569-1571.
Meyer, D., Schneider-Fresenius, C., et al. (1997). *Molecular characterization of glucokinase from Escherichia coli K-12.* Journal of bacteriology 179(4): 1298-1306.
Miesenböck, G., De Angelis, D. A., et al. (1998). *Visualizing secretion and synaptic transmission with pH-sensitive green fluorescent proteins.* Nature 394(6689): 192-195.
Miyawaki, A., Griesbeck, O., et al. (1999). *Dynamic and quantitative Ca2+ measurements using improved cameleons.* Proceedings of the National Academy of Sciences 96(5): 2135-2140.
Miyawaki, A., Llopis, J., et al. (1997). *Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin.* Nature 388(6645): 882-887.
Moeyaert, B. (2010). *Structural insight into fluorescent protein properties by rational design and characterization of photoconvertible, photoswitchable and phototoxic fluorescent proteins*, Katholieke Universiteit Leuven.
Monteiro, R. A., Souza, E. M., et al. (2000). *Use of lactose to induce expression of soluble NifA protein domains of Herbaspirillum seropedicae in Escherichia coli.* Canadian journal of microbiology 46(11): 1087-1090.

Nagai, T., Ibata, K., et al. (2002). *A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications.* Nature biotechnology 20(1): 87-90.
Neubauer, P., Lin, H., et al. (2003). *Metabolic load of recombinant protein production: inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in Escherichia coli.* Biotechnology and bioengineering 83(1): 53-64.
Ng, T., Squire, A., et al. (1999). *Imaging protein kinase Cα activation in cells.* Science 283(5410): 2085-2089.
Nguyen, A. W. and Daugherty, P. S. (2005). *Evolutionary optimization of fluorescent proteins for intracellular FRET.* Nature biotechnology 23(3): 355-360.
Niwa, H., Inouye, S., et al. (1996). *Chemical nature of the light emitter of the Aequorea green fluorescent protein.* Proceedings of the National Academy of Sciences 93(24): 13617-13622.
Nordström, K. and Uhlin, B. E. (1992). *Runaway–Replication Plasmids as Tools to Produce Large Quantities of Proteins from Cloned Genes in Bacteria.* Nature biotechnology 10(6): 661-666.
Novick, A. and Weiner, M. (1957). *Enzyme induction as an all-or-none phenomenon.* Proceedings of the National Academy of Sciences of the United States of America 43(7): 553.
Oehler, S., Amouyal, M., et al. (1994). *Quality and position of the three lac operators of E. coli define efficiency of repression.* The EMBO Journal 13(14): 3348.
Oehler, S., Eismann, E. R., et al. (1990). *The three operators of the lac operon cooperate in repression.* The EMBO Journal 9(4): 973.
Ormö, M., Cubitt, A. B., et al. (1996). *Crystal structure of the Aequorea victoria green fluorescent protein.* Science 273(5280): 1392-1395.
Ozbudak, E. M., Thattai, M., et al. (2004). *Multistability in the lactose utilization network of Escherichia coli.* Nature 427(6976): 737-740.
Palomares, L. A., Kuri-Breña, F., et al. (2002). *Industrial recombinant protein production*, EOLSS Publishers, Oxford. 6.
Pastan, I. and Adhya, S. (1976). *Cyclic adenosine 5'-monophosphate in Escherichia coli.* Bacteriological reviews 40(3): 527.
Patterson, G. H., Knobel, S. M., et al. (1997). *Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy.* Biophysical journal 73(5): 2782.
Pédelacq, J.-D., Cabantous, S., et al. (2006). *Engineering and characterization of a superfolder green fluorescent protein.* Nature biotechnology 24(1): 79-88.
Pei, X.-L., Wang, Q.-Y., et al. (2011). *Efficient production of a thermophilic 2-deoxyribose-5-phosphate aldolase in glucose-limited fed-batch cultivations of Escherichia coli by continuous lactose induction strategy.* Applied biochemistry and biotechnology 165(2): 416-425.
Pinsach, J., de Mas, C., et al. (2008). *Influence of process temperature on recombinant enzyme activity in Escherichia coli fed-batch cultures.* Enzyme and microbial technology 43(7): 507-512.
Prasher, D. C., Eckenrode, V. K., et al. (1992). *Primary structure of the Aequorea victoria green-fluorescent protein.* Gene 111(2): 229-233.
Prendergast, F. G. and Mann, K. G. (1978). *Chemical and physical properties of aequorin and the green fluorescent protein isolated from Aequorea forskalea.* Biochemistry 17(17): 3448-3453.
Ramirez, D. M. and Bentley, W. E. (1995). *Fed-batch feeding and induction policies that improve foreign protein synthesis and stability by avoiding stress responses.* Biotechnology and bioengineering 47(5): 596-608.
Rattenholl, A., Lilie, H., et al. (2001). *The pro-sequence facilitates folding of human nerve growth factor from Escherichia coli inclusion bodies.* European Journal of Biochemistry 268(11): 3296-3303.
Reid, B. G. and Flynn, G. C. (1997). *Chromophore formation in green fluorescent protein.* Biochemistry 36(22): 6786-6791.

Rizzo, M. A., Springer, G. H., et al. (2004). *An improved cyan fluorescent protein variant useful for FRET.* Nature biotechnology 22(4): 445-449.
Rodríguez-Carmona, E., Cano-Garrido, O., et al. (2012). *Recombinant Fab expression and secretion in Escherichia coli continuous culture at medium cell densities: Influence of temperature.* Process Biochemistry 47(3): 446-452.
Rosano, G. L. and Ceccarelli, E. A. (2014). *Recombinant protein expression in Escherichia coli: advances and challenges.* Frontiers in microbiology 5.
Russo, E. (2003). *Special Report: The birth of biotechnology.* Nature 421(6921): 456-457.
Sagmeister, P., Schimek, C., et al. (2014). *Tunable recombinant protein expression with E. coli in a mixed-feed environment.* Applied microbiology and biotechnology 98(7): 2937-2945.
Sandén, A. M., Prytz, I., et al. (2003). *Limiting factors in Escherichia coli fed-batch production of recombinant proteins.* Biotechnology and bioengineering 81(2): 158-166.
Santillán, M., Mackey, M., et al. (2007). *Origin of bistability in the lac operon.* Biophysical journal 92(11): 3830-3842.
Schlesinger, M. J. (1990). *Heat shock proteins.* J Biol Chem 265(21): 12111-12114.
Schumann, W. and Ferreira, L. C. S. (2004). *Production of recombinant proteins in Escherichia coli.* Genetics and Molecular Biology 27(3): 442-453.
Seeger, A., Schneppe, B., et al. (1995). *Comparison of temperature-and isopropyl-β-D-thiogalacto-pyranoside-induced synthesis of basic fibroblast growth factor in high-cell-density cultures of recombinant Escherichia coli.* Enzyme and microbial technology 17(10): 947-953.
Shaner, N. C., Campbell, R. E., et al. (2004). *Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein.* Nature biotechnology 22(12): 1567-1572.
Shiloach, J. and Fass, R. (2005). *Growing E. coli to high cell density—a historical perspective on method development.* Biotechnology advances 23(5): 345-357.
Shimomura, O. (1979). *Structure of the chromophore of Aequorea green fluorescent protein.* Febs Letters 104(2): 220-222.
Shimomura, O., Johnson, F. H., et al. (1962). *Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea.* Journal of cellular and comparative physiology 59(3): 223-239.
Singh, S. M. and Panda, A. K. (2005). *Solubilization and refolding of bacterial inclusion body proteins.* Journal of bioscience and bioengineering 99(4): 303-310.
Skerra, A. (1994). *Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in Escherichia coli.* Gene 151(1): 131-135.
Smith, T. (1892). *A new method for determining quantitatively the pollution of water by fecal bacteria.* 13th Annual Report of the State Board of Health of New York for: 712-722.
Sørensen, H. P. and Mortensen, K. K. (2005a). *Advanced genetic strategies for recombinant protein expression in Escherichia coli.* Journal of biotechnology 115(2): 113-128.
Sørensen, H. P. and Mortensen, K. K. (2005b). *Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli.* Microbial cell factories 4(1): 1.
Soriano, E., Borth, N., et al. (1999). *Flow cytometric analysis of metabolic stress effects due to recombinant plasmids and proteins in Escherichia coli production strains.* Metabolic engineering 1(3): 270-274.
Stepanenko, O. V., Verkhusha, V. V., et al. (2008). *Understanding the role of Arg96 in structure and stability of green fluorescent protein.* Proteins: Structure, Function, and Bioinformatics 73(3): 539-551.
Striedner, G., Cserjan-Puschmann, M., et al. (2003). *Tuning the Transcription Rate of Recombinant Protein in Strong Escherichiacoli Expression Systems through Repressor Titration.* Biotechnology progress 19(5): 1427-1432.
Studier, F. W. and Moffatt, B. A. (1986). *Use of bacteriophage T7 RNA polymerase to direct selective high-level expression of cloned genes.* Journal of molecular biology 189(1): 113-130.

Studier, F. W., Rosenberg, A. H., et al. (1990). *Use of T7 RNA polymerase to direct expression of cloned genes.* Methods in enzymology 185: 60-89.
Takahashi, A., Zhang, Y., et al. (2001). *Measurement of mitochondrial pH in situ.* Biotechniques 30(4): 804-815.
Tegel, H., Ottosson, J., et al. (2011). *Enhancing the protein production levels in Escherichia coli with a strong promoter.* FEBS journal 278(5): 729-739.
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(2): 211-222.
Tomosugi, W., Matsuda, T., et al. (2009). *An ultramarine fluorescent protein with increased photostability and pH insensitivity.* Nature methods 6(5): 351-353.
Trepod, C. and Mott, J. (2002). *A spontaneous runaway vector for production-scale expression of bovine somatotropin from Escherichia coli.* Applied microbiology and biotechnology 58(1): 84-88.
Tsien, R. Y. (1998). *The green fluorescent protein.* Annual review of biochemistry 67(1): 509-544.
Tsien, R. Y. (2009). *Constructing and exploiting the fluorescent protein paintbox (Nobel Lecture).* Angewandte Chemie International Edition 48(31): 5612-5626.
Turner, P., Holst, O., et al. (2005). *Optimized expression of soluble cyclomaltodextrinase of thermophilic origin in Escherichia coli by using a soluble fusion-tag and by tuning of inducer concentration.* Protein expression and purification 39(1): 54-60.
Valax, P. and Georgiou, G. (1993). *Molecular characterization of. beta.-lactamase inclusion bodies produced in Escherichia coli. 1. Composition.* Biotechnology progress 9(5): 539-547.
Van Hoek, M. and Hogeweg, P. (2006). *In silico evolved lac operons exhibit bistability for artificial inducers, but not for lactose.* Biophysical journal 91(8): 2833-2843.
Vera, A., González-Montalbán, N., et al. (2007). *The conformational quality of insoluble recombinant proteins is enhanced at low growth temperatures.* Biotechnology and bioengineering 96(6): 1101-1106.
Walsh, G. (2014). *Biopharmaceutical benchmarks 2014.* Nature biotechnology 32(10): 992-1000.
Ward, W. W. (2005). *Green fluorescent protein: Properties, applications, and protocols.* New York: Wiley. M. Chalfie and S. Kain: 39-66.
Ward, W. W. and Bokman, S. H. (1982). *Reversible denaturation of Aequorea green-fluorescent protein: physical separation and characterization of the renatured protein.* Biochemistry 21(19): 4535-4540.
Waugh, D. S. (2005). *Making the most of affinity tags.* Trends in biotechnology 23(6): 316-320.
Xu, J., Banerjee, A., et al. (2012). *Galactose can be an inducer for production of therapeutic proteins by auto-induction using E. coli BL21 strains.* Protein expression and purification 83(1): 30-36.
Yang, F. (1997). *The molecular structure of green fluorescent protein.*
Zaccolo, M., De Giorgi, F., et al. (2000). *A genetically encoded, fluorescent indicator for cyclic AMP in living cells.* Nature Cell Biology 2(1): 25-29.
Zacharias, D. A., Violin, J. D., et al. (2002). *Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells.* Science 296(5569): 913-916.
Zapata-Hommer, O. and Griesbeck, O. (2003). *Efficiently folding and circularly permuted variants of the Sapphire mutant of GFP.* BMC biotechnology 3(1): 5.
Zimmer, C. (2009). *Microcosm: E. coli and the new science of life*, Vintage Books USA.
Zimmer, M. (2002). *Green fluorescent protein (GFP): applications, structure, and related photophysical behavior.* Chemical reviews 102(3): 759-782.

Paper draft

Expression tuning via a glucose/lactose mixed feeding strategy in *Escherichia coli*

David Wurm[§], Julian Quehenberger[§], Sophia Ulonska, Tadej Bosilj, Christoph Herwig, Oliver Spadiut*

Vienna University of Technology, Institute of Chemical Engineering, Research Division Biochemical Engineering, Gumpendorfer Strasse 1a, 1060 Vienna, Austria

*Corresponding author
E-mail: oliver.spadiut@tuwien.ac.at

[§]These authors contributed equally to this work.

Abstract

Short production phases due to high metabolic burden and formation of insoluble inclusion bodies are known limitations for recombinant protein production in *Escherichia coli*. Up to now, these issues are mainly tackled by genetic engineering.

In this work the following issues are studied: (1) the transferability of the relationship between the specific uptake rates of glucose and lactose during their concomitant uptake is investigated. Therefore specific substrate uptake rates of two *E. coli* BL21(DE3) strains producing different target proteins under the control of the T7*lac* promotor system are compared during limiting glucose concentrations, while lactose was available in excess; (2) the tunability of the recombinant protein production rate and the ratio of soluble protein (SP) to inclusion bodies (IBs) are assessed; (3) the degree of induction is investigated on the cellular level.

Our findings show that the maximum specific lactose uptake rate is strongly influenced by the expressed target protein. Furthermore we find that lactose induction via a limiting mixed feeding strategy can increase the recombinant protein

production rate and the ratio of SP to IBs while differently induced subpopulation can be avoided. This meets the demands of pharmaceutical bioprocesses to efficiently yield high amounts of active target protein and to extend the production phase by reducing metabolic burden and enhancing cell fitness.

Keywords

Escherichia coli, lactose, mixed feed, expression tuning, cellular level, all-or-none induction, recombinant protein expression

Introduction

E. coli is the most popular prokaryotic host organism for the production of recombinant proteins due to fast growth kinetics, the possibility of high cell density cultures in inexpensive media, high final titers and well established genetics and metabolism (Terpe 2006; Baeshen et al. 2015). The large number of potent cloning techniques and expression systems and the ease of strain engineering greatly facilitate the development of bioprocesses for the production of heterologous proteins (Sørensen and Mortensen 2005a; Makino et al. 2011; Rosano and Ceccarelli 2014). The strain most commonly used for recombinant protein expression is *E. coli* BL21(DE3) in combination with the T7*lac* promoter system due to its outstandingly high yields of target protein (Baneyx 1999; Terpe 2006; Tegel et al. 2011).

Unfortunately high expression rates of target protein lead to reduced amounts of SP and promote formation of IBs (Baneyx and Mujacic 2004; Fahnert et al. 2004). When translation exceeds the folding capacities these inactive protein agglomerates are formed (Kiefhaber et al. 1991). Although some bioprocesses exploit the specific properties of IBs, solubilization and refolding are expensive, time and labor intensive, have to be optimized for each product and yields of final active protein are often low (Clark 2001). Therefore much effort is put in the development of strategies that enhance the solubility of the recombinant protein (Schein 1989; Sørensen and Mortensen 2005b).

Another drawback of high expression rates is that cells are subject to metabolic burden (Glick 1995; Kilikian et al. 2000; Lecina et al. 2013). Metabolic burden can be defined as the amount of resources (raw material and energy) that is withdrawn from the host's metabolism for maintenance and expression of the foreign DNA (Glick 1995). Since protein production is highly energy demanding, competition for energy

and capacity of the cellular machinery arises between expression of the recombinant protein and the synthesis of housekeeping proteins and maintenance metabolism (Neubauer et al. 2003). Negative effects caused by high metabolic burden resulting from overexpression of heterologous protein manifest in diverse ways. Reduced growth rates (Schmidt et al. 1999), decline of the fraction of dividing cells (Andersson et al. 1996), leakage of the outer membrane (Georgiou et al. 1988), breakdown of the synthesis of housekeeping proteins (Dong et al. 1995), enhanced proteolysis of heterologous protein (Ramirez and Bentley 1995; Rozkov and Enfors 2004), reduced substrate uptake (Neubauer et al. 2003) and consequently reduction of the synthesis rate of target protein and short production phases (Hoffmann and Rinas 2004) can be consequences of metabolic burden. As a result metabolic burden can severely limit the overall yield of target protein. A more moderate induction with the metabolizable inducer lactose instead of nonmetabolizable isopropyl β-D-1-thiogalactopyranoside (IPTG) or thiomethyl-β-D-galactoside (TMG) is a strategy that can potentially solve both the issues of target protein solubility (Neubauer et al. 1992; Gombert and Kilikian 1998) and reduced growth due to metabolic burden (Kilikian et al. 2000; Striedner et al. 2003; Pei et al. 2011; Zou et al. 2014; Fruchtl et al. 2015).

Another condition that counteracts full exploitation of the cell factory is the formation of distinct populations with different induction levels. Formation of differently induced subpopulations was observed for various inducible utilization systems (Afroz et al. 2014) and is especially well documented during induction of the *lac* operon or expression systems based thereon (Novick and Weiner 1957; Keasling 1999; Ozbudak et al. 2004; Afroz et al. 2014). This behavior is known as all-or-none induction, meaning that at the same time a culture consists of a fully induced population of cells (“all-“) and an uninduced population (“-none”). All-or-none induction should be avoided because it reduces the efficiency of a bioprocess in two ways: (1) the uninduced population consumes substrate without producing the target protein and (2) the highly induced population is subject to increased metabolic burden. A number of more recent studies report that the occurrence of all-or-none induction depends on the inducer type (Ozbudak et al. 2004; Van Hoek and Hogeweg 2006; Santillán et al. 2007; Dreisigmeyer et al. 2008; Afroz et al. 2014): Inducers like IPTG and TMG are called “gratuitous inducers” because they are capable of inducing the *lac* operon while not serving as substrates for the induced enzymes. These gratuitous inducers are reported to promote all-or-none induction,

while induction with lactose results in uniform induction. Predominantly this is explained with negative feedback control of the internal inducer concentration through lactose catabolism.

Short production phases due to high metabolic burden and reduced protein solubility are known limitations for the recombinant protein production in *E. coli*. As described above both issues are aggravated by high expression rates of target protein. In a previous study (Wurm et al. manuscript in preparation) we investigated the impact of the specific glucose uptake rate (q_{Glc}) on the specific lactose uptake rate (q_{Lac}) under nonlimiting lactose concentrations. We found a dependency of q_{Lac} on q_{Glc} and developed a mechanistic model capable of describing this linkage between the specific substrate uptake rates. This dependency furthermore impacts the expression of target protein as different amounts of inducer are taken up. In the present study we now carry out similar fermentations with an *E. coli* BL21(DE3) strain producing a different target protein, namely EGFP. By fitting the previously developed model (Wurm et al. manuscript in preparation) to the experimental data we (1) aim to investigate the transferability of the concomitant specific uptake rates of glucose and lactose between the two *E. coli* BL21(DE3) strains that are producing different target proteins. After decoupling q_{Lac} from $q_{Glc,}$ achieved via a mixed feed strategy with limiting glucose and lactose supply, we (2) aim to tune the recombinant protein production rate by controlling the specific inducer uptake. In addition to tuning the protein expression we hypothesize that it is possible to increase the ratio of SP to IBs and that the phase of high specific recombinant protein production can be extended. For all these experiments with varying q_{Lac} we chose a constant q_{Glc} that yielded the highest amounts of SP in our previous study (Wurm et al. manuscript in preparation). Finally (3) we investigate whether our induction strategy leads to differently induced subpopulations (“all-or-none induction”), a condition that should be avoided due to the above mentioned deleterious effects on the efficiency of a bioprocess.

We present an induction strategy that utilizes a cheap and nontoxic inducer and that significantly improves protein solubility, while homogeneous induction of the culture is achieved throughout the whole production phase.

Material and Methods

Strain and plasmid

For this study an *E. coli* BL21(DE3) strain was used. The strain was transformed with a pET 21a(+) plasmid harboring the T7*lac* promoter and an ampicillin resistance as well as the gene for the enhanced green fluorescent protein (EGFP) which was used as model protein.

Media

A defined minimal medium according to DeLisa et al. (1999) was used (Table 4). The preculture medium contained 8 g/L glucose and the batch medium 20 g/L. Glucose and lactose concentrations of the feed media varied depending on the experiment. The lactose concentration of the pulse medium was 200 g/L.

Table 4: Composition of DeLisa media.

Component	Preculture/batch medium (per liter)	(Mixed) feed media (per liter)	Lactose pulse medium (per liter)
Glucose	8.0/20.0 g	200 to 350 g	-
Lactose	-	31 to 96 g	200 g
KH_2PO_4	13.3 g	-	-
$(NH_4)_2HPO_4$	4.0 g	-	-
Citric acid	1.70 g	-	-
Antifoam PPG	-/100 mg	-	-
$MgSO_4 \cdot 7\ H_2O$	1.2 g	12.5 g	9.55 g
Fe(III) citrate	100 mg	25.0 mg	19.1 mg
EDTA	8.4 mg	8.1 mg	6.2 mg
$Zn(CH_3COO)_2 \cdot 2\ H_2O$	13.0 mg	10.0 mg	7.6 mg
$CoCl_2 \cdot 6\ H_2O$	12.5 µg	12.5 µg	12.5 µg
$MnCl_2 \cdot 4\ H_2O$	75.0 µg	75.0 µg	75.0 µg
$CuCl_2 \cdot 2\ H_2O$	6.0 µg	6.0 µg	6.0 µg
H_3BO_3	15.0 µg	15.0 µg	15.0 µg
$Na_2MoO_4 \cdot 2\ H_2O$	12.5 µg	12.5 µg	12.5 µg
Thiamine HCl	4.5 mg	-	-
Ampicillin	100 mg	-	-

Cultivation

Fermentations consisted of a batch phase followed by an uninduced fed batch phase (both on glucose). Subsequently the induction phase was initiated via an exponential feed containing glucose and lactose.

Preculture

Precultures were performed overnight (20 h) in 2,500 mL shake flasks (250 rpm, 37 °C). 500 mL of DeLisa preculture medium were inoculated from frozen stocks (1.5 mL, -80 °C).

Bioreactor setup and fermentation parameters

Fermentations were carried out in a 15 L stainless steel Satorius Biostat Cplus bioreactor (Satorius, Göttingen, Germany). 4.5 L batch medium were inoculated with 500 mL preculture. Batch phase was performed at 35 °C and the following fed batch and induced fed batch phases were performed at 30 °C for enhanced protein solubility. The fermentations were either induced via a limiting mixed feed containing glucose and lactose or with lactose pulses during limiting glucose feeding to maintain an excess of lactose. Dissolved oxygen (dO_2) was measured with a fluorescence dissolved oxygen electrode Visiferm DO425 (Hamilton, Reno, NV, USA) and dO_2 levels were maintained abouve 30% by aerating with 7.5 L/min pressurized air. Air was substituted with pure oxygen if necessary. pH was monitored with an Easyferm electrode (Hamilton, Reno, NV, USA) and kept at 7.2 by addition of NH_4OH (12.5%v/v) via the pump module of the bioreactor. Feed was supplied by a PRECIFLOW peristaltic pump (Lambda, Zurich, Switzerland). Mixing was performed at 1,400 rpm. CO_2 and O_2 content in the offgas were analyzed with a DASGIP GA, (Eppendorf, Hamburg, Germany). Process parameters were adjusted and recorded via a process information management system (PIMS; Lucullus, Biospectra, Switzerland).

Feed strategy

Until induction all fermentations followed the same scheme. The batch phase yielded 9 g/L dry cell weight (DCW) and was continued with an exponential fed batch until approximately 27 g/L DCW. During induction q_{Glc} was controlled via the feed rate which was calculated according to Equation 1 and Equation 2. The initial biomass concentration at the start of the feed was determined with optical density measurements (see below). In the experiments for the investigation of $q_{Lac,\ MAX}$ the culture was induced with an initial lactose pulse of 20 g/L (final concentration). During the induction phase the lactose concentration was determined at-line with HPLC and if necessary consecutive pulses were performed to maintain an excess of lactose of at least 3 g/L. In the experiments for the investigation of the tunability of the recombinant protein production rate lactose had to be limited and was supplied via a mixed feed containing both glucose and lactose. Lowering the q_{Lac} was achieved by decreasing the lactose to glucose ratio in the feed. The performed fermentations with their average specific substrate uptake rates during the induction phase and the goals of the experiment are listed in Table 5.

$$F_t = F_0 * e^{q_{Glc} * Y_{X/S} * t}$$

Equation 1: Calculation of the feed rate during fed batch

$$F_0 = \frac{x_0 * V_0 * q_{Glc}}{w}$$

Equation 2: Calculation of the initial feed rate at the beginning of the fed batch

F_0	initial feed rate [g_{Feed}/h]
F_t	feed rate t hours after start of the fed batch [g_{Feed}/h]
$Y_{X/S}$	biomass yield in DCW per substrate [g_x/g_{Glc}]
t	time after start of the fed batch [h]
q_{Glc}	specific glucose uptake rate [g_{Glc}/g_x/h]
x_0	initial biomass concentration in DCW per Volume [g_x/L]
V_0	initial reactor Volume [L]
w	glucose fraction in the feed medium [g_{Glc}/g_{Feed}]

Table 5: Performed experiments and average specific substrate uptake rates during the induction phase and the goals of the experiment. Specific substrate uptake rates in [g/g/h] of the experiments for the investigation of the concomitant uptake of glucose and lactose (fed batches 1 to 4) and for the investigation of the impact of specific lactose uptake on product formation and the ratio of SP to IBs (fed batches 4 to 7). If possible, depicted rates are mean values ± standard deviation of a series of calculated rates over a period of approximately 6 hours.

	q_{Glc}	q_{Lac}	% of $q_{Lac, MAX}$ [§]	Goal
Fed batch 1	0	0.023	-	Investigation of the transferability of the concomitant specific glucose and lactose uptake rates between *E. coli* BL21(DE3) expressing different products during lactose excess.
Fed batch 2	0.88 ± 0.05	0	-	
Fed batch 3	0.044 ± 0.004	0.21 ± 0.06	-	
Fed batch 4	0.24 ± 0.04	0.14 ± 0.04	-	
Fed batch 4	0.24 ± 0.04	0.14 ± 0.04	100	Investigation of the tunability of the recombinant protein expression rate and ratio of SP to IBs by controlling q_{Lac} via a mixed feed containing glucose and lactose.
Fed batch 5		0.075 ± 0.003	54	
Fed batch 6		0.020 ± 0.001	14	
Fed batch 7		0.0052 ± 0.0001	4	

[§] The value of the actual q_{Lac} divided by the value of $q_{Lac, MAX}$ at q_{Glc} = 0.24 g/g/h in per cent

Calculation of specific metabolic rates

All metabolic rates (specific substrate uptake rates and specific product formation rates) were calculated as mean values in between two sampling points. The difference of total substrate or product of two consecutive sampling points was divided by the arithmetic mean of the total biomass values of the two sampling points and the time between sampling (Equation 3).

$$q_A = \frac{A_n - A_{n-1}}{\bar{X} * \Delta t}$$

Equation 3: Calculation of a specific metabolic rate

q specific metabolic rate [$g/g_X/h$]
A substrate or product [g]
n sample number
$\overline{X}$ mean value of total biomass in DCW between two sampling points [g_X]
Δt time between two sampling points [h]

Sampling and analytics

Sampling

During the induction phase samples were taken every 1 to 1.5 hours and were centrifuged (8,600 rcf, 4°C, 10 min). Supernatant and cell pellets were collected for substrate and product quantification respectively and stored at -20 °C until analysis.

Biomass determination

5 mL fermentation broth were centrifuged (4,500 rcf, 4°C, 10 min) washed once with 0.9%(w/v) NaCl solution and subsequently dried at 95 °C for at least 72 h. Dry cell weight (DCW) was determined gravimetrically.

Optical density (OD) was determined photometrically at 600 nm. Samples were diluted with deionized water to stay in the linear range of the photometer (OD 0.2 to 0.8). The initial biomass concentration (x_0, Equation 2) was calculated via an established correlation (DCW = OD x 0.44; $R^2 = 0.99$).

Substrate and metabolite quantification

Glucose, lactose, galactose and acetate were quantified in the supernatant of the fermentation broth with HPLC (Agilent Technologies, Santa Clara, CA, USA) using an anion exchange column (Supelcogel C-610 H; Sigma-Aldrich, St. Louis, MO, USA) and a refractive index detector (Agilent Technologies, Santa Clara, CA, USA). Mobile phase was a 0.1%(v/v) H_3PO_4 solution with a constant flow rate of 0.5 mL/min at 30 °C.

Product quantification

Cell pellets of 5 mL fermentation broth were diluted (100 mM Tris, 10 mM EDTA pH 7.4) to 4.0 g/L DCW and resuspended. The resuspended samples were homogenized at 1,500 bar in 6 passages (EmulsiflexC3; Avestin, Ottawa, Canada). After centrifugation (15 min, 13,000 rcf, 4 °C) the supernatant was stored at -20 °C for quantification of the SP fraction. The remaining pellet was washed once (50 mM Tris, 0.5 M NaCl, 0.02%(w/v) Tween 80, pH 8), resuspended in 2 mL Tris-buffer (50 mM, 5mM EDTA, pH 8.0) and stored at -20 °C for quantification of IBs. 0.3 mL IB suspension were solubilized for 2 h on a shaker at room temperature in 1.2 mL

solubilization buffer (6 M guanidinium hydrochloride, 50 mM Tris, pH 8.0) with 5.0%(v/v) 2-mercaptoethanol added right before use. The samples were vortexed every 30 min. Product quantification was carried out with HPLC (UltiMate 3000; Thermo Fisher, Waltham, MA, USA) using a reversed phase column (EC 150/4.6 Nucleosil 300-5 C8; Macherey-Nagel, Düren, Germany). The product was quantified with an UV detector (Thermo Fisher, Waltham, MA, USA) at 280 nm. Mobile phase was composed of acetonitrile and water both with 0.1%(v/v) tetrafluoride acetic acid. A linear gradient from 30%(v/v) acetonitrile to 100% acetonitrile was applied. Bovine serum albumin was used as standard. The SP fraction was analyzed directly using the same HPLC method.

Fluorescence microscopy

To analyze the protein expression on a cellular level and to assess the formation of differently induced subpopulation fluorescence microscopy was performed. Fermentation broth was diluted 1:100 in 0.9%(w/v) saline solution. 20 µL diluted sample were stained on glass slides with 14 µmol/L RH414 dye, an unspecific membrane stain (maximum excitation/emission wavelengths (ex/em): 500/635 nm; tebu-bio, Le Perray-en-Yvelines, France). Images were acquired without prior incubation with a confocal laser scanning microscope (CLSM; C1 system with TE2000 inverted microscope; Nikon, Chiyoda, Japan). A 100x oil immersion objective was used and the gain of the photo multiplier tube was set to 65 Volt. EGFP ((ex/em: 488/508 nm) was exited with a blue Ar laser (488 nm) and emitted light was filtered with a 515±15 nm bandpass filter. Cells stained with RH414 were exited with a green He/Ne laser (543 nm) and emitted light was filtered with a 605±35 nm bandpass filter. For image processing, cell counting and quantification of the induction levels the software ImageJ (U. S. National Institutes of Health, Bethesda, MD, USA) was used.

Results and Discussion

In this study we (1) aimed at investigating the transferability of the concomitant specific uptake rates of glucose and lactose between two *E. coli* BL21(DE3) strains producing different target proteins under the control of the T7*lac* promotor system. Thereby we wanted to validate a previously developed model (Wurm et al. manuscript in preparation) by fitting the new data points. Furthermore we (2) aimed at tuning the recombinant protein production rate by controlling the specific inducer uptake. Here we hypothesized that it is possible to increase the ratio of SP to IBs and

that the phase of high specific recombinant protein production can be extended. Finally (3) we investigated if our induction strategy leads to all-or-none induction, a condition that should be avoided due to deleterious effects on the efficiency of a bioprocess as stated at the introductory section.

Concomitant uptake of glucose and lactose

Although the phenomenon of diauxic growth only occurs under nonlimiting substrate concentrations glucose nevertheless influences the uptake of lactose when both sugars are metabolized concomitantly. Prerequisite for concomitant uptake is that glucose is limited so the specific uptake rate is below the maximum metabolic capacity for sugar catabolism. The rest of this bottleneck can then be filled with lactose, the sugar with lower uptake priority. As we found out in our previous study (Wurm et al. manuscript in preparation), the bottleneck widens and narrows depending on the specific glucose uptake rate (q_{Glc}).

In previous experiments (Wurm et al. manuscript in preparation) with *E. coli* BL21(DE3) producing an antibody fragment, cloned into pET28a(+), we found out that the maximum specific lactose uptake ($q_{Lac,\ MAX}$) was dependent on the concomitant specific glucose uptake (Figure 9A). Via a simple modeling approach we were able to describe $q_{Lac,\ MAX}$ as a function of q_{Glc} as shown Equation 4. In the present work we conducted four similar fed batch experiments with EGFP as target protein to investigate the transferability of the concomitant uptake of glucose and lactose. To assess the applicability of the model as a platform for *E. coli* BL21(DE3) we tested the validity of the model by fitting our previously described model (Wurm et al. manuscript in preparation) to the new data points.

Upon reaching approximately 27 g/L DCW, cultures were supplied with glucose at a defined rate determined via Equation 1 and Equation 2 and lactose was supplied in excess. $q_{Lac,\ MAX}$ and the actual values for q_{Glc} were then determined via Equation 3. The actual specific substrate uptake rates are listed in Table 7 and plots of the concomitant sugar uptake depicted as q_{Lac} vs. q_{Glc} are shown in Figure 9A for the antibody fragment (previous study) and Figure 9B for EGFP (this study).

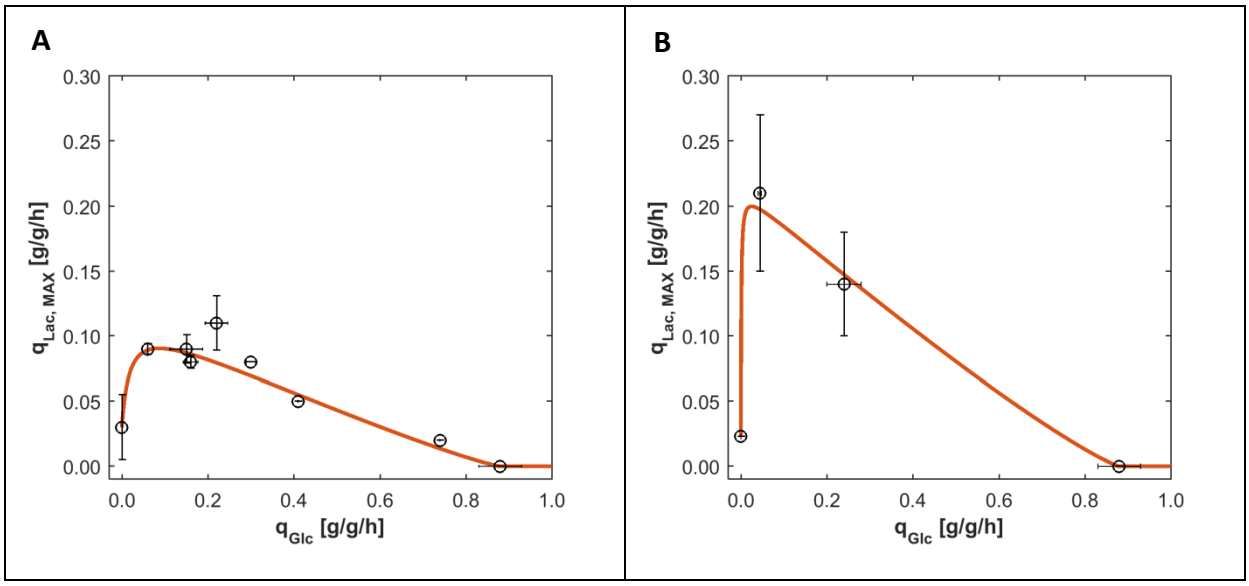


Figure 9: Maximum specific lactose uptake rate ($q_{Lac,\ MAX}$) plotted against specific glucose uptake rate (q_{Glc}). q_{Glc}-values were set as process parameters and lead to different values of $q_{Lac,\ MAX}$. As a result $q_{Lac,\ MAX}$ can be described as a function of q_{Glc}. While the trend of the resulting curve is the same for both products, the culture producing the antibody fragment (A) exhibits q_{Lac}-values approximately half as high compared to the culture that is producing EGFP (B).
The solid line represents the fit of the model shown in Equation 4 to the experimental data. For both products the parameters $q^*_{Lac,\ MAX}$, q^*_{Glc} and $q_{Lac,\ noGlc}$ were determined experimentally. The parameters K_m and n were identified in our previous study by minimizing an objective function (Wurm et al. manuscript in preparation). For the adaption of the model to the strain producing EGFP n was adopted from the previous study and only the value of K_m was optimized. n describes the type of inhibition which is not likely to change, while K_m describes the affinity of the cells to lactose which readily changes with the amount of lactose permease. Values of the used parameters for both model fits are listed in Table 6.

$$q_{Lac,\ MAX} = q^*_{Lac,\ MAX} \cdot \max\left(\left(1 - \frac{q_{Glc}}{q^*_{Glc}}\right)^n, 0\right) \cdot \left(\frac{q_{Glc}}{q_{Glc} + K_m} + \frac{q_{Lac,\ noGlc}}{q^*_{Lac,\ MAX}}\right)$$

Equation 4: Dependency of $q_{Lac,\ MAX}$ on q_{Glc}; adapted from Han and Levenspiel (1988). The model is described in detail by Wurm et al. (manuscript in preparation).

$q_{Lac,\ MAX}$	maximum specific lactose uptake rate
q_{Glc}	specific glucose uptake rate
$q^*_{Lac,\ MAX}$	highest maximum specific lactose uptake rate shifted by $q_{Lac,\ noGlc}$
q^*_{Glc}	critical specific glucose uptake rate, up to which lactose is consumed
$q_{Lac,\ noGlc}$	specific lactose uptake rate at $q_{Glc} = 0$
K_m	Monod constant for lactose uptake
n	type of inhibition (noncompetitive, uncompetitive, competitive) as described by Han and Levenspiel (1988)

Table 6: Parameter values used to fit our previously described model (Wurm et al. manuscript in preparation) to the experimental data as shown in Figure 9.

	$q^*_{Lac,\ MAX}$	q^*_{Glc}	$q_{Lac,\ noGlc}$	K_m	n
previous study: antibody fragment (Figure 9A)	0.088	0.88	0.034	0.019	1.16
this study: EGFP (Figure 9B)	0.191	0.88	0.023	0.0010	1.16

Table 7: Actual specific substrate uptake rates of experiments conducted for the investigation of the concomitant uptake of glucose and lactose. The specific glucose uptake rate (q_{Glc}) and the specific lactose uptake rate (q_{Lac}) for the time after the cultures reached approximately 27 g/L DCW are given in [g/g/h]. If possible, depicted rates are mean values ± standard deviation of a series of calculated rates over a period of approximately 6 hours.

	q_{Glc}	q_{Lac}
Fed batch 1	0	0.023
Fed batch 2	0.88 ± 0.05	0
Fed batch 3	0.044 ± 0.004	0.21 ± 0.06
Fed batch 4	0.24 ± 0.04	0.14 ± 0.04

During concomitant sugar uptake the culture producing EGFP exhibited a $q_{Lac, MAX}$ nearly twice as high compared to the culture that is producing the antibody fragment at the same q_{Glc}. In contrast, the endpoints (maximum specific glucose uptake and no glucose uptake, respectively) are very similar for both products. When glucose was available in excess, no lactose was taken up due to repression of the *lac* operon. As a consequence the cells were not induced and the specific sugar uptake was independent of the target protein. When no glucose was supplied, $q_{Lac, MAX}$ dropped to nearly zero as well. This indicates that for efficient initiation of the lactose uptake the cells need at least basal amounts of energy for the synthesis of lactose permease and beta-galactosidase in order to take up and metabolize lactose.

As shown in Figure 9, $q_{Lac, MAX}$ is strongly influenced by the expressed target protein. The high expression rate of heterologous proteins interferes with the host metabolism. Additionally, differences in the amino acid depletion pattern and protein solubility can activate different metabolic pathways and occupy cellular capacities to a varying degree (Hoffmann and Rinas 2004). Conceivably recombinant production of the antibody fragment resulted in the synthesis of fewer lactose permeases compared to when EGFP was produced, consequently leading to a reduced $q_{Lac, MAX}$. Nevertheless, when plotting q_{Lac} vs. q_{Glc} of the two strains producing different products they shared the same trend and therefore transferability of our model for the sugar uptake with adjustment of parameters (shown in Table 6) was possible. For the utilization of mixed feed systems it is important to determine the working range in which lactose is still taken up (and uptake is not inhibited by q_{Glc}) and glucose is not accumulating due to overload of the metabolic bottleneck. As we could show in this section, by using the described model this working range can be completely described with a total of three experiments determining the parameters $q^*_{Lac, MAX}$, q^*_{Glc} and $q_{Lac, noGlc}$.

Impact of q_{Lac} on product formation and the ratio of SP to IBs

In our previous study (Wurm et al. manuscript in preparation) we investigated the impact of q_{Glc} on q_{Lac} when lactose was available in excess. By changing q_{Glc} we could furthermore impact the formation of SP and IBs as different amounts of inducer are taken up. We identified an optimum regarding SP formation at a q_{Glc} of 0.25 g/g/h, which corresponds to a q_{Lac} of 0.08 g/g/h.

In this study we decoupled q_{Lac} from q_{Glc} and aim at tuning the recombinant protein production rate by controlling not only q_{Glc} but also q_{Lac} and therefore the specific inducer uptake. We further hypothesized that it is possible to increase the ratio of SP to IBs and that the phase of high specific recombinant protein production can be extended. We planned four different experiments at a q_{Glc} of 0.25 g/g/h and changed q_{Lac} for each fermentation (Table 8; fed batches 4-7).

Table 8: Actual specific substrate uptake rates of experiments conducted for the investigation of the impact of lactose uptake on product formation and the ratio of SP to IBs. The specific glucose uptake rate (q_{Glc}) and the specific lactose uptake rate (q_{Lac}) are given in [g/g/h] during the induction phase, after the cultures reached approximately 27 g/L DCW. Additionally the ratio of the amount of SP to IBs 6 hours after start of the induction is shown. Depicted rates are mean values of a series of calculated rates over a period of approximately 6 hours.

	q_{Glc}	q_{Lac}	% of $q_{Lac, MAX}$ [§]	Ratio SP to IBs
Fed batch 4	0.24 ± 0.04	0.14 ± 0.04	100	1.9
Fed batch 5		0.075 ± 0.003	54	3.6
Fed batch 6		0.020 ± 0.001	14	7.6
Fed batch 7		0.0052 ± 0.0001	4	7.9

[§] the value of the actual q_{Lac} divided the value of $q_{Lac, MAX}$ at q_{Glc} = 0.24 g/g/h in per cent

Figure 10 shows the amount of SP and IBs per DCW. Cells induced with $q_{Lac, MAX}$ yielded the highest amount of IBs per DCW whereas the formation of SP could not match these high amounts. Reduction of q_{Lac} to 54% and 14% of $q_{Lac, MAX}$, respectively, led to an increase in the amount of SP (Figure 10A) while the formation of IBs was delayed (Figure 10B) and overall reduced. In the case of q_{Lac} = 4% of $q_{Lac, MAX}$ neither high quantities of SP nor IBs were formed.

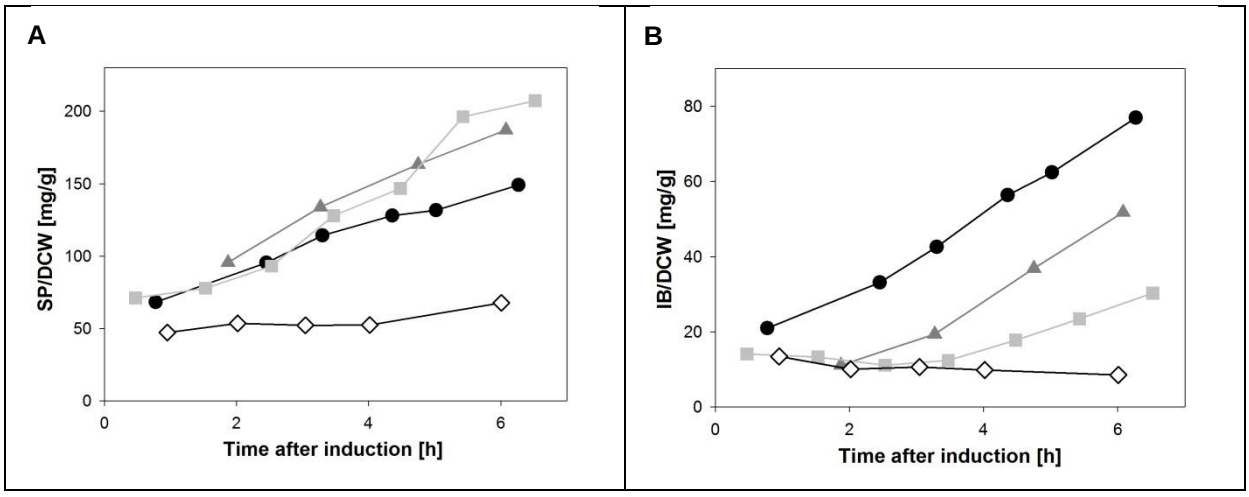


Figure 10: Ratio of soluble protein (SP; Figure A) and inclusion bodies (IB; Figure B) to DCW in mg/g. Fed batch 4, $q_{Lac, MAX}$ (●); Fed batch 5, q_{Lac} = 54% of $q_{Lac, MAX}$ (▲); Fed batch 6, q_{Lac} = 14% of $q_{Lac, MAX}$ (■); Fed batch 7, q_{Lac} = 4% of $q_{Lac, MAX}$ (◇). While the ratio of SP to DCW increases upon reduction of q_{Lac} until 14% of $q_{Lac, MAX}$, the formation of IB decreases. A q_{Lac} of 4% of $q_{Lac, MAX}$ was not sufficient to significantly induce the culture and neither high amounts of SP nor IBs were produced.

One would assume that a culture supplied with lactose at a rate of $q_{Lac, MAX}$ is strongly induced. Nevertheless the measured amount of SP was lower and the amount of total target protein was nearly the same as compared to the fermentations induced with lower q_{Lac}. High metabolic burden due to stronger induction can give an explanation for this phenomenon: Ample examples for product degradation by activation of proteolysis caused by increased heterologous protein expression can be found in the literature (Maurizi et al. 1985; Kitano et al. 1987; Kosinski et al. 1992; Ramirez and Bentley 1995). We conclude that in the case of higher q_{Lac}, produced target protein is only stable as IBs whereas target protein as SP is subject to proteolysis, resulting in low amounts of SP and high amounts of IBs (Ramirez and Bentley 1995; Harcum and Bentley 1999; Carrió et al. 2000; Rozkov et al. 2000; Fahnert et al. 2004). Reduction of the induction level due to lower q_{Lac} resulted in lower metabolic burden and therefore the effective amount of SP was higher and less protein was deposited as IBs. In the case of q_{Lac} = 4% of $q_{Lac, MAX}$ the amount of supplied lactose was not high enough to induce the expression of the target protein.

Figure 11 shows that within a range of q_{Lac} from 100% to 14% of $q_{Lac, MAX}$ the specific production rates of SP and IBs are inversely correlated to each other. As mentioned above q_{Lac} = 4% of $q_{Lac, MAX}$ was not sufficient to induce the culture and both the specific production rate of SP and IBs remained very low.

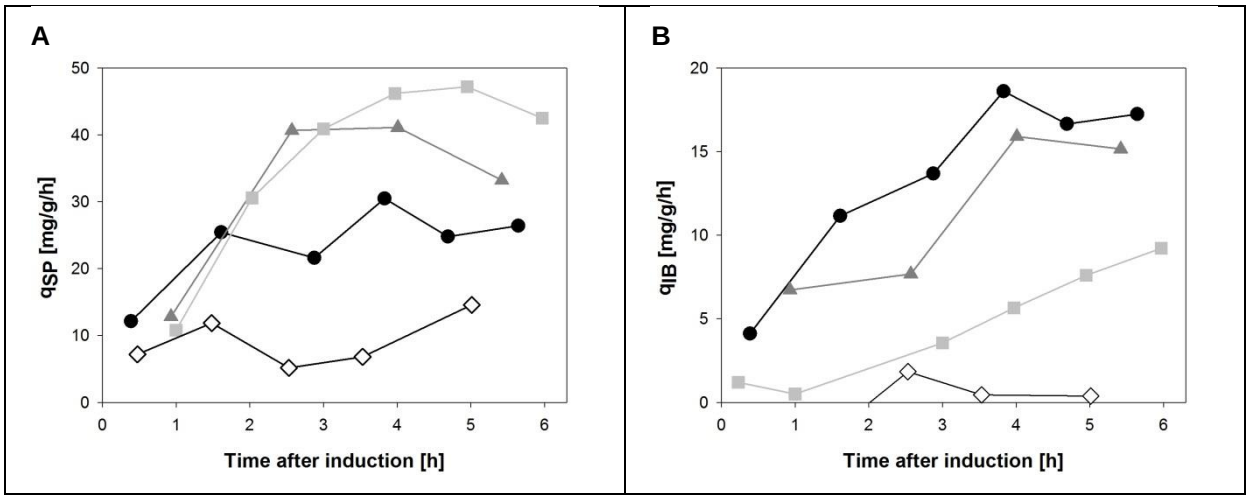


Figure 11: Specific rates of recombinant protein production of fermentations induced by applying different specific lactose uptake rates. The specific production rate of soluble EGFP (q_{SP}) is shown in (A) and the specific production rate of EGFP in IBs (q_{IB}) is shown in (B). Applied specific lactose uptake rates: Fed batch 4, $q_{Lac, MAX}$ (●); Fed batch 5, q_{Lac} = 54% of $q_{Lac, MAX}$ (▲); Fed batch 6, q_{Lac} = 14% of $q_{Lac, MAX}$ (■); Fed batch 7, q_{Lac} = 4% of $q_{Lac, MAX}$ (◇).

It can be inferred from Figure 11 that the production of SP can be maximized by choosing the optimal q_{Lac} for the induction of the culture. Figure 12 illustrates the location of this optimum in a surface plot. This means, that we have found a viable operating for maximum SP production.

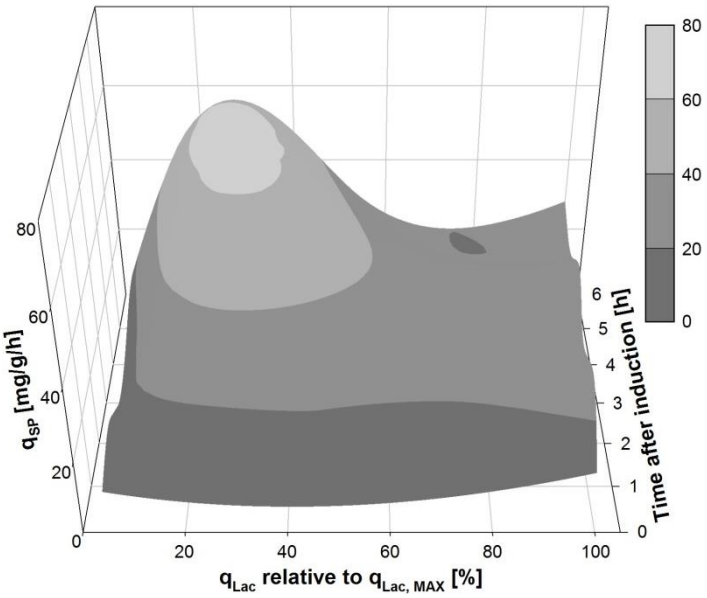


Figure 12: Surface plot showing the dependency of the specific production rate of soluble protein (q_{SP}) on the specific lactose uptake rate (q_{Lac}) during the induction phase. Protein synthesis is delayed approximately 1 hour (purple area at the bottom of the plot) which corresponds to the adaption time of the expression machinery. After that delay, highest specific lactose uptake (q_{Lac} = $q_{Lac, MAX}$) does not lead to highest q_{SP}, although then a maximum amount of inducer is to be expected inside the cell. Rather relatively low q_{Lac}-values correlate with the highest q_{SP}. It is important to find the optimal q_{Lac}, since too low values lead to a rapid decrease of q_{SP}.

Our results indicate that lower induction levels can result in higher effective productivity and increase the ratio of SP to IBs. We explain this with a reduction of metabolic burden combined with reduced proteolytic activity as explained above.

We could further show that by changing q_{Lac} it is possible to tune and optimize the specific recombinant protein production rate and the ratio of SP to IBs. Decreasing q_{Lac} increases the ratio of SP to IBs by two means, namely (1) by increasing the effective yield of SP, which, in the case of EGFP, holds true until a certain threshold between 14 and 4% of $q_{Lac, MAX}$, and (2) by delaying the formation of IB (Figure 10).

Subpopulations during lactose induction

The topic of all-or-none induction of the *lac* operon has been extensively investigated in the literature with different results depending on the type of the inducer. Induction with nonmetabolizable IPTG or TMG was reported to yield all-or-none responses (Novick and Weiner 1957; Ozbudak et al. 2004), while Afroz et al. (2014) state that lactose uptake in *E. coli* K-12 results in no formation of distinct subpopulations. Anyway, results in the literature have to be assessed with care because rarely time resolved data are available and results are based on end point measurements with induction times of up to 20 hours (Ozbudak et al. 2004; Afroz et al. 2014).

To ascertain that our induction strategy is capable of tuning the recombinant protein production rate on a cellular level, we investigated the occurrence of subpopulations during the induction phase. Tuning on a cellular level is a prerequisite for an efficient bioprocess since otherwise substrate is lost to the uninduced population not producing the target protein and the induced population is subject to increased metabolic burden. Time resolved fluorescence data were acquired with CLSM. The fluorescence of individual cells was calculated as integrated density and the results are depicted in histograms showing the frequency distribution of fluorescence intensities over the course of the induction phase (Figure 13).

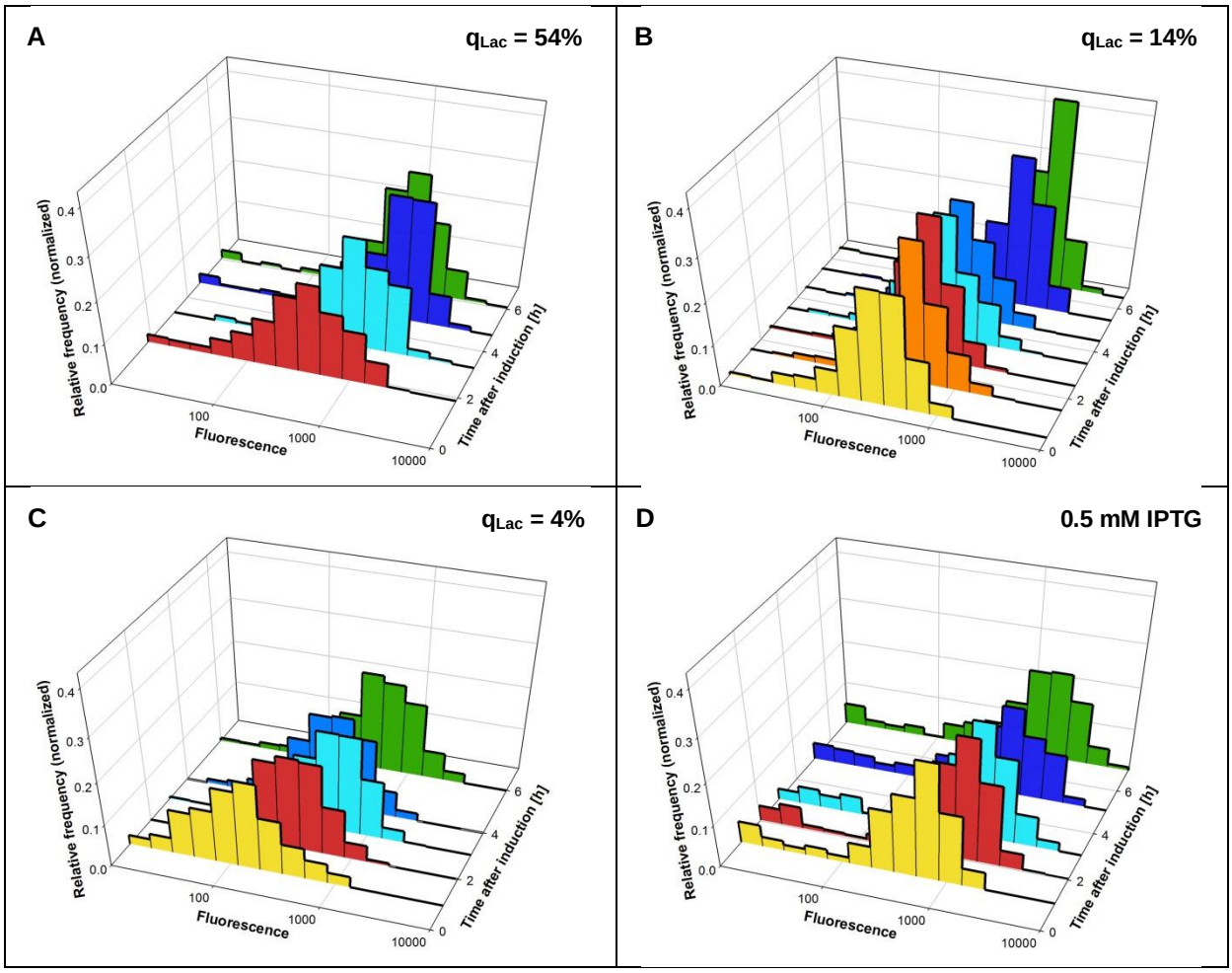


Figure 13: Frequency distributions of fluorescence intensities of fermentations induced by different specific lactose uptake rates (q_{Lac}) or with IPTG. A: q_{Lac} = 54% of $q_{Lac, MAX}$ (fed batch 5), B: q_{Lac} = 14% of $q_{Lac, MAX}$ (fed batch 6), and C: q_{Lac} = 4% of $q_{Lac, MAX}$ (fed batch 7); in these fermentations induction is completely unimodal and no subpopulations are present during the induction phase. D: induction with 0.5 mM IPTG, with the same fermentation parameters that were used for the lactose inductions; Here a bimodal induction pattern can be observed. About 10% of the cells remain uninduced even more than 6 hours after induction. Type and amount of inducer are indicated in the top right corner of each plot (q_{Lac} in % of $q_{Lac, MAX}$). Fluorescence values of individual cells are calculated from CLSM images as the fluorescent area per cell times the average fluorescence intensity measured in this area.

We did not assess to which extent EGFP deposited in IBs contributes to the fluorescence of the cells, although reportedly the specific fluorescence of IBs depends on process parameters like temperature and growth rate (Iafolla et al. 2008) and on the induction regime (Peternel et al. 2009). Nevertheless, when comparing the measured fluorescence intensities of the fermentations as shown in Table 8 with each other, the results were in accordance with the amount of EGFP per DCW measured with the (orthogonal) HPLC method: The culture induced with q_{Lac} = 14% of $q_{Lac, MAX}$ exhibits the highest fluorescence followed by the cultures induced with q_{Lac} = 54% of $q_{Lac, MAX}$ and q_{Lac} = 4% of $q_{Lac, MAX}$.

All fermentations induced with lactose resulted in unimodal fluorescence distributions, independent of the value of q_{Lac}. All-or-none induction in terms of two

distinguishable populations could not be observed. This implies the possibility of tuning the recombinant protein production on the cellular level. For comparison we performed a fermentation induced with 0.5 mM IPTG, applying the same fermentation parameters that were used for the lactose inductions (Figure 13D). Here we observed a bimodal induction pattern: About 10% of the cells maintained basal fluorescence levels more than 6 hours after induction. All-or-none induction of the *lac* operon has been subject to elaborate modeling efforts. During the induction with gratuitous inducers like IPTG and TMG these models predict the occurrence of all-or-none induction, while induction with metabolizable lactose is supposed to result in homogeneously induced cultures (Van Hoek and Hogeweg 2006; Santillán et al. 2007; Dreisigmeyer et al. 2008; Afroz et al. 2014). Generally the occurrence of subpopulations is explained with low catabolic activity paired with high inducer uptake like in the case of the gratuitous inducers. In the case of lactose, degradation of the actual inducer allolactose is strongly coupled with the induction level, since degradation is carried out by beta-galactosidase (LacZ) a gene product of the (induced) *lac* operon. Thereby it is not possible that individual cells overtake others in regard to their induction levels, what would allow them to outcompete the lower induced cells in the inducer uptake. In contrast to this great number of theoretical considerations experimental data are scarce. Ozbudak et al. performed a number of experiments with TMG (Ozbudak et al. 2004) and many of the before mentioned theoretical studies refer to these results. In the present study we provide further experimental data and by gathering time resolved data, possible time effects and behavior during adaption to lactose were monitored in detail. Our results are in accordance with the findings of the theoretical studies. We found no subpopulations during lactose induction, while IPTG yielded a small uninduced subpopulation of approximately 10% of all cells.

Conclusion

Recovery of active, soluble protein is crucial for the successful commercialization of a bioprocess for the production of heterologous protein.

The main findings of this study are:

- Our previously described model for the characterization of the concomitant utilization of glucose and lactose (Wurm et al. manuscript in preparation), is transferable between two *E. coli* BL21(DE3) strains producing different products

- Induction with a feed containing limiting amounts of glucose and lactose enables tuning of the recombinant protein expression rate
- Reducing the inducer uptake is a potent strategy to increase the ratio of SP to IBs without having to sacrifice high overall yields
- The developed induction strategy successfully prevents the formation of differently induced subpopulations

This work describes all necessary steps for maximizing the production of SP in a *lac* based expression system by process technological means, while utilizing the cheap and nontoxic inducer lactose: The described model can be used for the fast determination of a working range of physiologically feasible q_{Lac}-values. Within this range an optimal q_{Lac} can be found for maximal SP production. Providing the culture with a constantly limiting supply of lactose during a fed batch fermentation is simple to implement into an existing bioreactor setup and the strategy is readily transferable to different strains and products.

References

Afroz, T., Biliouris, K., et al. (2014). *Bacterial sugar utilization gives rise to distinct single-cell behaviours.* Molecular microbiology 93(6): 1093-1103.

Andersson, L., Yang, S., et al. (1996). *Impact of plasmid presence and induction on cellular responses in fed batch cultures of Escherichia coli.* Journal of biotechnology 46(3): 255-263.

Baeshen, M., Al-Hejin, A., et al. (2015). *Production of Biopharmaceuticals in E. coli: Current scenario and future perspectives.* Journal of microbiology and biotechnology.

Baneyx, F. (1999). *Recombinant protein expression in Escherichia coli.* Current opinion in biotechnology 10(5): 411-421.

Baneyx, F. and Mujacic, M. (2004). *Recombinant protein folding and misfolding in Escherichia coli.* Nature biotechnology 22(11): 1399-1408.

Carrió, M. M., Cubarsi, R., et al. (2000). *Fine architecture of bacterial inclusion bodies.* Febs Letters 471(1): 7-11.

Clark, E. D. B. (2001). *Protein refolding for industrial processes.* Current opinion in biotechnology 12(2): 202-207.

DeLisa, M. P., Li, J., et al. (1999). *Monitoring GFP-operon fusion protein expression during high cell density cultivation of Escherichia coli using an on-line optical sensor.* Biotechnology and bioengineering 65(1): 54-64.

Dong, H., Nilsson, L., et al. (1995). *Gratuitous overexpression of genes in Escherichia coli leads to growth inhibition and ribosome destruction.* Journal of bacteriology 177(6): 1497-1504.

Dreisigmeyer, D. W., Stajic, J., et al. (2008). *Determinants of bistability in induction of the Escherichia coli lac operon.* IET systems biology 2(5): 293-303.

Fahnert, B., Lilie, H., et al. (2004). *Inclusion bodies: formation and utilisation.* Physiological Stress Responses in Bioprocesses, Springer: 93-142.

Fruchtl, M., Sakon, J., et al. (2015). *Expression of a collagen-binding domain fusion protein: Effect of amino acid supplementation, inducer type, and culture conditions.* Biotechnology progress 31(2): 503-509.

Georgiou, G., Shuler, M. L., et al. (1988). *Release of periplasmic enzymes and other physiological effects of β-lactamase overproduction in Escherichia coli.* Biotechnology and bioengineering 32(6): 741-748.

Glick, B. R. (1995). *Metabolic load and heterologous gene expression.* Biotechnology advances 13(2): 247-261.
Gombert, A. and Kilikian, B. (1998). *Recombinant gene expression in Escherichia coli cultivation using lactose as inducer.* Journal of biotechnology 60(1): 47-54.
Han, K. and Levenspiel, O. (1988). *Extended Monod kinetics for substrate, product, and cell inhibition.* Biotechnology and bioengineering 32(4): 430-447.
Harcum, S. W. and Bentley, W. E. (1999). *Heat-shock and stringent responses have overlapping protease activity in Escherichia coli.* Applied biochemistry and biotechnology 80(1): 23-37.
Hoffmann, F. and Rinas, U. (2004). *Stress induced by recombinant protein production in Escherichia coli.* Physiological Stress Responses in Bioprocesses, Springer: 73-92.
Iafolla, M. A., Mazumder, M., et al. (2008). *Dark proteins: effect of inclusion body formation on quantification of protein expression.* Proteins: Structure, Function, and Bioinformatics 72(4): 1233-1242.
Keasling, J. D. (1999). *Gene-expression tools for the metabolic engineering of bacteria.* Trends in biotechnology 17(11): 452-460.
Kiefhaber, T., Rudolph, R., et al. (1991). *Protein aggregation in vitro and in vivo: a quantitative model of the kinetic competition between folding and aggregation.* Nature biotechnology 9(9): 825-829.
Kilikian, B., Suárez, I., et al. (2000). *Process strategies to improve heterologous protein production in Escherichia coli under lactose or IPTG induction.* Process Biochemistry 35(9): 1019-1025.
Kitano, K., Fujimoto, S., et al. (1987). *Intracellular degradation of recombinant proteins in relation to their location in Escherichia coli cells.* Journal of biotechnology 5(1): 77-86.
Kosinski, M. J., Rinas, U., et al. (1992). *Proteolytic response to the expression of an abnormal\-galactosidase in Escherichia coli.* Applied microbiology and biotechnology 37(3): 335-341.
Lecina, M., Sarró, E., et al. (2013). *IPTG limitation avoids metabolic burden and acetic acid accumulation in induced fed-batch cultures of Escherichia coli M15 under glucose limiting conditions.* Biochemical Engineering Journal 70: 78-83.
Makino, T., Skretas, G., et al. (2011). *Strain engineering for improved expression of recombinant proteins in bacteria.* Microb Cell Fact 10(32): 30.
Maurizi, M. R., Trisler, P., et al. (1985). *Insertional mutagenesis of the lon gene in Escherichia coli: lon is dispensable.* Journal of bacteriology 164(3): 1124-1135.
Neubauer, P., Hofmann, K., et al. (1992). *Maximizing the expression of a recombinant gene in Escherichia coli by manipulation of induction time using lactose as inducer.* Applied microbiology and biotechnology 36(6): 739-744.
Neubauer, P., Lin, H., et al. (2003). *Metabolic load of recombinant protein production: inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in Escherichia coli.* Biotechnology and bioengineering 83(1): 53-64.
Novick, A. and Weiner, M. (1957). *Enzyme induction as an all-or-none phenomenon.* Proceedings of the National Academy of Sciences of the United States of America 43(7): 553.
Ozbudak, E. M., Thattai, M., et al. (2004). *Multistability in the lactose utilization network of Escherichia coli.* Nature 427(6976): 737-740.
Pei, X.-L., Wang, Q.-Y., et al. (2011). *Efficient production of a thermophilic 2-deoxyribose-5-phosphate aldolase in glucose-limited fed-batch cultivations of Escherichia coli by continuous lactose induction strategy.* Applied biochemistry and biotechnology 165(2): 416-425.
Peternel, S., Gaberc-Porekar, V., et al. (2009). *Bacterial growth conditions affect quality of GFP expressed inside inclusion bodies.* Acta Chim Slov 56(4): 860-867.
Ramirez, D. M. and Bentley, W. E. (1995). *Fed-batch feeding and induction policies that improve foreign protein synthesis and stability by avoiding stress responses.* Biotechnology and bioengineering 47(5): 596-608.

Rosano, G. L. and Ceccarelli, E. A. (2014). *Recombinant protein expression in Escherichia coli: advances and challenges.* Frontiers in microbiology 5.

Rozkov, A. and Enfors, S.-O. (2004). *Analysis and control of proteolysis of recombinant proteins in Escherichia coli.* Physiological Stress Responses in Bioprocesses, Springer: 163-195.

Rozkov, A., Schweder, T., et al. (2000). *Dynamics of proteolysis and its influence on the accumulation of intracellular recombinant proteins.* Enzyme and microbial technology 27(10): 743-748.

Santillán, M., Mackey, M., et al. (2007). *Origin of bistability in the lac operon.* Biophysical journal 92(11): 3830-3842.

Schein, C. H. (1989). *Production of soluble recombinant proteins in bacteria.* Nature biotechnology 7(11): 1141-1149.

Schmidt, M., Viaplana, E., et al. (1999). *Secretion-dependent proteolysis of heterologous protein by recombinant Escherichia coli is connected to an increased activity of the energy-generating dissimilatory pathway.* Biotechnology and bioengineering 66(1): 61-67.

Sørensen, H. P. and Mortensen, K. K. (2005a). *Advanced genetic strategies for recombinant protein expression in Escherichia coli.* Journal of biotechnology 115(2): 113-128.

Sørensen, H. P. and Mortensen, K. K. (2005b). *Soluble expression of recombinant proteins in the cytoplasm of Escherichia coli.* Microbial cell factories 4(1): 1.

Striedner, G., Cserjan-Puschmann, M., et al. (2003). *Tuning the Transcription Rate of Recombinant Protein in Strong Escherichiacoli Expression Systems through Repressor Titration.* Biotechnology progress 19(5): 1427-1432.

Tegel, H., Ottosson, J., et al. (2011). *Enhancing the protein production levels in Escherichia coli with a strong promoter.* FEBS journal 278(5): 729-739.

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(2): 211-222.

Van Hoek, M. and Hogeweg, P. (2006). *In silico evolved lac operons exhibit bistability for artificial inducers, but not for lactose.* Biophysical journal 91(8): 2833-2843.

Zou, C., Duan, X., et al. (2014). *Enhanced extracellular production of recombinant Bacillus deramificans pullulanase in Escherichia coli through induction mode optimization and a glycine feeding strategy.* Bioresource technology 172: 174-179.

Conclusions

In this work I could show that it is possible to tune the recombinant protein production rate in *E. coli* by adjusting the lactose uptake via a mixed feeding strategy. Furthermore, reducing the inducer uptake has proven to be a potent strategy to increase the ratio of soluble protein to inclusion bodies without having to sacrifice high overall yields. According to single-cell fluorescence data, cultures were induced uniformly upon induction with lactose.

Up to now tuning of T7*lac* based expression systems on a cellular level has never been reported. Nevertheless this is the most popular expression system for *E. coli* and demand for tuning the protein expression in order to reduce metabolic burden and extend production phases is given. This Thesis fills this gap and provides a strategy to successfully achieve expression tuning on a cellular level.

In the following paragraphs the questions raised at the beginning of this Thesis will be answered.

1st Scientific question

Is the relationship between the concomitant specific substrate uptake rates of* E. coli *BL21(DE3) transferable between strains producing different target proteins?

The results of fermentations conducted with a *E. coli* BL21(DE3) producing EGFP were compared with preliminary results of *E. coli* BL21(DE3) producing an antibody fragment cloned into pET28a(+). Thereby I could show that at the same q_{Glc}, $q_{Lac, MAX}$ of the strain producing EGFP is roughly twice as high compared to the strain producing the antibody fragment.

Interestingly there were nearly no differences in the concomitant sugar uptake when both lactose and glucose were available in excess: The critical q_{Glc} at which the conjoint bottleneck of sugar catabolism was completely occupied alone by the favored sugar, glucose, was in both cases about 0.9 g/g/h.

The two strains exhibit similar kinetics of sugar uptake when uninduced, while there are severe differences in the concomitant sugar uptake when the strains are induced.

This stresses the influence of the expressed target protein on the host metabolism. This work shows that not only the yield and solubility differ between heterologous proteins due to their specific properties but also the metabolism of the host is severely influenced. The most direct way how expressed proteins can influence the host is via their amino acid composition (Gallant 1979; Cashel et al. 1996). Unfavorable amino acid depletion pattern can induce stress answers leading to decreased protein synthesis capacities (stringent response). In a more indirect way, cells that suffer from metabolic burden induced by strong expression systems are prone to higher proteolytic activity resulting in accelerated degradation of the overexpressed protein (Maurizi et al. 1985; Kitano et al. 1987; Kosinski et al. 1992; Ramirez and Bentley 1995).

Main finding: The specific lactose uptake rate during concomitant glucose uptake of the strain *E. coli* BL21(DE3) depends on the expressed recombinant protein, but can be described with the same mechanistic model.

2nd Scientific question

Are the protein production rate and the ratio of SP to IBs tunable by adjusting the specific lactose uptake rate?

In this Thesis I showed that tuning of the specific protein production rate (q_P) is possible. Highest q_{SP} could be achieved by reducing the supply of lactose until eventually lactose levels were too low, leaving cells uninduced. This enables for optimizing q_{Lac} to maximize the ratio of SP to IBs combined with high protein expression resulting in highest yields of SP.

The fact that a lower inducer concentration yielded higher total amounts of target protein and increased the amount of SP (i.e. active protein) is counterintuitive. Nevertheless similar results were reported by Pei et al. (2011).

Pei et al. suggest that residual lactose and/or galactose in the broth inhibited the protein production, or alternatively that internal “lactose seemed to partially repress the T7 promoter due to the glucose liberated by hydrolysis of lactose”. To argue against the influence of residual sugars it has to be mentioned that both in the experiments conducted by Pei et al. and in this work, the residual concentrations

during the production phases in question were well below sugar concentrations generally used in standard batch fermentations and are regarded as not deleterious for the growth of *E. coli*. Also the idea of glucose produced by lactose hydrolysis to repress the T7 promoter is questionable, especially during glucose based fed batches, when glucose is present regardless of lactose hydrolysis.

Alternatively, increased proteolytic activity at higher inducer levels can be responsible for the reduced effective protein production rates. High proteolysis rates during elevated heterologous protein expression are documented in the literature (Ramirez and Bentley 1995; Yang and Enfors 1995; Rozkov et al. 2000). Therefore it is possible that especially strong expression systems like the T7*lac* promoter system do not suffice the idea of direct proportionality of inducer to produced amount of target protein but rather can have the opposite effect. Since the effective recombinant protein production is influenced by the whole cellular metabolism the strict reduction and focus on the expression system, in not enough to explain the complexity of heterologous protein production.

The observed increase of IBs at higher q_{Lac} can be caused by different effects: If translation exceeds folding kinetics, accumulating unfolded proteins can be deposited of in IBs. As discussed in the Appendix in the chapter on “Online product quantification and correlating product data” this is not likely, since proper folding is a precondition for chromophore formation in GFP (see introduction) and IBs in the broth exhibited the same fluorescence per mass as SP. Another reason for the increased amount of IBs at higher q_{Lac} could be the increased stability against proteolytic activity of IBs (Fahnert et al. 2004). Proteases cannot attack proteins in IBs, what results in the turnover of SP, while the amount of IBs increases.

Main finding: Induction with a feed containing limiting amounts of glucose and lactose enables tuning of the recombinant protein expression rate and reduction of the inducer uptake is a potent strategy to increase the ratio of SP to IBs without having to sacrifice high overall yields.

3rd Scientific question

Does our induction strategy promote the phenomenon of all-or-none induction or is it possible to tune the expression on a cellular level?

Fluorescence microscopy revealed a unimodal induction of the cells when the lactose induction strategy was applied.

All-or-none induction in terms of two distinguishable populations was not observed. This implies a graded response and supports the possibility of tuning the protein production on the cellular level. Graded response during lactose induction when lactose is supplied in subsaturating amounts, is in accordance with Afroz et al. (2014).

Main finding: Induction with a limiting mixed feed containing glucose and lactose does not promote the formation of differently induced subpopulations.

Outlook

Production of a high value periplasmic product

To further establish the mixed glucose/lactose feeding strategy as platform technology we will switch to the production of a high value periplasmic product. When producing periplasmic proteins solubility and the rate of protein production are of utmost importance, since only soluble proteins can be transported across the inner cell membrane into the periplasm and accumulation in the cytoplasm has to be avoided. Therefore production of a periplasmic protein can serve as a practical test for the proposed tunability of the expression rate and enhanced protein solubility facilitated by the lactose induction strategy.

Investigating cell viability and metabolic burden by quantifying stress indicators

It was beyond the scope of this work to perform assays for the quantification of stress indicators. Nevertheless this is a key element to assess the success of the lactose induction strategy because one possible benefit is the reduction of metabolic burden resulting in prolonged production phases.

The lactose induction strategy will be continued with periplasmic horse radish peroxidase (HRP) as target protein. There one focus will be the monitoring of cell lysis. Via assessing the activity of HRP and alkaline phosphatase, a natural periplasmic protein (method described by Wurm et al., manuscript submitted), in the fermentation broth the integrity of the outer membrane will be assessed. DNA and beta-galactosidase (which both are originally present in the cytoplasm/nucleolus) will be determined in the fermentation broth for the indication of complete cell lysis.

Further options to assess cell viability and metabolic burden are flow cytometric measurements combined with appropriate dyes and quantification of guanosine 3′,5′-tetraphosphate (ppGpp) (Cserjan-Puschmann et al. 1999), an indicator of the stringent response.

Assessing the effective recombinant protein production rate

In general only the effective recombinant protein production rate ($q_{P,\ eff}$) can be assessed when protein contents are measured. Naturally, $q_{P,\ eff}$ is the difference of protein synthesis rate and protein degradation rate. Observed low $q_{P,\ eff}$ at high inducer supply possibly can be explained with high protein synthesis rates combined with intense protein degradation. Measuring the actual q_P in addition to $q_{P,\ eff}$ will give rise to the extent of protein degradation. And if present, a possible connection with induction levels could be assessed. Alternatively to determination of the actual q_P on the protein level like described by Rozkov and Enfors (2004) assessment of the mRNA via quantification of cDNA via qPCR is an option.

Investigation of the role of galactose in the induction of *lac* based expression systems

Some reports state that galactose is capable of inducing *lac* operon based systems (Mattanovich et al. 1998; Xu et al. 2012). In this work this mode of induction was not taken into account. Nevertheless product data acquired in this Thesis of a fermentation that was continued overnight at residual galactose concentrations after complete lactose uptake suggest that the culture remained induced and target protein was still produced in the absence of lactose. The potential of galactose as inducer can be easily assessed in simple shake flask experiments followed by fluorometric end point determination of EGFP. One implication of galactose induction would be that severe differences in the induction strength of *gal*$^-$ (like BL21(DE3)) and *gal*$^+$ strains were to be expected.

To avoid potentially negative effects of galactose accumulation during prolonged fermentations and to fully utilize lactose as a secondary carbon source *gal*$^+$ strains should preferentially be used for lactose induction strategies.

Appendix

Results beyond the scope of the paper draft

Online product quantification and correlating product data

In addition to product quantification with reversed phase high performance liquid chromatography (RP-HPLC) the culture fluorescence was monitored online with a GFP-probe (modified AF 44 S/R turbidity probe; Aquasant Messtechnik AG, Bubendorf, Switzerland).

Special interest was put on the detection range of the GFP-probe and whether the measured signal originates only from the SP fraction or from IBs as well.

In Figure 9 the signal of the GFP-probe is plotted against SP and IB concentrations (in g/L broth) determined with RP-HPLC. These plots show that the probe is capable of monitoring the total amount of EGFP in the fermentation broth. Both SP and IBs contribute to the detected signal. Also correcting the IB concentration with an "emission factor" that would account for fluorescence attenuation of the IBs did not result in an increased regression coefficient. This shows that EGFP IBs are highly active and the maturation of the chromophore is not hindered for proteins aggregated in IBs. Since the maturation of the chromophore strictly requires a correctly folded protein as explained in the Introduction section, this implies that these IBs are not misfolded protein deposits but agglomerations of active protein. This contradicts results of García-Fruitós et al. (2005) who report a reduction of the specific fluorescence of EGFP IBs to about 20%. Nevertheless the specific activity of EGFP IBs has been reported to depend on the growth conditions (Iafolla et al. 2008) and induction regime (Peternel et al. 2009). Therefore it is conceivable that in the case of induction with the lactose feed strategy IBs are formed by active protein.

The utilization of the probe is somewhat limited by its narrow detection range (0 to 65,000 AU). This resulted in signal saturation at 10 g/L EGFP in the broth, although considerably higher concentrations were achieved during the fermentations.

By adjusting the signal amplification there is still some potential to measure a broader concentration range of total EGFP. Alternatively if monitoring phases of lower product concentrations is not of interest, the amplification can be reduced even further and greater EGFP concentrations in the broth can be measured.

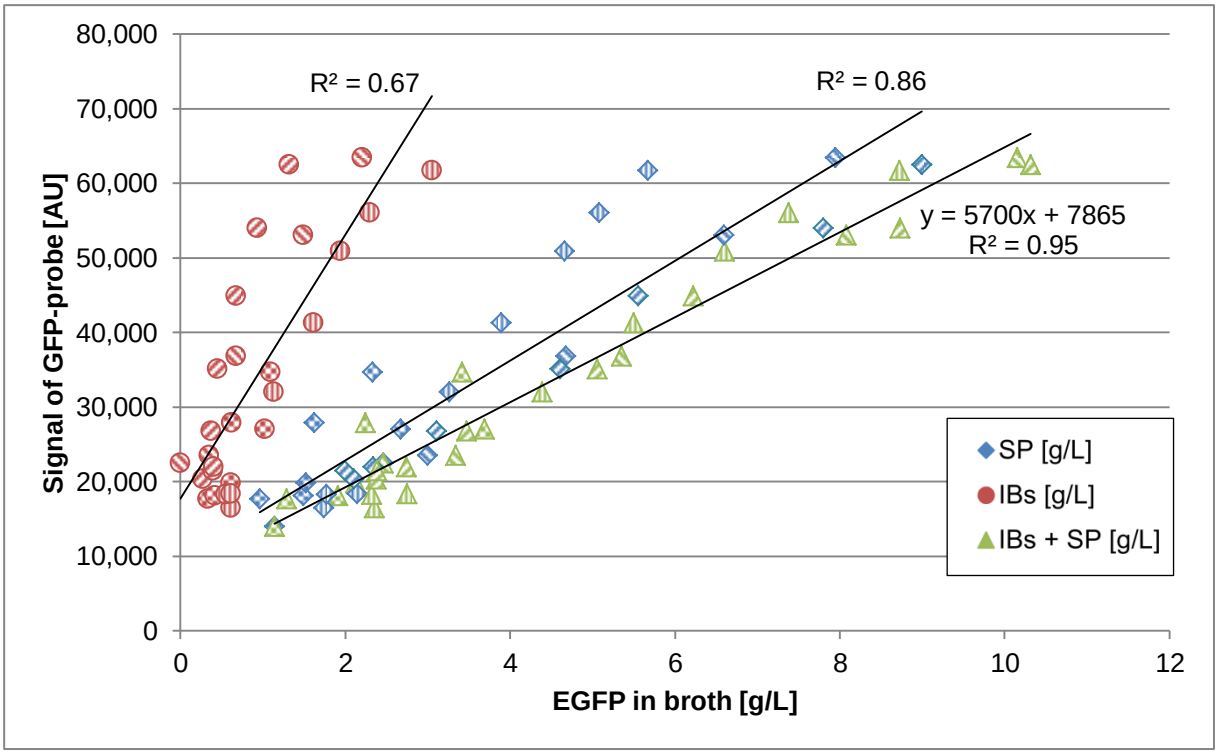


Figure 9: Correlating HPLC data with the signal of a probe for online GFP measurements. The graph shows that the signal of the GFP-probe linearly correlates with the total EGFP concentration in the broth (termed IBs + SP in the graph) over the whole detection range. Since the signal of the GFP-probe correlates best with the total EGFP concentration and not with the soluble fraction, this shows that not only soluble EGFP contributes to culture fluorescence, but also IBs exhibit fluorescence and are not inactive. Measurements belonging to the same fermentation are indicated with the same filling pattern.

Decreasing $q_{S,MAX}$ during the course of the induction phase

It could be shown that the specific substrate uptake in general (both glucose and lactose) was reduced several hours into the induction phase like reported by Neubauer et al. (2003). This can be attributed most probably to metabolic burden associated with heterologous protein production and accumulation of target protein inside the cell. The now following considerations are based on the substrate concentrations and rates depicted in Figure 10: Approximately 5 (Figure 10A), respectively 7 hours (Figure 10B) after induction this effect led to glucose accumulation during two fermentations. These fermentations were conducted with a mixed feed equivalent to a q_{Glc} of 0.24±0.04 g/g/h and two different levels of q_{Lac} (0.14 and 0.075 g/g/h, respectively). Interestingly during this period of excess glucose, lactose was still taken up even though q_{Lac} decreased slightly. This means, that although it is well established that high external glucose levels readily keep an uninduced *lac* operon in that uninduced state (carbon catabolite repression), once the operon is induced it cannot be switched off by external glucose. Possibly the numbers of LacY (lactose permease) were high enough, so that inducer exclusion via

interaction of dephosphorylated EIIA with LacY could only prevent lactose uptake to a small extent.

With the reduction of $q_{S, MAX}$, q_P decreased simultaneously. By induction with a lower initial q_{Lac} (0.020 g/g/h) accumulation of external glucose could be prevented and the decrease of q_P was significantly delayed. By conducting a fermentation induced with 0.5 mM IPTG which also led to glucose accumulation 7 hours after induction it could be shown, that accumulated external galactose from lactose hydrolysis was not the reason for the reduced glucose uptake, although both glucose and galactose share sugar transport systems like the PTS or none-PTS transporters like GalP (Kornberg and Riordan 1976; Postma and Roseman 1976). With these experiments it was possible to demonstrate the existence of a physiological time-effect that is initiated with the start of the induction phase and leads to a gradual of cellular productivity.

In conclusion, by reducing the specific inducer uptake rate (q_{Lac}) and therefore reducing the metabolic burden it was possible to extend the period of constantly high substrate uptake rates - this way circumventing the deleterious time-effect of decreasing maximum specific substrate uptake rates and productivity, and simultaneously gaining the potential for an extended production phase resulting in higher final titers and space time yields.

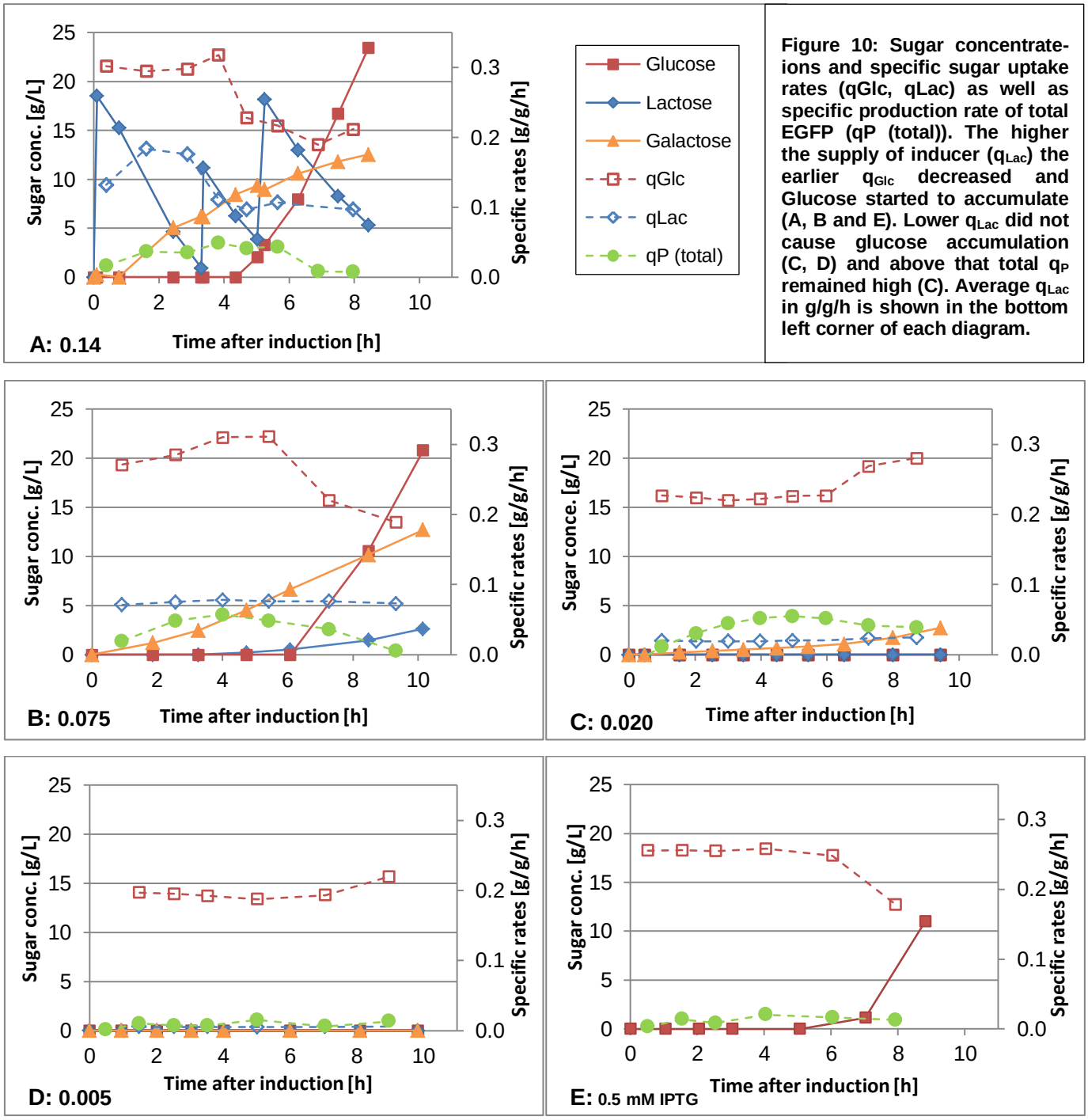


Figure 10: Sugar concentrate-ions and specific sugar uptake rates (qGlc, qLac) as well as specific production rate of total EGFP (qP (total)). The higher the supply of inducer (q_{Lac}) the earlier q_{Glc} decreased and Glucose started to accumulate (A, B and E). Lower q_{Lac} did not cause glucose accumulation (C, D) and above that total q_P remained high (C). Average q_{Lac} in g/g/h is shown in the bottom left corner of each diagram.

Indication of extracellular lactose degradation

To assess degradation of external lactose by extracellular beta-galactosidase a sample of supernatant taken 10 hours after induction was spiked with 0.6 g/L lactose. Sugar quantification with HPLC immediately after lactose addition and after 3 days at room temperature showed that lactose was degraded slowly (0.1 g/L/3 days) while the glucose and galactose concentrations increased. 19 days after lactose addition the concentration was at 0 g/L. Compared to the total lactose degradation that

ranged between 0.15 and 6 g/L/h the amount of external lactose degradation is insignificant. In succeeding fermentations with the lactose induction strategy, sensitive enzymatic beta-galactosidase assays will be applied to monitor cell lysis.

Problems encountered during this work and solutions

Arrested growth of cultures upon inoculation of the 15 L stainless steel Satorius Biostat Cplus bioreactor (Satorius, Göttingen, Germany)

Problem

After several successful fermentations, suddenly a series of fermentations died off upon inoculation of the bioreactor due to unknown cause. After ruling out problems with the preculture and media composition we found out, that the problem was caused by faulty pH measurements in the bioreactor despite successful calibration of the pH probe. Autoclavation induced a change in the pH probe resulting in measurement errors of 1.5 pH units. Because the medium naturally changes its pH during autoclavation due to solubilization of salts this shift in the pH measurements was not noticed at first. Only after measuring a sample with an external pH meter we realized this discrepancy. In earlier fermentation this was not an issue because a different pH probe was used and the error was smaller than 0.5 pH units.

Solution

To solve this problem we started to routinely measure the pH of a sample after autoclavation with an external pH meter. The offset of the calibration of the pH probe of the fermenter can then be adjusted via the console of the fermenter, so that the pH is measured correctly.

Low feed rates combined with high head pressure

Problem

Due to the high oxygen demand of high density cultures, fermentations were carried out at 0.5 bar head pressure to increase the solubility of oxygen. When using PRECIFLOW peristaltic pumps (Lambda, Zurich, Switzerland) for low feed rates, this counter pressure prevented feed medium from entering the fermenter. This is only an issue during the start of the fed batch phase, at low pump speed.

Solution

To solve this problem we diluted the feed, so higher feed rates and pump power were required to reach the desired substrate uptake. Drawback of this option is that at higher cell densities the required feed rate can exceed the pump power, requiring a change to a higher concentrated feed.

Reduction of the head pressure is also an option, but we utilized this solution only when we conducted fermentations at low growth rates. Otherwise the excessive use of oxygen for aeration would have been necessary.

Switching to a different pump would also be an option, but no suitable alternatives were available at the time of the fermentations.

Clogging and pressure issues using the homogenizer EmulsiflexC3 (Avestin, Ottawa, Canada)

Problem

The homogenizer repeatedly failed to generate the pressure necessary for cell disruption (~1,500 bar). This was due to clogging of the high pressure valve with viscose cell suspensions.

Solution

To clear the clogging the homogenizer is to be flushed with compressed air and consecutively rinsed thoroughly with ethanol 70%(v/v) and deionized water both with and without application of working pressure.

To avoid this problem we adapted the sample preparation and routinely introduced the following steps:

- Upon resuspension the biomass pellet is diluted to 4 g/L DCW instead of 5 g/L. This adaption may have to be reconsidered when investigating analytes of low concentration.
- The suspension is sonicated in an ultrasound bath for 1 minute.
- The suspension is treated with a lab mixer (Ultra Durrax; IKA, Staufen, Germany) to disrupt viscose agglomerations.

Standard operating procedures (SOPs)

SOP: Quantification of EGFP inclusion bodies and soluble protein via RP-HPLC
This SOP describes the process steps and settings for quantification of EGFP inclusion bodies (IBs) via RP-HPLC.

IB and soluble protein (SP) samples require sample preparation as described in the SOP “Homogenistation and Solubilisation of IBs” authored by Wieland Reichelt and adapted by Lukas Veiter, attached at the very end of this section.

Product quantification was carried out with HPLC (UltiMate 3000; Thermo Fisher, Waltham, MA, USA) using a reversed phase column (EC 150/4.6 Nucleosil 300-5 C8; Macherey-Nagel, Düren, Germany). The product was quantified with an UV detector (Thermo Fisher, Waltham, MA, USA) at 280 nm. Mobile phase was composed of acetonitrile and water both with 0.1%(v/v) tetrafluoride acetic acid. A linear gradient from 30%(v/v) acetonitrile to 100% acetonitrile was applied. Bovine serum albumin was used as standard. The soluble product fraction was analyzed directly using the same HPLC method.

Materials

- Solubilized IB samples or SP samples
- Bovine serum albumin (BSA)
- Solubilization buffer prepared according to SOP “Homogenistation and Solubilisation of IBs”
- Acetonitrile HPLC grade (ACN)
- MQ water
- Trifluoro acetic acid HPLC grade (TFA)
- HPLC vials, caps and septa
- syringe filters (0.2 µm pore size), syringes and hollow needles

Equipment

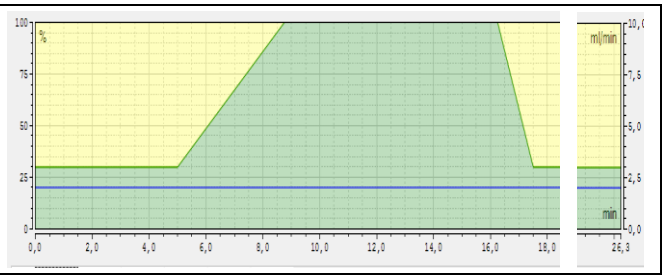
- 50 mL volumetric flask
- HPLC system
- RP-HPLC column EC 150/4.6 Nucleosil 300-5 C8 (Macherey-Nagel, Düren, Germany)

Procedure

- Preparation of Buffer A
    - MQ water + TFA (0.1%(v/v))
    - Sonicate for 15 minutes
- Preparation of Buffer B
    - ACN + TFA (0.1%(v/v))
    - Sonicate for 15 minutes
- Preparation of BSA standards
    - Dissolve 50 mg BSA in a 50 mL volumetric flask with deionized water

    - Prepare standards: 1.0 g/L, 0.75 g/L, 0.5 g/L, 0.4 g/L and 0.3 g/L by diluting with deionized water
    - Store at 4 °C
- Sample preparation
    - Dilute samples below 1.0 g/L EGFP if necessary
    - Samples must be filtered with syringe filters (0.2 µm)
    - Transfer into HPLC vials
    - Store at 4 °C
- HPLC settings
    - Flush the column for 30 min with 30% Buffer B
    - Flow rate: 2 mL/min
    - Column temperature: 30 °C
    - Upper pressure limit: 300 bar
    - Injection volume: 10 µL
    - Detection wavelength: 280 nm
    - After each sample and standard perform one blank run with solubilization buffer to remove residual protein off the column
- Elution profile

Time [min]	Buffer B [%]
0.00	30
5.00	30
8.75	100
16.25	100
17.50	30
26.25	30



- Retention times (in minutes)
    - BSA standard (can vary with supplier and production method): 6.95
    - EGFP (IBs): 7.10
    - EGFP (SP): 7.08
- Peak integration
    - Set the baseline manually according to the blank runs with solubilization buffer
    - Vertically split adjacent peaks
- Finishing the measurement
    - Flush the column for at least 30 minutes
    - Store the column in 40% Buffer A + 60% Buffer B
    - Remove the column and shut down the HPLC system according to the standard procedure

SOP: Imaging acquisition with the confocal laser scanning microscope system C1/TE2000

Note: Usage of the confocal microscope is only allowed after introduction by a qualified person. This SOP does not explain how to handle a confocal microscope.

This SOP describes imaging process of *Escherichia coli* cell producing the fluorescent marker EGFP with the confocal laser scanning microscope system C1/TE2000, but it can be consulted as a guideline for imaging in general.

Materials

- Cell suspension of *E. coli* cells of interest
- 0.9%(w/v) NaCl solution
- RH414 unspecific membrane stain; 14 µM diluted in DMSO or water
- Coverslips
- Glass slides (for inverted microscopy)

Equipment

- confocal laser scanning microscope system C1/TE2000 (Nikon, Chiyoda, Japan)
- 10 µL and 20 µL pipettes and tips
- Box with ice

Procedure

- Keep cell suspensions on ice during the procedure whenever possible
- Keep the fluorescent dye protected from light (use dark tubes for storage)
- Start the microscopy system according to the user manual. Imaging RH414 (Ex/Em: 500/635 nm) requires: green He/Ne laser (543 nm), 605±35 nm bandpass filter; Imaging EGFP (Ex/Em: 488/508 nm) requires: blue Ar laser (488 nm), 515±15 nm bandpass filter. 100x oil immersion objective is recommended.
- Dilute the cell suspension to 0.2 to 0.5 g DCW/L with NaCl solution
- Shake cell suspension thoroughly and apply 20 µL on a glass slide
- Add 1.5 µL RH414 dye, no mixing required
- Add the coverslip and put the specimen onto the microscope
- Select your detail of interest and focus on the cells using transmission microscopy
- Switch to confocal microscopy; fine-tune the focus in “live” mode. Keep the time that the specimen is illuminated with laser light at a minimum since both fluorescent dyes and fluorescent marker proteins suffer from photobleaching.
- Acquisition of single images and exemplary settings (optimize settings for individual problems):
    - Make sure that correct lasers and filters are activated
    - Gain: 60-65 Volt
    - Offset: 127 (Choose Gain and Offset to maximize the information content of your image by avoiding saturation of pixels. Also take into

account that fluorescence can increase over the course of a time series.)
    - Resolution: 1024x1024 (minimum resolution for publication)
    - Pixel dwell time: 1.92 µs (increasing dwell time enhances signals but increases photobleaching and slows down the imaging process)
    - Pinhole: small (larger pinholes enhances signals on cost of resolution)
    - Once settings are optimized do not change them between specimens that are to be compared with each other.
    - Select “single” mode
    - Save the image in ICS format (Image Cytometry Standard)
- Instead of single images it is also possible to acquire z-stacks (3D image):
    - Use settings like described for the acquisition of single images
    - Additionally select “acquire z-stack”
    - After focusing the cells select current plane as reference plane
    - Choose the height of the z-stack and thickness of slices. Recommended settings: 4.0 µm (z-stack), 0.25 µm (slice)
    - Select “single” mode
    - Save the image in ICS format (Image Cytometry Standard)
- After finishing your measurements
    - Shut down the microscopy system according to the user manual
    - clean the objectives
    - dispose of the glass slides in an appropriate (yellow) sharps container

SOP: Generating false color images from microscopy raw data using ImageJ

This SOP lists basic operations in the public domain image processing program ImageJ. With these operations raw microscopy data files can be converted into images suitable for presentation or publication.

Materials

- Raw image data in ICS format (Image Cytometry Standard), consisting of two separate files: a text header file (.ics extension) and a much bigger file with the actual image data (.ids extension). Acquired according to the SOP "Imaging acquisition with the confocal laser scanning microscope system C1/TE2000"

Equipment

- ImageJ 1.50e freeware program (U. S. National Institutes of Health, Bethesda, MD, USA)

Procedure

- Open the .ids file
- Select following "Bio-Formats Import Options":
  - "View stack with": "Standard ImageJ"
  - "Split channels"
  - Confirm with "OK"
- Acquired channels are displayed in separate windows:
  - C=0: confocal image, 450±20 nm bandpass filter, "blue" channel
  - C=1: confocal image, 515±15 nm bandpass filter, "green" channel
  - C=2: confocal image, 605±35 nm bandpass filter, "red" channel
  - C=3: transmission microscopy image, "white" channel
- Correctly display the images by selecting the header "Image" > "Type" > "32-bit"
- To assign a false color to a stack or slice select the button "Select LTU menu"  and choose your desired color scheme
- To browse through the slices of an image stack move the scroll bar at the bottom of a channel window
- To merge channels select the button "RGB Merge/Split Tool"  and choose the channels to be merged. Tick "Create composite" and "Keep source images"
- To generate a z-projection composed of multiple slice of an image stack or to generate an image of a single slice of an image stack choose the header "Image" > "Stacks" > "Z project…". Select the slices that are to be used for the z-projection and choose the appropriate "Projection type" ("Average" or "Maximum").
- To add a scale bar to your images choose the button "Scale Bar Tool"  and choose your desired settings

- To view metadata of your images either choose “Display metadata in results window” in the “Bio-Formats Import Options” when opening the image file or select the header “Image” > “Show Info…”
- To view pixel dimensions in (width and height in µm) select the header “Image” > “Properties…”
- To correctly display your processed images in standard picture viewers save your files in .png format

SOP: Quantification of cellular fluorescence signal using ImageJ (U. S. National Institutes of Health, Bethesda, MD, USA)

This SOP gives detailed instructions on how to quantify the fluorescence intensity of cellular markers. Intensities of individual cells can be measured in order to be visualized in a histogram showing the frequency distribution of fluorescence levels.

Materials

- Raw image data in ICS format (Image Cytometry Standard), consisting of two separate files: a text header file (.ics extension) and a much bigger file with the actual image data (.ids extension). Acquired according to the SOP "Imaging acquisition with the confocal laser scanning microscope system C1/TE2000"

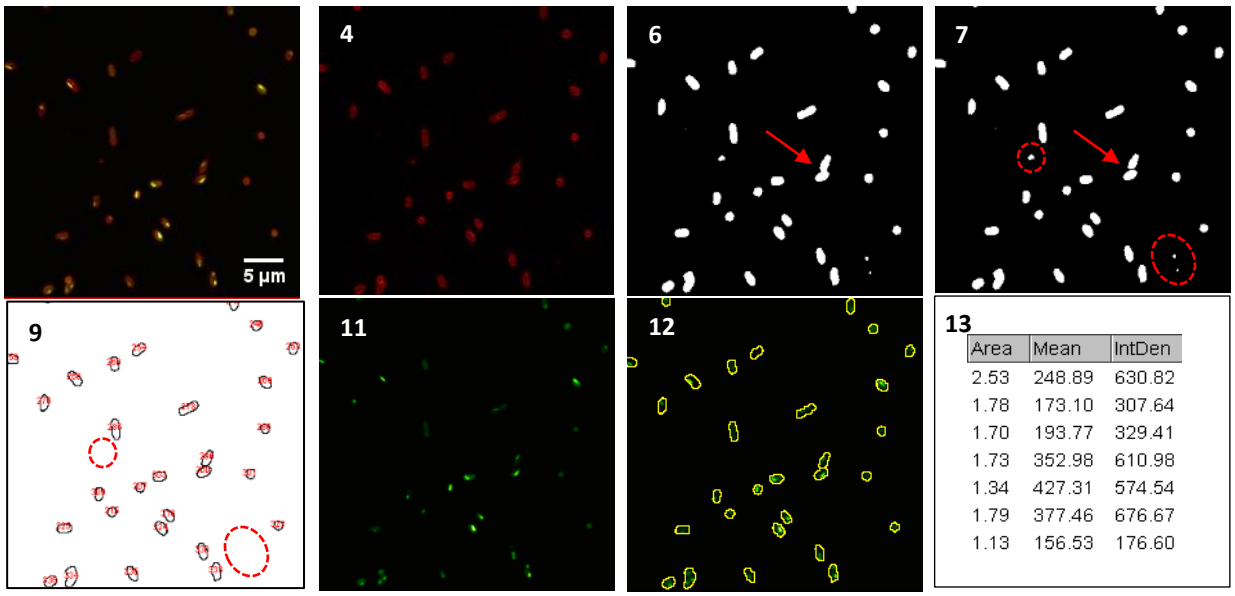
Equipment

- ImageJ 1.50e freeware program (U. S. National Institutes of Health, Bethesda, MD, USA)

Procedure

1. Open the .ids file
2. Select following "Bio-Formats Import Options":
    - View stack with: "Standard ImageJ"
    - "Split channels"
    - Confirm with "OK"
3. Select channel 2 (C=2, confocal image, 605±35 nm bandpass filter, "red" channel). This channel should picture the outline of your cells.
4. Correctly display the images by selecting the header "Image" > "Type" > "32-bit"
5. When investigating stacked images, select the slice that pictures your cells best
6. Select the header "Image" > "Adjust" > "Threshold…"
    - Adjust the threshold level so that individual cells grow to saturated white ovals while minimizing none-cellular artefacts.
    - Switching "Default" to "Huang" is sufficient most of the time, but additional fine-tuning with the scrollbars can be necessary.
    - Confirm your settings by pressing "Apply" then deselect "Set background pixels to NaN" and press "OK" twice
7. To separate cells that are sticking together select the header "Process" > "Binary" > "Watershed". This option can cause trouble, when cells are not completely white after adjusting the threshold. Then this step should be skipped.
8. Select the header "Analyze" > "Analyze Particles…"
    - Set "Size" to "0.5-Infinity" to ignore artefacts produced during imaging and image processing
    - Select "Add to Manager" and "Include Holes"
    - Press "OK"

9. ImageJ has now generated a mask on basis of the cells pictured in channel 1. The program has recorded position and area of every separated object in the selected slice.
10. Make sure that all relevant data are measured in the next steps by selecting the header “Analyze” > “Set Measurements…”
    - Select “Area”, “Mean gray value” and “Integrated density”
    - Press “OK”
11. Select channel 1 (C=1: confocal image, 515±15 nm bandpass filter, “green” channel). This channel should picture EGFP in the cells.
12. Select an entry in the ROI Manager
    - Select all entries by using the shortcut “Ctrl+A”
    - Press “Measure”
13. In the Results window “Area”, Mean” and “IntDen” of the EGFP fluorescence signal are now listed for each individual cell.
    - **“IntDen” is the product of “Area” times “Mean” and can be used to create frequency distributions of the culture´s fluorescence levels.**
    - The results can be exported into excel via copy & paste



Selected steps for the quantification of fluorescence of EGFP in *E. coli*. Top left: E. coli cells (red) expressed different amounts of EGFP. The clockwise following images show the same image section and represent the steps described under the number pictured in the top left corner of each image. Images 4, 11 and 12 are false colored for a better overview – this is not necessary for fluorescence quantification and not described in the SOP. The scale bar shown in the top left image also applies to all other images.

SOP: Homogenistation and Solubilisation of IBs by Wieland Reichelt, adapted by Lukas Veiter

TU VIENNA Technische Universität Wien Vienna University of Technology	**Standard Operating Procedure**	BioVT
	Title **Homogenistation and Solubilisation of IBs**	Page:83/86
Research Division ***Biochemical Engineering***	SOP Number: SOP Status: Effective/Editable	Date: 24.05.2015

Version	Number.0 for authorized versions; Number.1 for editions of a version, etc
Replaced version	
Author	Wieland Reichelt; Lukas Veiter
Date	
Signature	
Authorized by	
Date of authorization	
Signature	

Summary	Preparation solubilisation puffer

Materials

- Tris
- Guanidin Hydrochlorid **(98% for DSP/ 99,5 for HPLC analytics)**
- 2-Mercaptoethanol

Equipment

- Emulsiflex©-C3 HOMOGENIZER (read Operation of Emulsiflex©-C3 HOMOGENIZER before usage)
- Table top centrifuge
- Shaker
- pH meter

Procedure

<u>Resuspension of Cell Pellets</u>

i) Prepare lysis-buffer 100mM Tris (12,14 g/l), 10mM Na2EDTA (3,7 g/L) and adjust pH to 7,4 with HCl

ii) The pellet of 5 ml fermentation broth is resuspended on a scale in ice cold lysis buffer to a maximal final biomass concentration of 5 g/l (BDW). Due to the dead volume of 13 ml the minimal required volume for homogenisation is 20 ml.

Homogenisation and Washing of IBs

i) Prepare Buffer A: 50mM Tris, 0.5M NaCl, 0.02% Tween 80 (w/v) pH 8

ii) Prepare Buffer B: 50 mM Tris, 5mM EDTA, pH 8

iii) The cell suspension is homogenised at 1500 bar in 6 passages; it takes 13 ml 21 seconds for one passage (3min 30sec for 20ml)

iv) Rinse homogeniser after each sample with 30ml water.

v) Store sample on ice

vi) Take 15 mL of homogenate and centrifuge it at 13000 g for 15 minutes at 4 °C.

vii) Wash pellet with 15 mL buffer A and centrifuge again (13000 g, 15 min, 4°C)

viii) Resuspend pellet in __ ml buffer B, sample can now be stored at -20 °C

Solubilisation

i) Prepare solubilisation buffer 6M Guanidin Hydrochlorid (573,18g/l) and 50mM Tris (6,06 g/l) and adjust the pH to 8 (HCl) [(1)]

ii) Add 2-Mercaptoethanol (5 % volume) right before use.

iii) Take 300 µl of washed IB sample from step viii.

iv) Add 1200 µl of solubilisation buffer and vortex.

v) Put samples on shaker for 2h, vortex every 30 min.

vi) Centrifuge samples at 3000 rpm for 10 min.

vii) Samples are filled in HPLC vials through a syringe filter (0.2 µm)

Literature

- (1) 2010 H.F. Cueto-Rojas; „Interferon-α 2b quantification in inclusion bodies using Reversed Phase-Ultra Performance Liquid Chromatography (RP-UPLC)“

References

Afroz, T., Biliouris, K., et al. (2014). *Bacterial sugar utilization gives rise to distinct single-cell behaviours.* Molecular microbiology 93(6): 1093-1103.

Cashel, M., Gentry, D., et al. (1996). *The stringent response.* Escherichia coli and Salmonella typhimurium Cellular and Molecular Biology. F. Neidhardt. Washington, ASM Press: 1458–1496.

Cserjan-Puschmann, M., Kramer, W., et al. (1999). *Metabolic approaches for the optimisation of recombinant fermentation processes.* Applied microbiology and biotechnology 53(1): 43-50.

Fahnert, B., Lilie, H., et al. (2004). *Inclusion bodies: formation and utilisation.* Physiological Stress Responses in Bioprocesses, Springer: 93-142.

Gallant, J. A. (1979). *Stringent control in E. coli.* Annual review of genetics 13(1): 393-415.

García-Fruitós, E., González-Montalbán, N., et al. (2005). *Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins.* Microbial cell factories 4(1): 27.

Iafolla, M. A., Mazumder, M., et al. (2008). *Dark proteins: effect of inclusion body formation on quantification of protein expression.* Proteins: Structure, Function, and Bioinformatics 72(4): 1233-1242.

Kitano, K., Fujimoto, S., et al. (1987). *Intracellular degradation of recombinant proteins in relation to their location in Escherichia coli cells.* Journal of biotechnology 5(1): 77-86.

Kornberg, H. and Riordan, C. (1976). *Uptake of galactose into Escherichia coli by facilitated diffusion.* Journal of general microbiology 94(1): 75-89.

Kosinski, M. J., Rinas, U., et al. (1992). *Proteolytic response to the expression of an abnormal\-galactosidase in Escherichia coli.* Applied microbiology and biotechnology 37(3): 335-341.

Mattanovich, D., Kramer, W., et al. (1998). *Rational design of an improved induction scheme for recombinant Escherichia coli.* Biotechnology and bioengineering 58(2-3): 296-298.

Maurizi, M. R., Trisler, P., et al. (1985). *Insertional mutagenesis of the lon gene in Escherichia coli: lon is dispensable.* Journal of bacteriology 164(3): 1124-1135.

Neubauer, P., Lin, H., et al. (2003). *Metabolic load of recombinant protein production: inhibition of cellular capacities for glucose uptake and respiration after induction of a heterologous gene in Escherichia coli.* Biotechnology and bioengineering 83(1): 53-64.

Pei, X.-L., Wang, Q.-Y., et al. (2011). *Efficient production of a thermophilic 2-deoxyribose-5-phosphate aldolase in glucose-limited fed-batch cultivations of Escherichia coli by continuous lactose induction strategy.* Applied biochemistry and biotechnology 165(2): 416-425.

Peternel, S., Gaberc-Porekar, V., et al. (2009). *Bacterial growth conditions affect quality of GFP expressed inside inclusion bodies.* Acta Chim Slov 56(4): 860-867.

Postma, P. and Roseman, S. (1976). *The bacterial phosphoenolpyruvate: sugar phosphotransferase system.* Biochimica et Biophysica Acta (BBA)-Reviews on Biomembranes 457(3): 213-257.

Ramirez, D. M. and Bentley, W. E. (1995). *Fed-batch feeding and induction policies that improve foreign protein synthesis and stability by avoiding stress responses.* Biotechnology and bioengineering 47(5): 596-608.

Rozkov, A. and Enfors, S.-O. (2004). *Analysis and control of proteolysis of recombinant proteins in Escherichia coli.* Physiological Stress Responses in Bioprocesses, Springer: 163-195.

Rozkov, A., Schweder, T., et al. (2000). *Dynamics of proteolysis and its influence on the accumulation of intracellular recombinant proteins.* Enzyme and microbial technology 27(10): 743-748.

Xu, J., Banerjee, A., et al. (2012). *Galactose can be an inducer for production of therapeutic proteins by auto-induction using E. coli BL21 strains.* Protein expression and purification 83(1): 30-36.

Yang, S. and Enfors, S. O. (1995). *The influence of energy sources on the proteolysis of a recombinant staphylococcal protein A in Escherichia coli.* European Journal of Biochemistry 233(3): 969-975.