Unterschrift des Betreuers

TECHNISCHE UNIVERSITÄT WIEN
Vienna University of Technology

D I P L O M A R B E I T

Development of a Mixed-Feed Bioprocess for the Production of rhBMP-2 using an E.coli pBAD Expression System

Thema

Ausgeführt am Institut für

Chemical Engineering

der Technischen Universität Wien

unter der Anleitung von Prof. Dr. Christoph Herwig

durch

Magdalena Verena Kment, Bakk. techn.

Name

Fuchsthallergasse 15/7 1090 WIEN

Anschrift

06. Juni 2013 _____

Datum Unterschrift (Student)

Master Thesis

Production of human Bone Morphogenetic Protein-2 Upstream Design Space Development

Development of a Mixed-Feed Bioprocess for the Production of rhBMP-2 using an *E.coli* pBAD Expression System

Principal Investigator & Head of Laboratory:

Prof. Dr. Christoph Herwig

Supervisor:

DI Patrick Sagmeister

Master Student:

Magdalena Kment, Bakk. techn.

Faculty:

Institute of Chemical Engineering
Technical University of Vienna
Getreidemarkt 9, 1060 Wien

Associated Departments:

RCPE – Research Center Pharmaceutical Engineering
Inffeldgasse 21a/II, A-8010 Graz

BIRD-C Biotech Innovation Research Development & Consulting
Erne-Seder-Gasse 4/ Stiege 2/ Lokal 3
1030 Vienna, Austria

Morphoplant GmbH
Universitätsstrasse 136 ,44799 Bochum

06/06/2013

Table of Content

Abstract

Background: Recent regulatory initiatives encourage the use of Process Analytical Technologies (PAT) and Quality-by-Design (QbD) in order to gain a science- and risk-based understanding of bioprocesses. PAT and QbD are increasingly recognized as versatile means to characterize process parameters with respect to product quality.

Aims: In this thesis, the application of QbD and PAT is demonstrated on the upstream processing of the production of recombinant human bone morphogenetic protein-2 (rhBMP-2), a pharmacologically relevant inductor of osteogenesis. The primary objective of this work is to elucidate physiological process parameters using dynamic experimentation and design-of-experiments (DoE).

Material and Methods: For the fermentative production of rhBMP-2 a recombinant *E. coli* mixed-feed expression system under the control of the p_{BAD} promoter is used. For the mixed feed strategy, L-arabinose is employed as inducer of the recombinant construct and D-glucose is used as main C source. Pulse and ramp experimets are conducted to elucidate physiological boundaries of the rhBMP-2 production. Soft-sensors as well as in-line measurements derived from FTIR are implemented as PAT tools to monitor and control process conduct. Design-of-experiments (3-factor CCF design) is performed to analyze the contribution and interaction of physiological process parameters.

Results: Maximum metabolic conversion rates of L-arabinose were determined as a function of temperature and uptake of D-glucose using a novel dynamic method. With respect to critical quality attributes, IB purity and product titer, IB purity was found to be positively associated with $q_{s\ gluc}$ and the cultivation temperature whereas, product titer was only found to be positively associated with the temperature.

Conclusion: In summary, the use of soft-sensor assisted dynamic experimentation and physiological design of experiments is successfully shown to characterize the contribution of physiological process parameters on the upstream processing of rhBMP-2 production.

Introduction

Background & Motivation

Worldwide sales of biological drugs reached the US$100 billion mark in 2010 and it is expected that, by 2015, more than 50% of all new approved drugs will be of biological origin [1–3]. This immense commercial but also tremendous clinical success of biological drugs had a significant impact on the pharmaceutical industry and poses great challenges for the development and manufacturing of these highly complex bio-molecules. In order to acknowledge these requirements the FDA and the ICH issued several guidelines that clearly define the manufacturing demands of biological drugs. More specifically these guidelines introduced the concept of QbD, PAT (“ A system for designing, analyzing and controlling manufacturing through timely measurements (e.g. inline-FTIR, soft sensor) of critical quality and performance attributes of raw and in-process materials and processes with the goal of ensuring final product quality”[4]) and the design space which is a multidimensional combination and interaction of input factors and process parameters that have been demonstrated to provide assurance of quality (e.g statistically designed experiments (DoE)) [5], [6]. Over the past decades, bioprocess development and biochemical engineering shifted from a mere product-oriented, empirical, “know-how” attitude to a process-oriented, science-based, “know-why” attitude [7]. This shift resulted in the implementation of quality requirements, which can be summarized under the terms “Quality by Design” (QbD) and “Process Analytical Technology” (PAT) [8], [9]. Briefly, “Quality by design means designing and developing manufacturing processes during the product development stage to consistently ensure a predefined quality at the end of the manufacturing process” [8]. The key term in this definition is: “during the product development stage”; which means that quality attributes need to be identified, critically appraised and defined as early as possible during bioprocess development. Furthermore, this definition also demands a detailed scientific understanding and thorough knowledge on the physiological properties of the biological host/vector system under consideration as a necessary prerequisite for ensuring the successful development and subsequent manufacturing of high quality biologics.

Within the past decades, many recombinant proteins have been successfully introduced on the market as therapeutic agents. These biological drugs are produced using so-called host-vector systems; briefly the heterologous gene of interest (after insertion into plasmid DNA, yielding the vector) is introduced into a cellular system (host), which is then responsible for efficient protein synthesis and production. In order to optimize heterologous expression within the host cell sophisticated expression systems (promoter systems) are commonly utilized to overexpress the protein of interest. These expression systems consist for instance, of the well-known *lac*-promoter, the *trp*-promoter or the *ara*-promoter. *E. coli* is one of the most widely used host organisms. It is a gram-negative, rod-shaped bacterium that harbors the advantages a rapid growth rate and ease of high-cell-density fermentation, low cost and the availability of many different promoter systems [10]. However, *E. coli* also shows some disadvantages. Above all, most heterologous expressed proteins are primarily produced in form of inclusion bodies. Furthermore, the proteins of interest are only produced as non-glycosylated proteins, since *E. coli* does not have the necessary molecular equipment to conduct these eukaryotic post-translational modifications (PTMs). Inclusion bodies contain insoluble products, that results from product aggregation [11]. Since these aggregates contain protein products that do not show any biological activity, inclusion bodies require solubilization and refolding into the native 3D-structure and of the protein of interest [12]. This poses great challenges for subsequent downstream processing.

During upstream processing the main focus is set on characterization of the host/vector system and also on enhancing protein production rates. With respect to the characterization of the host/vector system it is of importance to elucidate the physiological boundaries of the system under consideration. Characterization of the system is important in order to deal with inhomogeneity's in large bioreactors, like process parameters or mixing gradients of substrates [6] or metabolites [13], which can cause difficulties in scaling up bioprocesses for commercial production purposes on the one hand [14]. On the other hand the physiological strain characterization by means of adaptation times or maximum metabolic rates is also of importance [15], [16]. To achieve this aim dynamic experimentation, which includes the application of ramp, pulse and shifts, is a promising strategy [14]. In order to achieve the

second goal of upstream processing, increasing protein production rates, the application of statistical modeling by means of design of experiments (DoE) has gained more and more acceptance recently. DoE offers the advantage that multiple process parameters are changed at a time, which results in a fewer number of experiments and can therefore safe costly resources [17]. In summary both methods, dynamic experimentation and DoE allow for a science-based process understanding, which is a necessary prerequisite according to currently applicable guidelines [4].

The primary motivation of this thesis is to demonstrate the applicability and feasibility of contemporary technologies, such as dynamic experimentation and DoE for the production of the therapeutically used protein recombinant human bone morphogenetic protein-2 (rhBMP-2).

Escherichia coli – Metabolism and its impact on recombinant protein production

Escherichia coli is one of the most important bacterial model organisms in molecular biology and biotechnology. The bacterium has a simple structure (see Figure 1), contains very few genes and is easy to cultivate in the laboratory. This last characteristic makes this host organism amenable for batch and also fed-batch fermentation.

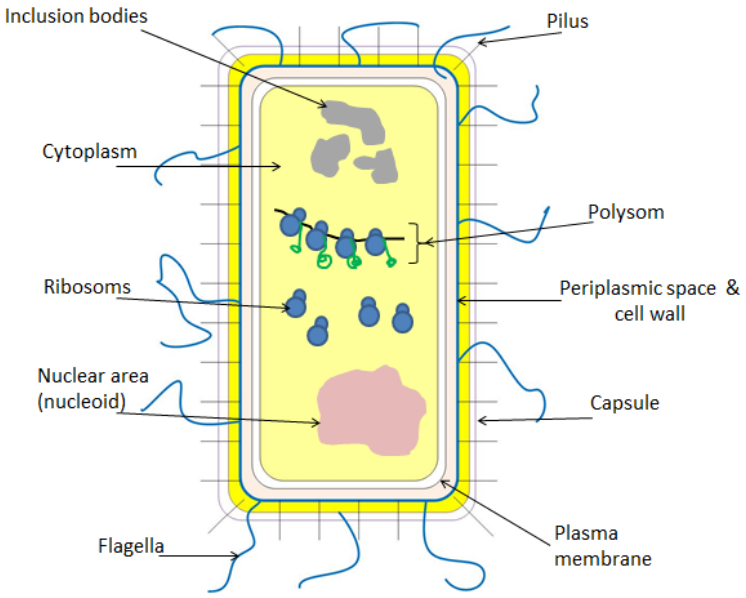

Figure 1: Structure of the gram-negative, rod-shaped bacterium *E. coli*. It is composed of two membranes and a peptidoglycan layer. The space between the membranes is called the periplasmic space. Furthermore, it must be noted that protein synthesis (transcription and translation) are spatially combined (compare to eukaryotes). Figure is adapted from Clark et al. [18] .

E. coli cells typically measure 3 µm in length and 0.5 µm in diameter [19]. However, the final cell volume highly depends on the availability of nutrients [19], [20].

In order to exploit *E. coli* as "protein-factory" in fermentation experiments, many highly sophisticated genetic tools have been developed that allow the introduction of foreign DNA into the host cell [21]. However, various metabolic characteristics of this organism need to be considered when using *E. coli* for heterologous gene expression. First of all, *E. coli* is able to metabolize hexoses (e.g. D-glucose) but also pentoses (e.g. L-arabinose). Glucose metabolism starts with its uptake by specific transporters and results in several

interconnected pathways with the major being: glycolysis, gluconeogenesis, the pentose-monophosphate, the TCA cycle, several anaplerotic reactions and finally acetate production and assimilation [21]. The central metabolite of the glycolysis and the TCA is Acetyl-CoA, which is the biologically active thioester of acetic acid. The production of acetic acid is a limiting and critical factor for recombinant protein production. Extensive amounts of acetic acid accumulate in *E. coli* as an extracellular co-product of aerobic fermentation. Acetic acid under fermentation/physiological conditions appears as the ion acetate at neutral pH [22]. The amount of acetate formed is directly related to the rate at which the cells grow and this is directly related to the rate at which they consume the preferred C-source, D- glucose. Under conventional fed-batch process conditions, the growth rate of the culture is determined by the feeding rate and therefore, *E. coli* generates acetate when glucose is the limiting nutrient and the cells grow above a threshold growth rate [23]. It has been shown, that acetate formation has a negative impact on the production of recombinant proteins on the one hand but also impairs ATP synthesis due to increasing intracellular pH, which results in an decrease in the necessary pH gradient for ATP synthesis [22], [24]. Various strategies that aim to reduce or avoid the production of acetate are reported in the literature [25].

Similar to the hexose D-glucose, the pentose L-arabinose is also taken up by *E. coli via* specific transporters. However, subsequent metabolisation of this sugar molecule primarily results in the production of D-xylulose-5-phosphate which is further utilized in the pentose phosphate pathway [26].The conversion of L-arabinose to D-xylulose-5-phosphate by *E. coli*-specific enzymes has been intensively studied by Englesberg and colleagues [27–29]. Briefly, the AraA protein acts as L-arabinose isomerase that converts L-arabinose to L-ribulose and the AraB proteins acts as a kinase that phosphorylates L-ribulose, and finally AraD works as an epimerase that converts L-ribulose-phosphate to D-xylulose-phosphate [26]. The expression of these genes, araA, araB and araD (often reffered to as the araBAD) is controlled *via* a specific promoter, p_{BAD}, which is under positive regulation of the substrate L-arabinose.

Further properties of the L-arabinose controlled pBAD expression system and the importance of assessing L-arabinose uptake rates as physiological boundaries

Regulation and Expression of pBAD genes. As introduced above, the *E. coli* host/vector system utilized in this thesis ectopically expresses the therapeutic protein rhBMP-2 and its expression is under the control of the pBAD system, a regulatory system for the metabolization of L-arabinose [30], [31]. This pBAD system is positively regulated by the regulator araC gene product, AraC, and the cAMP receptor protein (CRP, also known as CAP) [26], [32], [33]. AraC acts as activator for the expression of the pBAD genes in presence of L-arabinose, while it acts as repressor for the pBAD genes in the absence of L-arabinose [34], [35]. Repression by AraC of the p_{BAD} promoter is functionally done *via* DNA looping [36], [37]. Binding of L-arabinose to AraC prevents binding of AraC to DNA and thereby abolishes DNA looping [38] (see Figure 2). Expression of the pBAD genes can be further enhanced by CRP, a DNA-binding protein, that is activated in presence of the cAMP [38–40]. cAMP levels are usually elevated in the absence of D-glucose, while they are usually low in the presence of D-glucose.

Hence, the L-arabinose operon allows controlled expression of recombinant proteins inserted into the *E. coli* genome. An important feature of this system is that expression of the protein of interest is dependent on the amount of substrate, L-arabinose. Activation of this operon follows the “all-or-nothing”-principle. Small amounts of L-arabinose might not be able to induce all cells of the culture, whereas high concentrations of L-arabinose induce expression in all cells. Furthermore, the controlled addition of the substrate allows regulating protein expression (i.e. in case that toxic protein levels are achieved, reduction of L-arabinose addition is possible).

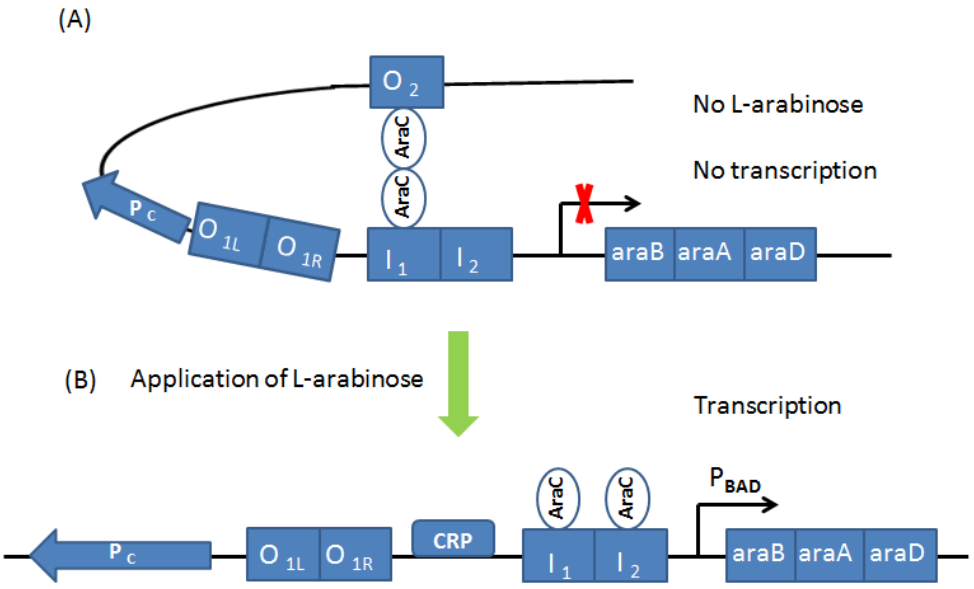

Figure 2 (A-B): Structure and regulation of the three Ara genes (araB, araA, araD) A). Repressed state. In the absence of the substrate L-arabinose, the physiological repressor, AraC, binds as dimer to upstream elements of the operon (O_2, I_1); thereby inducing a looped DNA conformation. This conformation does not allow transcription of the ara genes. B) Transcriptionally active state. In the presence of L-arabinose, the repressor, AraC adopts a different conformation that does no longer allow binding to the O_2 and I_1 sites and thereby causes the DNA loop to open. This event allows DNA-transcriptase to bind and to induce expression of the ara-genes. Full expression may be achieved if cAMP additionally binds to CRP. Adapted and modified from Schleif et al. [26].

In light of this complex interplay of multiple factors that influence pBAD expression, the detailed investigation of the influence of the two different feeds used in this study and their impact on product synthesis in terms of "catabolite sensititivity" is of utmost importance for process understanding.

Bacterial inclusion bodies

Besides the metabolic considerations that need to be taken into account when using *E. coli* for recombinant protein production, it is also of importance to acknowledge the formation of inclusion bodies (IBs). In general, IBs represent aggregates of unfolded/misfolded protein product that are insoluble and stable against degradation by proteases [41]. From a mechanistic point of view, IB formation is neither the consequence of incorrect transcription or translation nor it is an intrinsic property of the physicochemical characteristics (e.g.

molecular weight, lipophilicty) of the protein product [12], [41]. The occurrence of protein aggregates in form of IBs has been found to be due to conformational issues arising from incorrect chaperone-assisted folding [42]. One of the reasons why recombinant proteins show an increased tendency to form IBs is that the overexpression of the protein results in saturation of the chaperone capacity [42]. Additionally, proteins that contain S-S-bonds (e.g. insulin or rhBMP-2) also tend to form IBs because of incorrect formation of this covalent bond [43]. Furthermore, these aggregates are also widely considered to be biologically inactive [41]. However, recent reports showed that at least some IBs show biological activity [44]. For instance, de Groot el al. [45] and Peternel el al. [46–48] have shown that Green Fluorescent Protein (GFP) embedded in such aggregates is able to elicit fluorescence, when studied using confocal microscopy or fluorescence spectroscopy. Hence, it can be concluded that product trapped in IBs contains at least native-like conformation. Furthermore, the formation of IBs is highly dependent on process temperature; i.e. IBs have been found to be more stable when process temperature increases [45].

Despite the fact that IBs represent unwanted by-products, it has also been shown that after purification and solubilization of IBs, high overall yields can be achieved [12].

At lower process temperature the transcription and translation rate are slower, which results in less IB formation [49]. Decreasing temperature during protein production has been shown to be an effective way to raise solubility of proteins [50–52] and thereby prevent inclusion body formation [53]. Additionally a slower translation rate by reducing the cultivation temperature to 25-30°C supports appropriate folding of proteins [53], [54]. Apart from slower protein synthesis rate, also changes in folding kinetics of polypeptidic chain and smaller driving force for protein self-association are factors for better folding [55]. Furthermore low temperatures avoid degradation by heat shock proteases and prevent heat denaturation of recombinant proteins [53], [56]. Heat denaturation is more strongly associated with temperatures above 37°C [57]. Dragosits (2009) [58] showed that reducing cultivation temperature for yeast from 30°C to 20°C increased the specific product rate by 3-fold [58]. These results agree with Rodriguez-Carmona (2012)[59] who found an inverse relation between temperature and product yield, by comparing recombinant protein

production in *E. coli* at 37°C, 33°C and 30°C. In contrast to these positive effects of low temperature, plasmid copy number was more stable at 37°C than at 33°C, where it decreases over the process [59]. As already reported, lower temperature decreases protein production. Shin (1997) [60] showed that recombinant protein formation in a T7-based expression system in *E. coli* was that low that it was almost repressed at 30°C [60].

Usually, IBs do not only contain aggregated product but also contain impurities (i.e. bacterial host cell proteins). Batas and colleagues quantified IB purity of recombinant hen egg white lysozyme [12]. They found significant amounts of impurities and concluded that approximately 45% of product was present in IBs. Patra et al. made similar observations [61]. They expressed recombinant human growth hormone in *E. coli* as IBs and recovered ~50% of purified rhGH from the IBs. For DSP, inclusion body purity has a significant influence on the consecutive DSP steps, like solubilization and refolding. In general, higher IB purity results in higher refolding rates during DSP [12].

Recombinant human bone morphogenetic protein – 2

Human bone morphogenetic proteins (BMPs) constitute a protein family, which belongs to the highly conserved super-family of transforming growth factor-β (TGF-β) growth factors (GFs). BMPs play a critical role as key drivers for bone and cartilage formation [62]. Besides their osteo-inductive functions, BMPs exhibit also pleiotropic functions in embryonic neuronal development but also in cardiac development [63]. The biological/pharmacological conformation of rhBMP-2 is constituted by the homodimer. The two monomers consist of 114 AAs. Each of these two domains contains the well-known cysteine-knot motive, which is characteristic for the BMP-family, but also appears in other related growth factors, such as TGF-β and PDGF (see Figure 3)

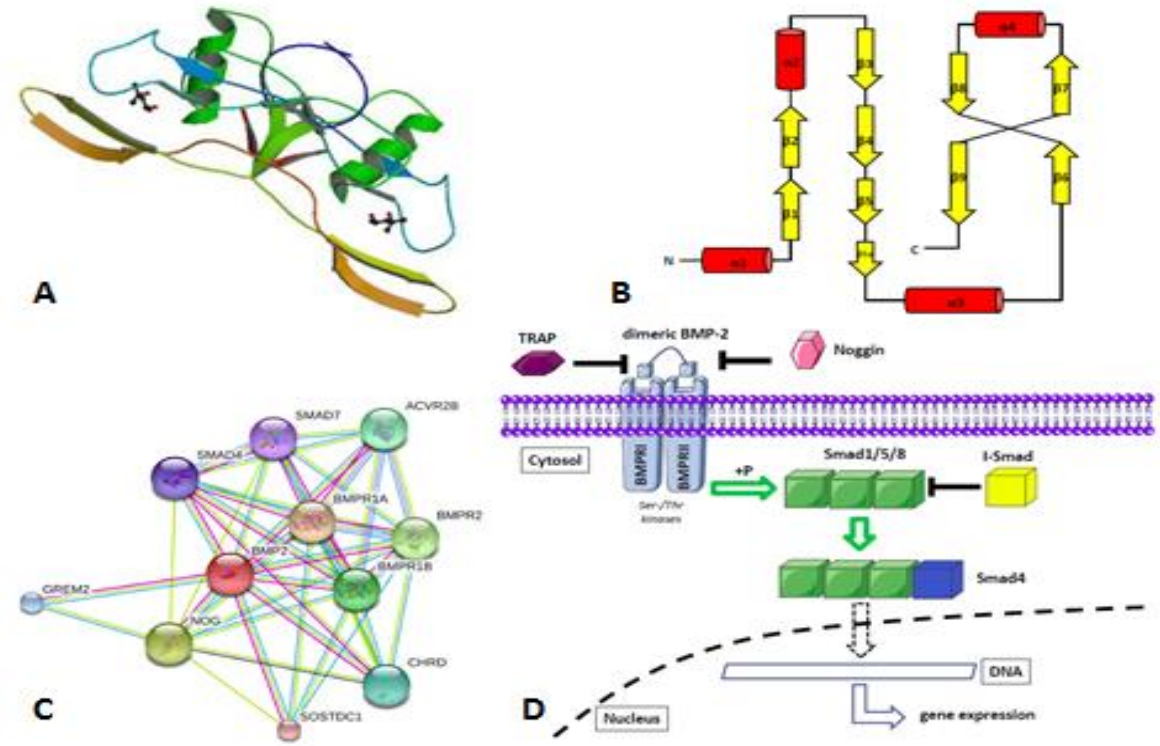

Figure 3: Structure and Function of rhBMP-2. A) 3D-conformation of dimeric-rhBMP-2. Graphic downloaded from www.rcsb.org (accession code: 3BMP) B) Secondary structure elements. RhBMP-2 consists of 4 hour α-heliches and 9 antiparallel β-sheets. (adapted and modified from Scheufler et al. [64]) C+D) Interaction partners of BMP-2. BMP-2 interacts with its membrane-bound receptors BMPR1A and BMPR1B, leading to the activation of the Smad signaling pathway that results in regulation of gene expression. C) interaction graph downloaded from STITCH2 database (stitch.embl.de). D) adapted and modified from KEGG database (pathway code: hsa4350).

This cysteine-knot is constituted by six highly conserved cysteine residues. This motive is required for the stability of the native conformation, since rhBMP-2 lacks the hydrophobic core known from other classical globular proteins [64], [65]. When BMP-2 reaches its target cells it binds to its cognate receptors, which are receptor-serine/threonine kinases. Signal transduction is then conferred *via* Smad-proteins and leads to changes in gene expression patterns of respective target genes. Owing to its demonstrated osteoinductive capacity in clinical studies, rhBMP-2 (incorporated into a medical device) was granted FDA approval in 2002 [66].

Rationale for using a mixed-feed *E. coli* system

In this thesis a mixed-feed system is employed in order to establish a highly efficient cultivation/production system for the fermentative production of this target protein, a recombinant bacterial expression system - equipped with the gene of interest and an inducible p_{BAD} promoter. This mixed-feed system is subjected to a dynamic mixed substrate feeding profile. In this context, mixed substrate feeding, describes the utilization of two substrates that are simultaneously administered to the system, instead of using only one substrate. Mixed-feed strategies aiming to optimize productivity of recombinant processes, have been successfully applied in other model organisms [16], [67]. In such a system, these one substrate may serve as the primary C-source (e.g. D-glucose), while the other substrate (e.g. L-arabinose) may act as secondary C-source and also functions as specific inducer for the heterologous expression of our target protein. Figure 5 depicts the main motivation is to control independently the transcription rate using a metabolizable carbon source (e.g. L-arabinose feed) while supplying sufficient energy to the cell (e.g. D-glucose feed). Furthermore, the mixed-feeding concept offers other benefits, like lower oxygen consumption, lower heat production, and increases biomass growth because of higher biomass yields due to the second substrate [68], [69]. The sole use of D-glucose is often associated with excessive acetate formation and other unwanted reductive overflow metabolites, which limit biomass yield and are detrimental to recombinant protein production [21], [23], [70]. The application of the secondary C-source is based on its property to selectively induce the expression of the target protein, with the aim to efficiently tune the transcription rate [71].

Besides these advantages, it must be noted that mixed-feed systems are also prone to reductive overflow metabolism (for illustration see Figure 4). Overflow metabolism describes the suboptimal utilization of an abundantly supplied energy source resulting in the production of an unwanted by-product (for illustration see Figure 4). Examples of the effects of overflow metabolism are shown in Amribt et al. and Sonnleitner et al. [72], [73].

First insides on the advantages of mixed-feed systems were obtained from the production of recombinant proteins using *P. pastoris* as expression system. In 2007, Jungo et al. showed

that the production of avidin is superior when using two substrates (sorbitol and methanol) compared to using methanol alone [68], [69].

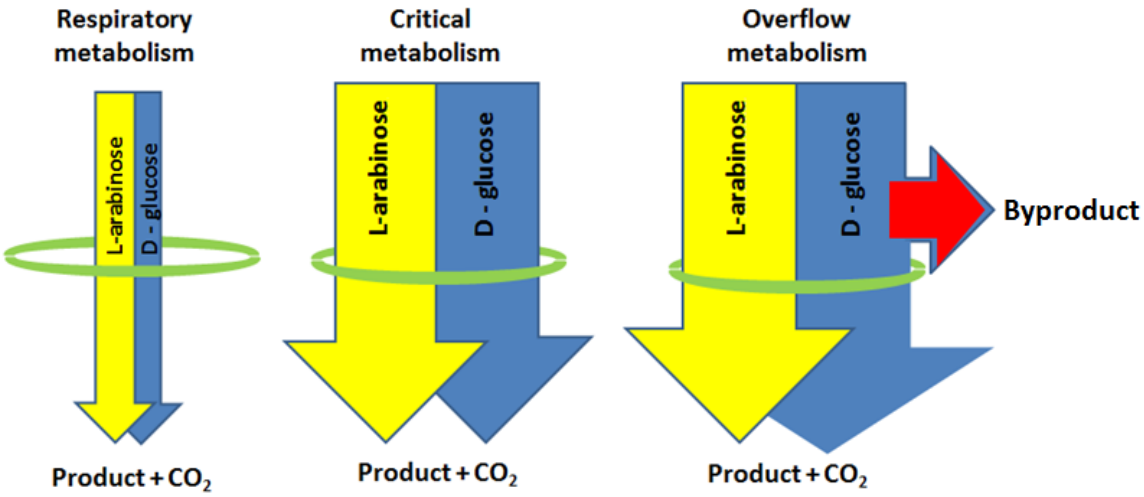

Figure 4: Respiratory metabolism: with D-glucose and L-arabinose completely consumed for cell growth. Critical metabolism (maximum respiratory capacity): with cells maximum specific growth rate. Overflow metabolism: with D-glucose and L-arabinose excess, and production of the associated metabolites (acetate). Adapted and modified from Amribt [72].

Our group also extended the approach of mixed-feeding expression platforms to *E. coli* as host protein. The pBAD mixed feed system was shown to allow tunable recombinant protein expression on cellular level, hence is not submitted to "all or none induction" [74]. Two other successful applications of a mixed-feed system in E. coli producing rhBMP-2 are given in this thesis (see Part 1 on pg. 22ff and Part 2 on pg. 51ff).

Furthermore this thesis provides valuable information on the mixed-feed system, by using L-arabinose as inducer (second C-source) while D-glucose serve as energy supply (primary C-source). The catabolite repressed uptake rates of these two carbon sources represent the physiological process descriptors, which are critically evaluated in this thesis. It is one objective to elucidate physiological boundaries for the simultaneous uptake of glucose and the catabolite repressed secondary carbon source arabinose, during changing cultivation temperature.

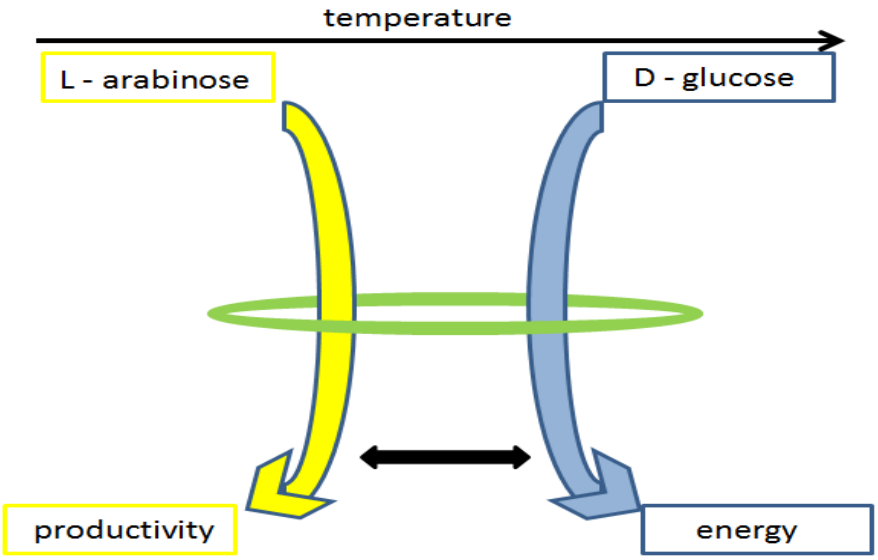

Figure 5: Demonstrate the impact of a mixed-feed system on the productivity and the energy supply. This mixed-feed system utilizes D-glucose as primary and L-arabinose as secondary carbon source. Furthermore, L-arabinose serves as inducer for the promoter-controlled expression of the recombinant protein of interest. Hence, this mixed-feed system enables transcription control of the target protein (rhBMP-2).

The advantages of dynamic experimentation

Dynamic experimentation enables bioprocess engineers to study multiple individual process parameters or multiple levels of process parameters within one fermentation run. Therefore, dynamic experimentation provides a fast and efficient means to study complex physiological processes in recombinant expression systems.

First impressing effects of dynamic experimentation were already observed in the 1980s [14]. In these experiments in yeast, dynamic substrate feeding was primarily used to decrease heat production. Since these pioneering experiments different dynamic experiments were successfully developed [14].

On principle dynamic experiments can be grouped into 4 categories: shift, ramp, pulse and oscillation experiments [14]. Shift experiments are characterized by sudden changes in process parameters to provoke certain process conditions. These process parameters are then held constant to monitor adaptation of the system to the new conditions. Shift experiments can be classified into two groups: shift-up and shift-down experiments [14]. Contrary to this fast and sudden change in process conditions, dynamic ramp experiments

are characterized by moderate to slow chances. Usually, in ramp experiments, changes of the system are slower than the physiological adaptation of the culture under investigation. Ramp experiments allow the investigation of the system at each single condition. Oscillation and pulse experiments are characterized by fast changes of process parameters. These changes are faster than the physiology of the cells can adapt [14].

In 2001, the group of von Stockar successfully applied shift-up and shift-down experiments for the development of methods for the on-line calculation of conversion rates and yield coefficients in baker`s yeast [74]. Furthermore, Dietzsch et al., performed dynamic pulse experiments to produce horseradish peroxidase in *P. pastoris* [15]. This method was used to determine the strain-specific substrate uptake rate (q_s) in a fast and easy-to-implement manner. More recently, Zalai et al. extended this approach by combining dynamic fed batch experiments and mixed feeing strategies [16].

From a methodological viewpoint, the dynamic experiments described in this thesis enclose pulse and ramp experiments.

Soft sensors to support the control strategy

Specific substrate uptake rates (q_s) have emerged as important physiological process parameters for fermentation [75]. Some authors showed that higher specific substrate uptake rates have been shown to be associated with increased specific productivity [15]. However, others reported controversial issues on this topic [76], [77]. For control of specific substrate uptake rates in fermentation experiments, rate-based soft-sensor can be applied. Soft-Sensors are process analytical technologies that provide access to critical non-measured process parameters *via* mathematical processing of readily available process data. Hence, these mathematical models are based on growth kinetics or multivariate statistical methods (e.g. PCA, MLR, neural networks) [78]. The structure of such a soft-sensor adapted for dynamic experiments is depicted in Figure 6. The soft-sensor estimates the biomass concentration during the fermentation run by using constants (e.g. biomass and substrate stoichiometries, feed densities and concentrations) and inputs (e.g. O_2 and CO_2 concentrations derived from off-gas analytics, air concentration as well as extracellular

substrate concentrations as measured using FTIR in-line). The above mentioned constants also serve as input parameters for the device "Volume Calculation". The constants, the inputs, as well as the result from the volume estimation are used for the approximation of the biomass concentration (Soft Sensor). These real-time process data are then delivered to the "Feed Rate Setpoint Calculator". The "Feed Rate Setpoint Calculator" provides a feed rate setpoint which reflects the actual process state (biomass concentration, bioreactor volume). Following a substrate balance approach, the flow rate setpoint for both feeds can be calculated according to $F_1 = qs_1 * x_1 * V/w_1$ and $F_2 = qs_2 * x_2 * V/w_2$ in real-time. In order to execute the feed rate setpoint a simple PI flow controller can be used [79].

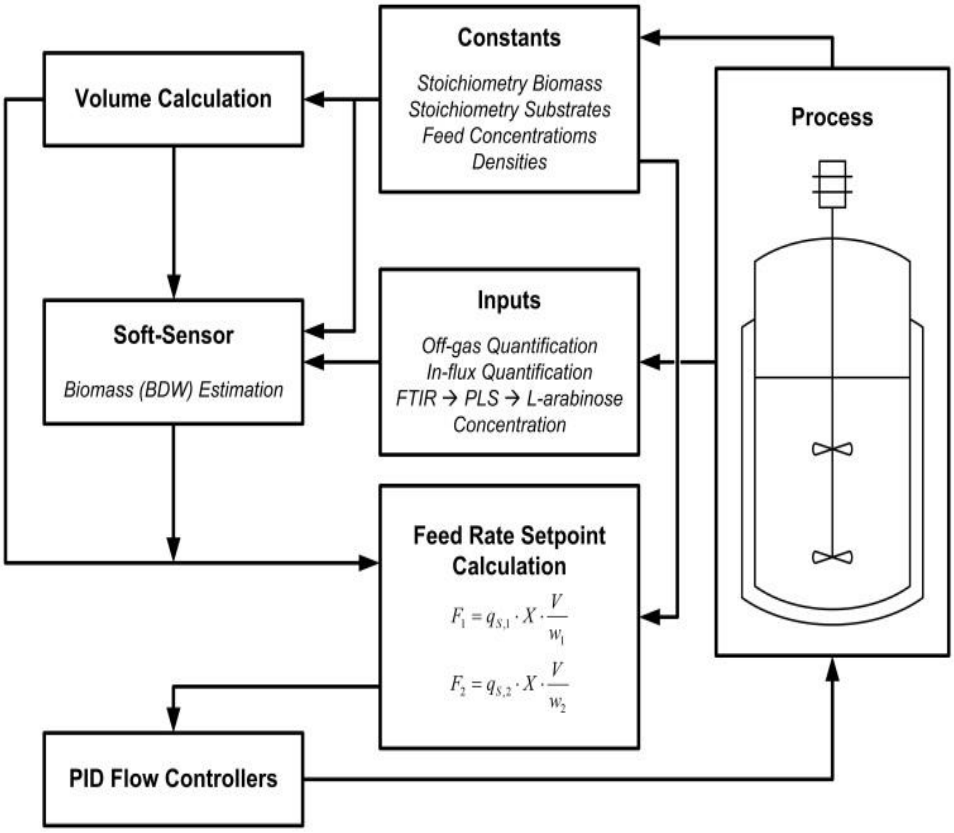

Figure 6: Experimental setup for soft-sensor assisted dynamic experimentation. Constants and inputs from the process are used for the estimation of the biomass concentration (Soft Sensor) as well as the estimation of the volume (Volume Calculation). These process data are delivered to the Feed Rate Set point Calculator that provides a feed rate set point reflecting the current process state. Execution of the feed rate set point is done *via* a simple PI flow controller. Adapted and modified from [80].

Several authors successfully applied soft-sensor technology for bioprocesses. For example, Liu et al. proposed a novel soft sensor method based on artificial neural network models for the estimation of mycelia concentration, sugar concentration and fermentation of macrolide antibiotics [81]. Additionally, Wechselberger et al. used soft sensors to derive substrate and metabolite concentrations using a kinetic model based on the respiratory limitations of baker`s yeast [6], [73]. Furthermore, in 2012, an expert panel provided detailed recommendations on soft sensor applications [82].

Design of experiments as a process investigation and optimization tool

Due to recent regulatory initiatiatives, the application of design-of-experiments (DoE) methodology has gained more and more acceptance in the development of biotechnological processes [4], [83]. DoE provides a versatile tool to optimize fermentation runs using a reduced number of experiments. On principle, the generic DoE scheme can be rather simple (see Figure 7). It relates defined input factors of a bioprocess to defined output responses.

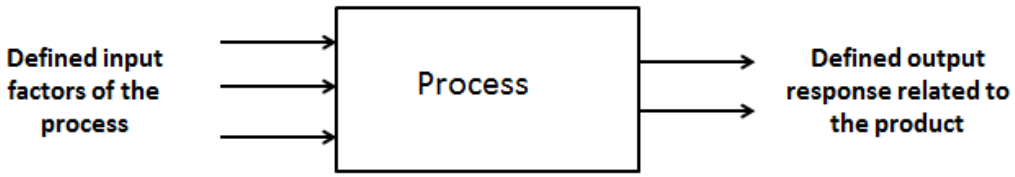

Figure 7: Basic scheme of the DoE technology. Input factors and output responses of a bioprocess.

The main advantage is, that fewer experiments are necessary on the one hand and that interaction effects between different input variables can be elucidated. Furthermore, DoE allows the signal to be decoupled from background noise and enables an estimation of the inherent experimental error by including replicate experiments [84].

In general application of DoE encompasses several steps [84] (see Figure 8):

- identification of input factors and output responses, (e.g. *via* risk assessment)
- choosing an appropriate design scheme (e.g. for screening or response surface modeling),
- generation of a design matrix,
- conduct of the respective experiments in a randomized manner in order to reduce bias,
- mathematical fitting of the generated data to describe relationships and interactions,
- model validation, and finally
- drawing final conclusions with respect to the underlying bioprocess.

In this thesis a novel approach will be carried out to gain DoE:

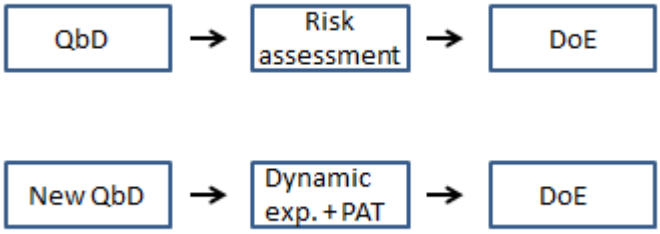

Figure 8: The figure above mentioned the state of the art QbD approach and the second depicts the new approach which was deposed in this thesis.

With respect to the choice of the design scheme, several options exist. First of all, it has to be decided if a screening experiment is conducted or if a response surface model is anticipated. In the first case, factorial design plans or fractional factorial design plans are the methods of choice. For the latter case, so-called central composite CCF designs (see Figure 9) or Box-Behnken designs are often applied. For the generation of such experimental plans, different software tools exist.

In order to derive mathematical models of the relationship between input and output parameters, different methods can be applied. The most straightforward approach is to conduct a multiple linear regression (MLR) model. In such a model the input parameters are

termed independent variables (x) and the output parameter of interest is called the dependent variable (y). The different independent variables constitute the input matrix (X). In case of MLR, the input matrix X is related to y using the classical linear equation.

$$\boldsymbol{y = k_1\, x_1 + k_2 x_2 + k_{...} x_{....} + k_n x_n * d}$$

Equation 1: General Form of a multiple linear equation

Although, the general multiple linear equation is easy to understand and intuitive only linear models can be fitted. However, it must be remembered that bioprocesses often show a non-linear behavior.

The last and final step in conducting DoE is to validate the mathematical model. For validation, different test statistics can be applied:

- R^2 = the fraction of the variation of the response variable explained by the model
- Q^2 = the fraction of the variation of the response variable predicted by the model

In principle, R^2 gives information on the internal validity of a model, whereas Q^2 provides information on the external validity (generalizability) of the model; i.e. how well new experiments with different input parameters can be predicted.

According to Mandenius, a reasonable model should show $R^2 > 0.75$ and $Q^2 > 0.6$, whereas values below 0.25 should be considered unreliable [85].

For a review of current concepts employing DoE in the context of QbD, the interested reader is referred to the review articles [7], [8], [86], [87]. Furthermore, Mandenius also gives a comprehensive overview on the application of DoE for the biotechnological production of secondary metabolites, the optimization of culture media, and the production of enzymes and other proteins [85].

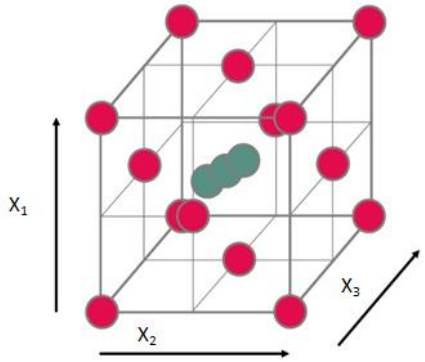

Figure 9: Central Composite CCF Design in three-dimensional input space. The CCF design approximates a sphere by additionally investigating center points.

Goals and key elements

This thesis provides valuable information on the upstream-processing of the production of the pharmaceutical protein rhBMP-2 expressed in *E. coli* using the pBAD mixed feed expression platform.

In order to gain a science and risk-based understanding of the rhBMP-2 production in *E. coli*, it is the primary objective of this work to characterize the upstream process. In order to achieve this primary objective dynamic experimentation (pulse and ramp experiments) are employed (Part 1) at the one hand, additionally a 3-factor ($q_{s\ ara}$, $q_{s\ gluc.}$ and temperature) DoE is employed (Part 2).

The specific goals of this thesis are:

Part 1: Dynamic Experimentation for the estimation of the maximum specific uptake rate for L-arabinose

- To Identify physiological boundaries for the rhBMP-2 pBAD mixed-feed system
- To determine the maximum specific uptake rate of the pBAD inducer and second C-source, L-arabinose
- To determine the maximum L-arabinose uptake rates as a function of q_s and temperature

Part 2: Application of DoE based on dynamic experiments

- Understanding of the interaction between process parameters ($q_{s\ ara}$, $q_{s\ gluc.}$ and temperature)
- Identification of the inclusion body purity
- Determination of rhBMP-2 titer and the interaction between ($q_{s\ ara}$, $q_{s\ gluc.}$ and temperature)

Roadmap and structure of the thesis

The first part of the thesis deals with the development and testing of a novel dynamic method for the physiological investigation of mixed feed systems using a combination of first-principle soft sensors and in-line Fourier transformation infrared spectroscopy. The developed method allows the investigation of mixed feed metabolic capabilities; hence how much inducing substrate and growth substrate a strain is able to simultaneously metabolize.

In the second part of the mixed feed metabolic capabilities are investigated as a function of temperature for the *E. coli* pBAD mixed feed expression platform. On the basis of the physiological information a three factor DoE is designed and performed aiming at the investigation of mixed feed ratios and temperature on inclusion body purity and final product titer.

Finally, the general benefits of the applied strategy using dynamic methods, soft-sensor control strategies and physiological DoE approaches as well as recommendations for further improvement are summarized in the final Conclusions and Outlook section.

Part 1: Dynamic Experimentation for the estimation of the maximum specific uptake rate for L-arabinose

Biological System:	rhBMP-2 expressing pBAD mixed-feed system
Scientific Question:	What is the critical value of the physiological process parameter $q_{s\ ara\ max}$ when $q_{s\ gluc}$ is constantly controlled *via* soft-sensors?
Methodological Approach:	Dynamic experimentation using pulse and ramp experiments using a soft sensor control strategy applying in-line spectroscopic measurements
Anticipated Added Value:	Characterization of the system in terms of the maximum specific uptake rate for L-arabinose

Part 2: Application of DoE based on dynamic experiments

Biological System:	rhBMP-2 expressing pBAD mixed-feed system
Scientific Question:	the impact of cultivation temperature on inclusion body purity and final product titer
Methodological Approach:	3-factor DoE using $q_{s\ ara}$, $q_{s\ gluc}$ and temperature
Anticipated Added Value:	optimal conditions in respect to inclusion body purity and final product titer

Manuscripts considered for peer-reviewed publication enclosed in this thesis:

Manuscript I: PART 1 on pg. 22 ff:

Sagmeister P., Kment M., Wechselberger P., Meitz A., Langemann T., and Herwig C*.

Soft-sensor assisted dynamic investigation of mixed feed bioprocesses

Process Biochemistry, currently in review

Indvidual Authorship contributions: PSA designed all the experiments and performed the experiments, MKM performed all the experiments and was responsible for data analysis, PWE implemented the soft-sensor technology, AME constructed and provided the E. coli strains, Analytics (SDS-page, Western blot, RP-HPLC) was done by MKM and AME, TLA assisted in performing the experiments and revised the manuscript, PSA wrote the manuscript, CHE was the principal investigator. The manuscript was circulated to all co-authors prior to submission.

Manuscript II: PART 2 on pg. 51 ff:

Kment M.a, Sagmeister P.a, Meitz A., Langemann T., and Herwig C*.

The E. coli pBAD mixed feed platform system: Investigation of temperature on mixed feed metabolic capabilities, inclusion body purity and product titer using dynamic methods and physiological design of experiments

Manuscript in preparation.

aboth authors contributed equally to this work

* corresponding author

Individual Authorship contributions: PSA designed all the experiments and assisted in performing the experiments, MKM performed all the experiments performed the analysis of all data and wrote the manuscript, AME constructed and provided the E. coli strains, Analytics (SDS-page, RP-HPLC) was done by MKM and AME, TLA assisted in performing the experiments, CHE was the principal investigator. This manuscript will be circulated to all co-authors prior to submission.

Results

PART 1

Soft-sensor assisted dynamic investigation of mixed feed bioprocesses

Patrick Sagmeister[1], Magdalena Kment[2], Patrick Wechselberger[1], Andrea Meitz[2], Timo Langemann[2,3] and Christoph Herwig*,[1]

*to whom the correspondence should be addressed to

1 Institute of Biochemical Engineering, Vienna University of Technology

2 Research Center of Pharmaceutical Engineering (RCPE) GmbH, Graz

3 BIRD-C GmbH, Kritzendorf, Austria

Keywords:

Bioprocess Technology; Mixed Feed, Dynamic Experiments, Physiological Process Control; Soft-Sensors, In-line FTIR Fourier Transformation Infrared Spectroscopy

*Christoph Herwig

Gumpendorferstrasse 1a / 166-4

A-1060 Wien, Österreich

emailto: christoph.herwig@tuwien.ac.at

Tel: +43 1 58801 166400

Fax: +43 1 58801 166980

Abstract

Recombinant mixed feed bioprocesses are characterized by the controlled feeding of multiple defined carbon sources for the benefit of increased physiological bioprocess control and enhanced productivity. Typically, a primary carbon source is applied for primary energy supply and an assimilable secondary substrate is supplied for the induction of the expression system. However, design and optimization of mixed feed bioprocesses is not straight forward due to physiological constraints such as adaptation times and catabolite repression, resulting in physiologically restricted uptake of the secondary substrate.

A novel soft-sensor assisted dynamic method that allows the physiological process design with respect to co-utilization of primary and secondary substrate in fed-batch processes was developed. The method is based on the independent control of the specific uptake rates of primary and secondary substrate *via* a combination of a rate-based first principle soft-sensor and in-line fourier transformation infrared spectroscopy. Maximum specific uptake rates and adaptation behavior of the secondary carbon source are determined by a) pulses of secondary substrate and b) the steady increase of the specific uptake rate of the secondary substrate until accumulation is observed while in both cases the specific primary substrate uptake rate is controlled at a fixed set-point.

The power of the presented method was demonstrated on a recombinant *E. coli* C41 pBAD expression system with D-glucose as primary and L-arabinose as catabolite-repressed secondary carbon source. Onset of catabolite repression was observed once a total specific substrate uptake threshold of 1.0 g/g/h was exceeded and adaptation times to L-arabinose were determined to be ~10 minutes.

The presented method can be considered generically applicable for the fast physiological investigation of mixed feed systems, paving the way for future developments in this field. Furthermore, metabolic capabilities of the promising but yet unexplored recombinant *E. coli* pBAD mixed feed system were explored for the first time.

Introduction

Fast-growing organisms like bacteria or yeast for the production of heterologous protein are widely used in the pharmaceutical industry. The added value of such production processes comes from a combination of high biomass yields and high productivity of the target protein. To speed up process development time, novel methods for the efficient design of recombinant bioprocesses are needed.

The term 'mixed feed bioprocess' describes the co-administration of multiple carbon sources for the benefit of increased bioprocess productivity [1–4]. Typically, a primary substrate supplying energy for growth (hereafter referred to as growth substrate, for example D-glucose or glycerol) and an assimilable secondary substrate inducing recombinant protein production (hereafter referred to as inducing substrate, for example L-arabinose, lactose or methanol) is administered. This opens novel degrees of freedom for bioprocessing since the energy supply through the growth substrate and the energy drain through recombinant protein production (a function of the uptake of the inducing substrate) is under process-technological control. Glycerol/methanol as well as sorbitol/methanol mixed feed processes using the *Pichia pastoris* AOX-system were reported to result in increased volumetric productivities mainly through an increase in the biomass yield coefficient [10].

The independent control of the cell's energy recourses and the transcription rate of recombinant protein *via* mixed feed approaches open new process-technological perspectives for a) obtaining proteins in the desired folding state and b) higher overall volumetric productivities during long term expression. For *E. coli* bioprocesses, transcription control approaches through repressor titration are reported [5]. However, although very promising, to the authors' knowledge no recombinant mixed feed bioprocess for *E. coli* is reported.

Under carbon limited conditions, bacteria are capable of simultaneous utilization of multiple carbon sources at a time, a physiological capability vital for survival in natural environments [6]. In case substrates are available in non-limiting concentrations, substrates are used sequentially [7], a phenomenon referred to as diauxic growth which can be attributed to

catabolite repression [8]. Kinetics studies on mixed substrate utilization in *E. coli* in continuous cultures demonstrated that *E. coli* is capable of utilizing multiple carbon sources at a time and the contribution of one substrate to the total carbon uptake is reflected by its steady state residual substrate concentration [6]. Mixed-feed bioprocess design demands knowledge about maximum metabolic capabilities of the simultaneous uptake of growth and inducing substrate. Hence, investigation of mixed substrate utilization including the investigation of physiological boundaries is mandatory for the design of recombinant mixed feed bioprocesses: Multiple substrates are administered at a time and the strains capabilities of mixed substrate utilization, described by the onset of catabolite repression, should not be violated. However, to the authors' knowledge there has been no physiological study so far dealing with the physiological investigation of bacterial recombinant mixed feed fed-batch bioprocesses, probably due to the absence of fast and efficient methods for the investigation of physiological process boundaries.

Dynamic (also referred to as transient) experimentation involves the controlled deflection of process states in order to identify transient regulatory behaviors [9] and quickly determine strain characteristics and process performance related attributes. Investigation of *P. pastoris* AOX mixed feed processes was carried out *via* transient experimentation in continuous cultures [3] , [10] and dynamic fed-batch experimentation with decoupled glycerol/methanol feed streams where optimal ratios of the specific glycerol/methanol uptake rates for maximizing the specific production rate were identified [4]. Following a recent review [11], dynamic approaches in bioprocess development can roughly be categorized in a) ramp experiments b) shift experiments and c) pulse experiments. Ramp experiments typically aim at submitting the culture to a slow dynamic change in process conditions where time regimes of the dynamic deflection of process states are above the time regimes of cellular adaptation mimicking (pseudo-) steady state conditions as in fed-batch cultivation. The application of ramp experimentation allowed the investigation of the respiratory bottleneck through decellerostat *E. coli* cultivations [12] as well as the characterization of *S. cerevisiae* strains [13], the investigation of cell physiology [14] and the determination of strain specific parameters [15]. In contrast to ramp experiments, pulse experiments offer the possibility to

observe transient behavior such as adaptation to novel substrates. Recently, a dynamic method using pulses for the design of feeding strategies in recombinant *P. pastoris* cultures was reported [16].

Soft-sensors are process analytical devices capable of estimating a difficult to be measured process variable by model-driven processing of readily available process data. Numerous data driven approaches are reported in literature, as reviewed recently [17]. Data driven soft-sensors rely on existing data sets for the construction of multivariate models, for example *via* Partial Least Squares (PLS), Artificial Neural Networks (ANNs) or Principal Component Regression (PCR). However, a sufficing data basis for the application of data driven soft-sensors is typically not available during bioprocess development. Soft-sensors using little prior knowledge on the basis of first principle relationships provide a promising alternative. Elemental balancing can be considered a first-principle approach [18]. On the basis of an over-determined equation system (carbon, nitrogen, degree of reduction and charge balances), unknowns can be estimated and the quality of the estimation can be checked. Although the principle of this approach dates back to the 1990's [19], it recently finds novel applications for bioprocess monitoring [20] and bioprocess control [12]. Furthermore, the benefits of data reconciliation, a constrained based statistical data processing method, was applied successfully for the real-time correction and consistency checking of multiple concentration estimates from in-line FTIR [21], [22].

The presented approach is demonstrated on an *E. coli* pBAD mixed feed bioprocess. The C-limited mixed feed fed-batch process under investigation is based on the simultaneous utilization of D-glucose (primary carbon source) and L-arabinose using the pBAD expression system in *E. coli* which has been described extensively [23–30]. L-arabinose acts as secondary carbon source as well as inducer, whereby recombinant gene expression can be considered tightly coupled to the uptake of L-arabinose [25]. The pBAD expression system allows tightly controlled and inducer concentration dependent recombinant gene expression in a wide host range (Newman 1999). In contrast to one pint additions with inducer negative strains, the pBAD mixed feed system intends to achieve higher overall productivities by process-technological control of assimilable inducer feeding.

The presented method aims at the efficient acquiring of reliable information on the physiological boundaries of mixed feed bioprocesses with a minimum of experiments. To achieve this goal, the benefits of i) dynamic pulse and ramp experimentation, ii) soft-sensor assisted bioprocess control and iii) real-time chemical information derived from in-line spectroscopic measurements were combined within one experimental method. A soft sensor assisted bioprocess control strategy capable of the control of specific substrate uptake rates is applied. Changes in the residual substrate concentrations during dynamic pulse experimentation are accounted for *via* in-line attenuated total reflection Fourier Transformation Infrared spectroscopy (in-line FTIR) and delivered to the soft-sensor. The dynamic and static control of specific substrate uptake rates within dynamic pulse and ramp experimentation is achieved by the soft sensor assisted bioprocess control strategy. This is prerequisite for the development and/or optimization of bioprocesses for the production of heterologous proteins.

Material & Methods

Strain

E. coli C41 (F– ompT hsdSB (rB- mB-) gal dcm (DE3); Lucigene, Middleton, WI, USA) carrying the plasmid pBK-BMP for production of recombinant human bone morphogenetic protein 2 (rhBMP-2) was used. pBK-BMP originates from pBAD24 (provided by BIRD-C, Vienna, Austria). The gene coding for rhBMP-2 was cloned under the control of the L-arabinose inducible pBAD. The ampicillin resistance cassette of pBAD24 was exchanged with a kanamycin resistance. *E. coli* C41 strains have an intact L-arabinose metabolism.

Media

A defined minimal medium with D-glucose as main carbon source (batch medium D-glucose concentration: 20 g/l; fed-batch medium D-glucose concentration 400 g/l) as described in detail elsewhere [31] was used.

Bioreactor setup

Experiments were conducted in a Techfors-S bioreactor (Infors, Bottmingen, Switzerland) holding 10 L working volume. Base and feed bottles were placed on balances (Sartorius,

Göttingen, Germany). A Techfors-S integrated analogue pump was used for the addition of the growth substrate feed and a peristaltic analogue pump (Lamda, Baar, Switzerland) was used for the addition of the assimilable inducing substrate. Dissolved oxygen (Hamilton, Reno, USA), pH (Hamilton, Reno, USA) and headspace pressure (Keller, Winterthur, Switzerland) were monitored during the fermentation. Off-gas concentrations of CO_2 and O_2 were measured by a gas analyzer system (Müller Systems AG, Egg, Switzerland) based on non-dispersive infra-red (CO_2) and paramagnetic (O_2) measurement principle. Signals were recorded by the process information management system (PIMS) Lucullus (Secure Cell, Schlieren, Switzerland).

Fermentation parameters

Dissolved oxygen levels (DO_2) were maintained above 40% saturation (100% saturation were set before inoculation at 35°C, 0.3 bar gauge, pH 7.2). The pH was kept constant at 7.2 by adding 12.5% NH4OH, which also served as nitrogen source. Temperature and feeding profiles were varied and controlled as described in the results section.

Off-line analytical methods

Biomass dry weight concentration

Biomass dry weight (BDW) concentrations were determined from the culture broth. 2 ml of the cell suspension were centrifuged (RZB 5171, 10 min, 4°C) in pre-weighted glass tubes. The pellets were washed twice using distilled water and dried at 105°C for 72 hours. The BDW concentration was determined in duplicates.

Metabolite concentrations

Cell-free supernatant samples for determination of residual substrate concentrations were taken from the vessel using an in-line ceramic 0.2 µm filtration probe (IBA, Heiligenstadt, Germany). Concentrations in the supernatant were measured by HPLC (Supelcogel C-610, Sigma Aldrich, St. Louis, USA), using an isocratic gradient (0.5 ml/min, 0.1% H3PO4 as eluent).

Soft sensor Rate calculation, reconciliation, biomass estimation

A soft sensor based on a redundant equation system (Carbon balance and Degree of Reduction balance) was used as described in detail elsewhere [18], [20]. Based on the calculation of oxygen-, carbon dioxide, and substrate turnover rates the soft-sensor estimates the unknown biomass turnover rate. The estimated biomass turnover rate is then numerically integrated to estimate the total biomass formed. The redundant equation system (degree of redundancy of 1) allows gross error detection and the reduction of random noise using a reconciliation approach adapted from literature [19].

The soft sensor was extended to account for changes in the residual concentration of the inducing substrate which was estimated from in-line FTIR measurements. The calculating tools were called in time intervals of 300 seconds.

In-line residual substrate measurement

An ATR-FTIR spectrometer (ReactIR, Columbus, Ohio, USA) interfaced with the bioreactor *via* a 25 mm Ingold nozzle was used. Infrared spectra (256 scans) were measured in intervals of 2 min and used for the estimation of residual arabinose concentrations *via* a partial least squares (PLS) model (2nd derivative spectra, one principal component). The corresponding model was established prior to fermentation using off-line HPLC data from previous fermentations. Supernatant concentrations [g/Lsup] were corrected with the cell-bio-volume to obtain concentrations in units of total reactor volume [g/Lrv] before being used for the establishment of the PLS model.

In-line real-time data exploitation

IC Quant (Mettler Toledo, Columbus, Ohio, USA) was used for the real time computation of a PLS model based on historical data for the estimation of the residual arabinose concentration. Respective concentrations were delivered to the soft-sensor which was implemented using the Sim-Fit tool of the Lucullus PIMS (Secure Cell, Schlieren, Switzerland).

Results and Discussion

Schematic description of the developed method

The presented fed-batch method for the determination of the specific uptake rates $q_{s,ind.}$ of a catabolite-repressed inducing substrate (here: L-arabinose) with simultaneous uptake of growth substrate (here: D-glucose) is composed of three distinct phases as illustrated in Figure 1. Following a batch phase for the initial accumulation of biomass, the specific substrate uptake rate of the growth substrate ($q_{s,grow.}$) is controlled on a defined set-point in all further process phases *via* a soft-sensor assisted control strategy (described below).

Simultaneous to the controlled uptake of growth substrate (straight line, Figure 1), a pulse of inducing substrate giving a final concentration of 12 g/l (dot-dashed line, Figure 1) in the broth is applied during phase 1. Evaluation of phase 1 provides information on a) adaptation kinetics of the culture to the inducing substrate and b) the maximum uptake rate of inducing substrate $q_{s,ind.,max.}$ Following the depletion of the inducing substrate a second feed stream with inducing substrate is started to continuously supply both growth and inducing substrate to the culture during phase 2. The specific uptake of inducing substrate (dashed line, Figure 1) is gradually increased until accumulation of inducing substrate is observed *via* in-line FTIR. The specific substrate uptake rate of inducing substrate at the time-point of accumulation corresponds to the maximum specific uptake rate $q_{s,ind.,max}$ of inducing substrate with simultaneous controlled uptake of primary substrate. Following the accumulation of inducing substrate, the feed of inducing substrate is terminated and the uptake of secondary substrate is monitored resembling a second pulse experiment in phase 3 (Figure 1). This strategy represents a combination of pulse experimentation (phase 1 and phase 3) and ramp experimentation (phase 2) which allows for the comparison of $q_{s,ind.,max}$ as obtained from the pulses and $q_{s,ind.,max}$ at the time point of its accumulation during the ramp within a single experimental run.

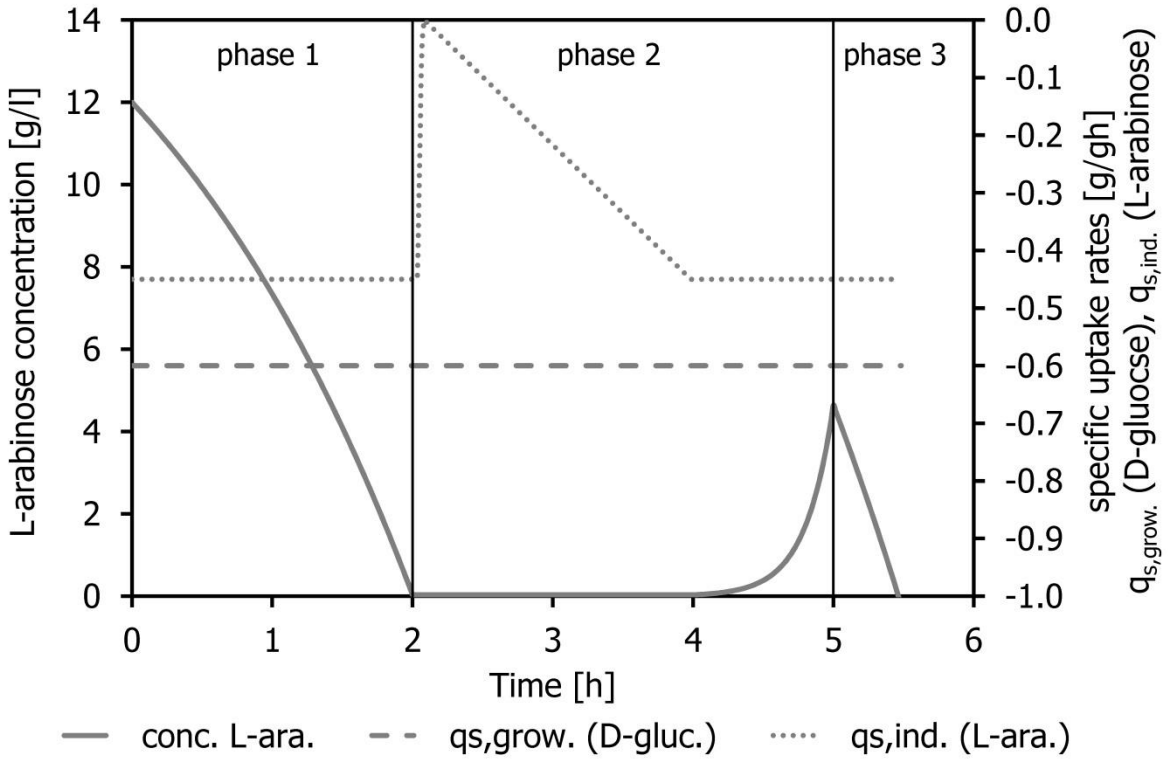

Figure 1: Schematic representation of the presented method for the determination of the maximum specific substrate uptake rate of inducing substrate with simultaneous controlled uptake of growth substrate.

Control strategy: Soft-sensor assisted control of the specific primary substrate uptake rate at simultaneous uptake of residual inducing substrate

The method described relies strongly on a robust control of the specific uptake rate of growth substrate. This is not straight forward as the control of specific substrate uptake rates demands real-time information on the biomass concentration, a highly dynamic process state that is influenced by the combined uptake of primary and secondary substrate in mixed feed systems. In order to cope with this challenge, a rate-based soft-sensor as described in detail elsewhere [12], [20] was extended to account for changes is the residual concentration of inducing substrate as determined *via* in-line FTIR. The structure of this soft-sensor is depicted in Figure 2. Given constants (such as biomass and substrate stoichiometry, feed densities and concentrations) and measured inputs (such as O_2 and CO_2 concentrations in the off-gas, supply gas concentrations and concentrations of extracellular inducing substrate) are used for calculation of the reactor volume (V) and estimation of the BDW concentration (X) during the fermentation. These real-time process data are delivered to a

feed rate calculator. Following a substrate balance approach, the flow rate setpoint was calculated according to Equation 1 in real-time for multiple feeds (growth and inducing substrate, with respective substrate mass fractions w). Execution of the feed rate set-points was done *via* PID flow controllers within the PIMS.

$$F = q_s \cdot X \cdot V / w$$

Equation 1: Calculation of the feed rate set-points.

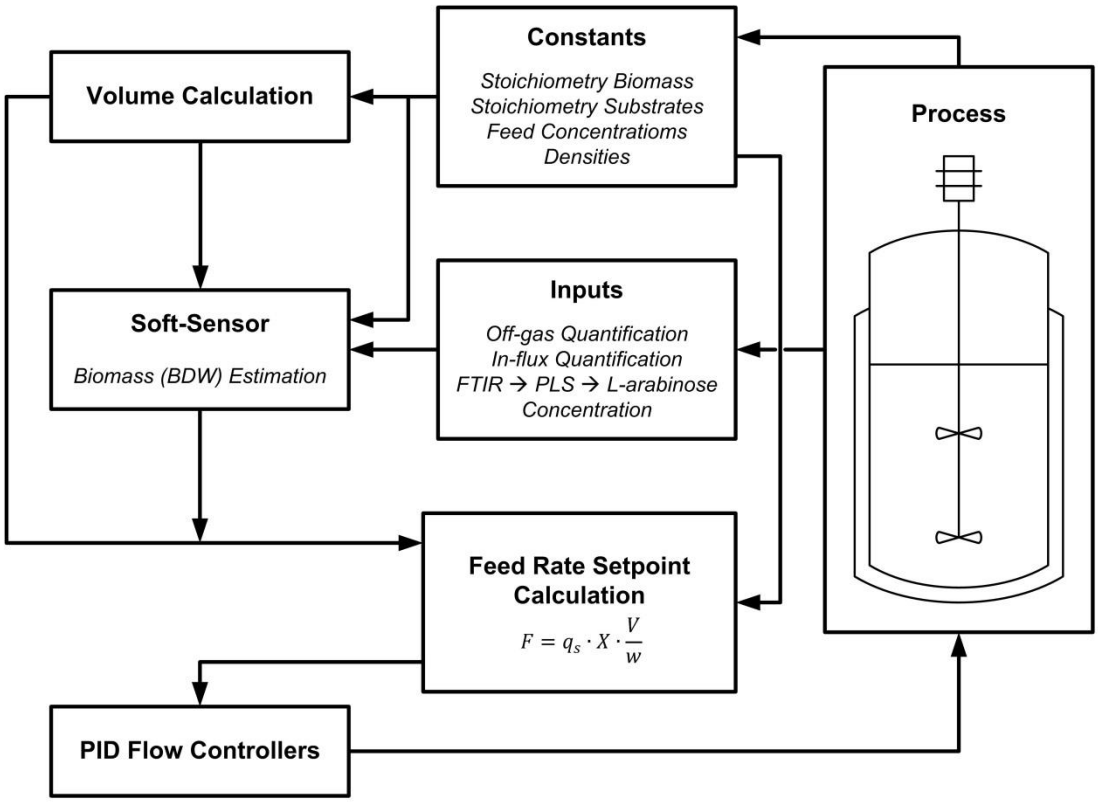

Figure 2: Experimental setup for soft-sensor assisted dynamic experimentation.

Performance of the presented strategy: Determination of the maximum specific substrate uptake of L-arabinose at controlled specific uptake of D-glucose

Here, the in-process performance of the described method for the determination of the maximum specific substrate uptake of the inducing substrate L-arabinose ($q_{s,ind.,max}$) with simultaneous controlled uptake of the growth substrate D-glucose ($q_{s,grow.)}$ is described.

Phase 1 - 1st Pulse of L-arabinose

Following a batch phase (data not shown) the culture was shifted to a fed-batch with soft-sensor controlled specific uptake of inducing substrate. In order to identify $q_{s,ind.,max}$, the specific uptake of the growth substrate was held constant at a set-point of $q_{s,grow.}$ = 0.6 g/g/h (Figure 3 C, dashed grey line) while a pulse of L-arabinose (12 g/l) was submitted to the system at 3.2 h process time (Figure 3 A-C, vertical arrow). Figure 3 A depicts the concentration of inducing substrate over process time and compares the data obtained from in-line FTIR measurements (solid grey line) to off-line measurement (empty circles) of the L-arabinose concentration from HPLC. It can be seen that - except for a low-level off-set in the FTIR data - the two methods align well. From the off-line data the maximum specific uptake rate of inducing substrate was determined as $q_{s,ind.,max}$ = 0.35 g/g/h and the adaption time of the pBAD expression systems was estimated on the basis of the in-line FTIR L-arabinose estimation with approximately 6 - 10 min (Figure 3 A, horizontal arrow).

The biomass estimation using the applied first principle soft-sensor provided the basis for the control strategy. Figure 3 B shows a comparison of the biomass concentration estimation of the soft-sensor (solid grey line) as well as off-line biomass dry cell weight measurements (filled squares). In general, the two data fit well and the soft-sensor reflects the strong increase in biomass concentration during the pulse accurately. However, from Figure 3 B it can be seen that soft-sensor overestimates the biomass concentration which can be explained by an off-set due to a wrong initial value implemented at the start of the soft-sensor.

Additionally, the formation of acetate was monitored during the process. Acetate formation started at 4.7 h process time, reached its peak of 1.5 g/L at 5.1 h and was fully metabolically

degraded again at 5.6 h (Figure 3 B, empty diamonds). Since the specific uptake rates of growth and inducing substrate were constant throughout the pulse experiment, the delayed start of acetate production might be attributed to an unidentified time dependent effect.

Summarizing, it can be concluded that the strain under investigation is capable of metabolizing the inducing substrate L-arabinose at a maximum specific uptake rate of $q_{s,ind.,max}$ = 0.35 g/g/h with a simultaneous uptake of $q_{s,grow.}$ = 0.65 g/g/h (D-glucose) following an adaptation phase of 6 - 10 min.

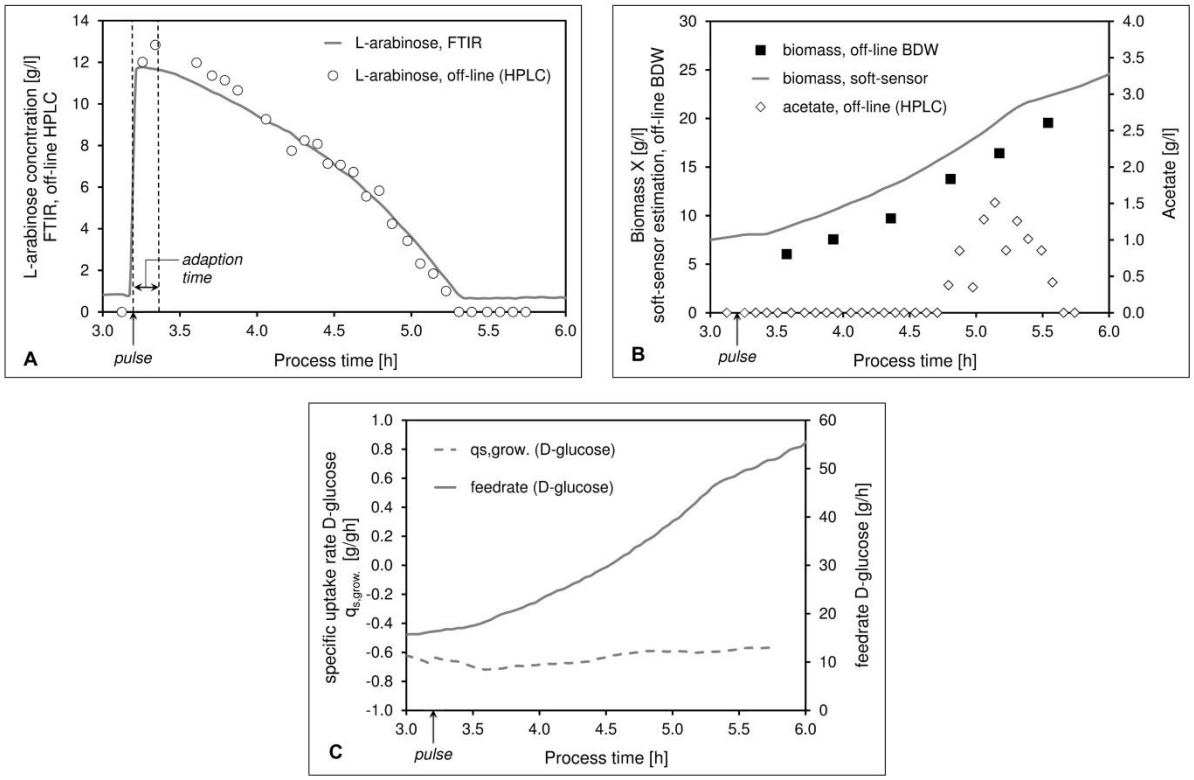

Figure 3 (A-C): Pulse experiment. A: Concentrations of L-arabinose over process time. Solid grey line: estimation *via* in-line FTIR, empty circles: off-line measurements *via* HPLC. The vertical arrow indicates the moment of application of the L-arabinose pulse. The horizontal arrow indicates the adaption time. B: Kinetics of changes in biomass dry cell weight and acetate production over process time. Grey line: soft-sensor assisted estimation of biomass, black squares: off-line biomass measurements BDW, empty triangles: acetate concentration in the culture broth supernatant. C: Control of the pulse experiment at a constant specific uptake rate of D-glucose ($q_{s,grow.}$). Dashed grey line: $q_{s,grow.}$, solid grey line: D-glucose feed rate.

Phase 2: Controlled ramp of $q_{s,ind.}$ until accumulation of L-arabinose

Following the complete consumption of L-arabinose from the pulse a ramp experiment was conducted by continuously increasing the specific uptake rate of inducing substrate. The goal of this experiment was to quantify the maximum specific uptake rate of inducing substrate $q_{s,ind.,max}$ at the moment of accumulation of L-arabinose. The specific uptake rate for D-glucose was again kept constant at a set-point of $q_{s,grow.}$ = 0.6 g/g/h (Figure 4 C, dashed grey line). The specific uptake rate of inducing substrate was subjected to a linear increasing ramp starting from 0.0 g/g/h to 1.0 g/g/h. Figure 4 C depicts the specific uptake rates of growth (dashed grey line) and inducing substrate (dashed black line) which were both controlled through the soft-sensor strategy. The change in the biomass concentration was estimated accurately by the soft-sensor (Figure 4 B, solid grey line). The actual uptake rate of growth substrate was found as $q_{s,grow.}$ = 0.58 g/g/h. Similar to phase 1, the concentration of L-arabinose estimated *via* in-line FTIR corresponded to the off-line measurements (HPLC) with a low-level off-set (Figure 4 A). The moment of accumulation of L-arabinose in the culture broth supernatant was at 7.45 h process time (indicated by vertical arrows in Figure 4). At this time the maximum specific uptake rate of inducing substrate with simultaneous controlled uptake of growth substrate was reached with $q_{s,ind.,max}$ = 0.35 g/g/h. At the same moment also the accumulation of acetate was observed. This observation indicates that at a total substrate flux of $q_{s,tot.}$ = 0.93 g/g/h the system switches to a reductive NADH generation probably due to a bottleneck in the TCA-cycle [9].

The ramp experiment in phase 2 identified the maximum specific uptake rate of inducing substrate as $q_{s,ind.,max}$ = 0.35 g/g/h; this is virtually the same value as identified by the 1st pulse experiment in phase 1. This observation is an orthogonal confirmation of the results obtained before.

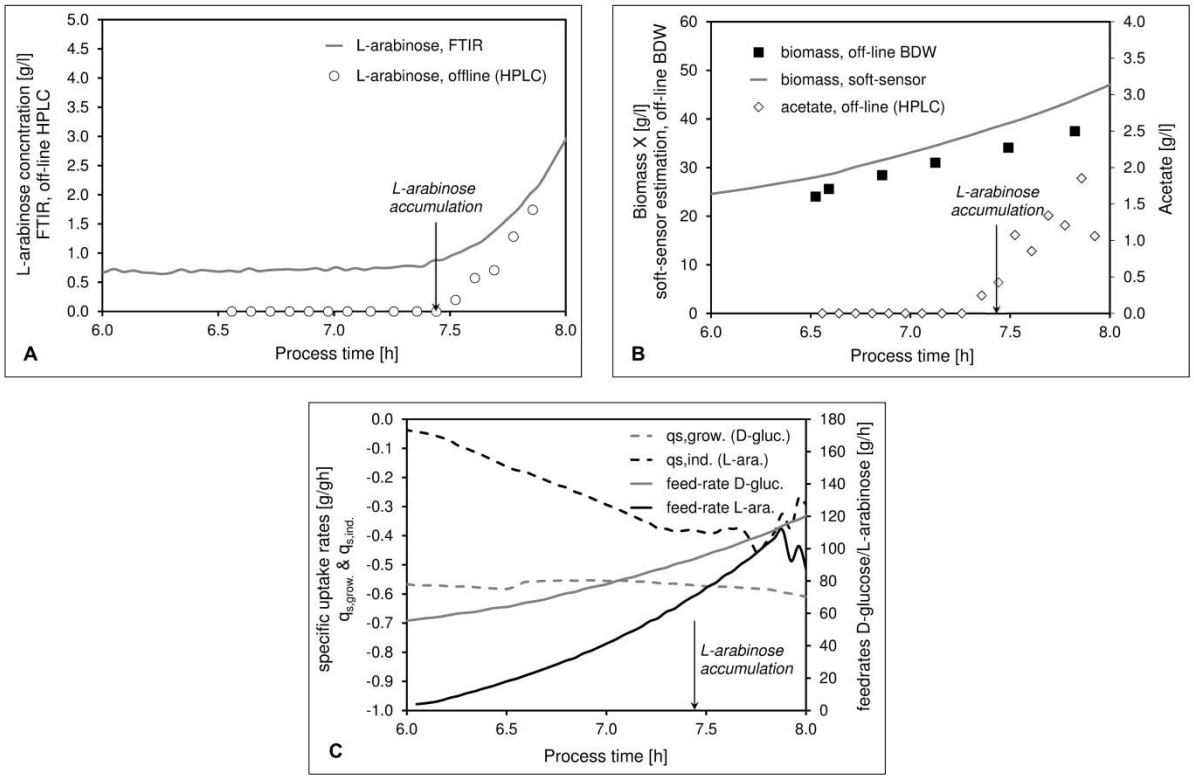

Figure 4 (A-C): Ramp experiment. A: Concentrations of L-arabinose over process time. Solid grey line: estimation *via* in-line FTIR, empty circles: off-line measurements *via* HPLC. The vertical arrow indicates the moment of L-arabinose accumulation. B: Kinetics of changes in biomass dry cell weight and acetate production over process time. Grey line: soft-sensor assisted estimation of biomass, black squares: off-line biomass measurements BDW, empty triangles: acetate concentration in the culture broth supernatant. C: Determination of the specific uptake rate of L-arabinose ($q_{s,ind.}$) and control of a constant specific uptake rate of D-glucose ($q_{s,grow.}$). Black lines: L-arabinose, grey lines: D-glucose, solid lines: feed-rates, dashed lines: specific uptake rates. The horizontal arrows identify the maximum specific uptake rate of inducing substrate under the given conditions.

Phase 3: Uptake of accumulated L-arabinose (2nd pulse)

At a process time of 7.45 h the maximum uptake of inducing substrate was reached and L-arabinose started to accumulate in the culture broth supernatant because the physiological capacity was exceeded. The feed of L-arabinose was stopped when an excess of approximately 3.5 g/l had accumulated (at 8.1 h process time, vertical arrows in Figure 5) and the subsequent uptake of the inducing substrate was investigated as a 2nd pulse experiment within the fermentation (phase 3). The ceramic 0.2 µm filtration probe allowed a high sampling frequency (45 seconds) of cell-free supernatant and therefore a quantitative evaluation of this highly dynamic process phase. As in phase 1 and 2 the concentration of L-arabinose in the culture broth supernatant estimated *via* in-line FTIR (Figure 5 B, solid grey line) was confirmed by off-line measurements (HPLC, Figure 5 B, empty circles). With this data the maximum specific uptake rate for L-arabinose was determined as $q_{s,ind.,max}$ = 0.36 g/g/h for this phase, while the specific uptake rate of D-glucose was again held constant at $q_{s,grow.}$ = 0.63 g/g/h by the soft-sensor control strategy as shown in Figure 5 C (dashed grey line, set-point: $q_{s,grow}$. = 0.6 g/g/h).

Acetate concentrations in the culture broth supernatant were also determined with off-line HPLC measurements (Figure 5 B, empty diamonds). The concentration of acetate kept increasing until all L-arabinose was depleted. Afterwards the culture again fell below the maximal physiological capacity and started to re-metabolize the excess acetate as an additional C-source.

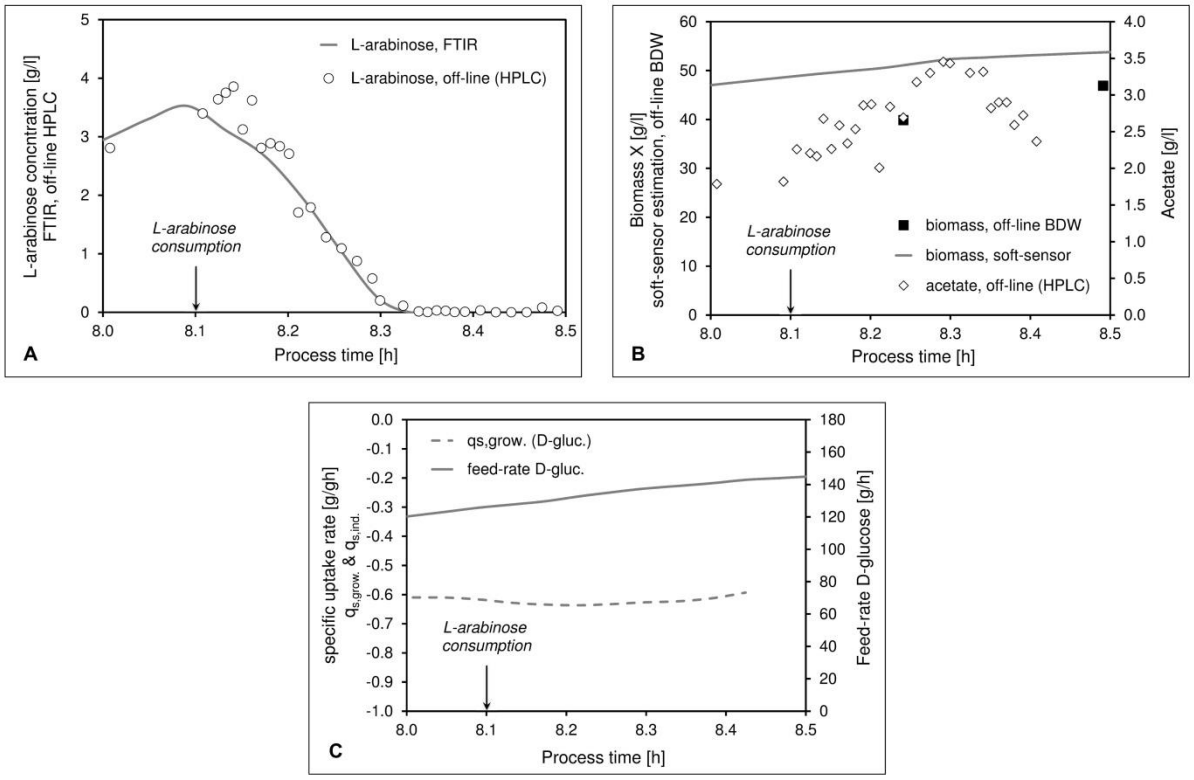

Figure 5 (A-C): Consumption of accumulated L-arabinose (2nd pulse). A: Concentrations of L-arabinose over process time. Solid grey line: estimation *via* in-line FTIR, empty circles: off-line measurements *via* HPLC. The vertical arrow indicates the termination of the L-arabinose feed stream. B: Kinetics of changes in biomass dry cell weight and acetate formation over time. C: Control of the specific uptake rate (qs,grow.) of D-glucose. Black arrows denote L-arabinose consumption.

Comparison of the process phases

The primary objective of the above experiment was the physiological investigation of *E. coli* C41 (pBAD) in the mixed feed system D-glucose/L-arabinose with respect to the maximum specific uptake rate of inducing substrate $q_{s,ind.,max.}$ For all three phases of the experiment highly similar values were observed (1st pulse: 0.35 g/g/h, ramp: 0.35 g/g/h, 2nd pulse: 0.36 g/g/h) while the simultaneous uptake of growth substrate was held constant at $q_{s,grow.}$ = 0.62 ± 0.04 g/g/h. This corresponds to an average maximum specific uptake rate for L-arabinose of 0.35 ± 0.01 g/g/h for the whole fermentation which highlights the accuracy and value of the proposed method.

In order to further validate the analyses a carbon balance (C-balance) was compiled (Table 1). A C-balance of 100% denotes full recovery. During phase 1 and 3 C-balance the recoveries were 106% and 101%, respectively. The C-balance during phase 2 gave a recovery of 89%. The observed error of 11% on the overall C balance corresponds approximately to an error of 3% on the individual rates (Wechselberger et al. 2010). This can probably be attributed to a higher error on the specific substrate uptake rate due to the combined use of two individual feeds compared to the pulse experiments where only one feed is administered.

These results were confirmed by compiling degree of reduction (DoR) balances which also gave recoveries of 113, 82 and 102% for the three phases, respectively. Hence, on the basis of the overall recoveries (C- and DoR balance) it can be concluded, that the highly dynamic pulse phases can be correctly and quantitatively described.

Table 1: Summary of rates and yields over all phases of the experiment.

	1st pulse	ramp	2nd pulse
specific rates			
qs ind.	0.35	0.35	0.36
qs grow.	0.65	0.58	0.63
qs ace.	0.05	0.03	0.05
qs tot.	1.00	0.93	0.99
μ	0.56	0.30	0.54
balances (recovery)			
carbon balance	1.06	0.89	1.01
DoR balance	1.13	0.82	1.02

Investigation of maximum specific L-arabinose uptake as a function of specific D-glucose uptake

Maximum specific L-arabinose uptake rates ($q_{s,ind.,max}$) as a function of the maximum specific D-glucose uptake rate ($q_{s,grow}$) were determined on three levels. Specific D-glucose uptake rates ($q_{s,grow}$) were controlled at defined set-points ($q_{s,grow.}$ = 0.6/0.25/0 g/g/h). The maximum uptake rate of L-arabinose ($q_{s,ind.,max}$) was measured (response). A triplicate fermentation run ($q_{s,grow.}$ = 0 g/gh) was done to estimate the experimental error, resulting in 5 individual fermentation runs.

Quantitative evaluation of the runs followed the method as described in the previous section and are summarized in Table 2. The results were used to create a linear regression model displaying the maximum specific uptake rate of inducing substrate as a function of the specific uptake rate of growth substrate as shown in Figure 6. The function obtained on the basis of five individual fermentation runs is given in Equation 2.

$$q\_(s, ind., max) = -1.15 \cdot q\_(s, grow.) - 1.03$$

Equation 2: Maximum specific uptake rate of inducing substrate (L-arabinose) as a function of the specific uptake rate of growth substrate (D-glucose). Coefficient of determination: R^2 = 0.962.

The maximum specific uptake of inducing substrate is the highest in case no additional growth substrate is supplied and linearly decreases as a function of the specific uptake of growth substrate.

Table 2: Maximum specific uptake of inducing substrate (L-arabinose) at controlled specific uptake levels of growth substrate (D-glucose).

$q_{s\ grow,}$ [g/g/h]	$q_{s\ ind,max}$ [g/g/h]	$q_{s\ tot,}$ [g/g/h]
0.00	1.08 ± 0.10	1.08 ± 0.10
0.26	0.65	0.91
0.62 ± 0.03	0.35 ± 0.01	1.02 ± 0.05

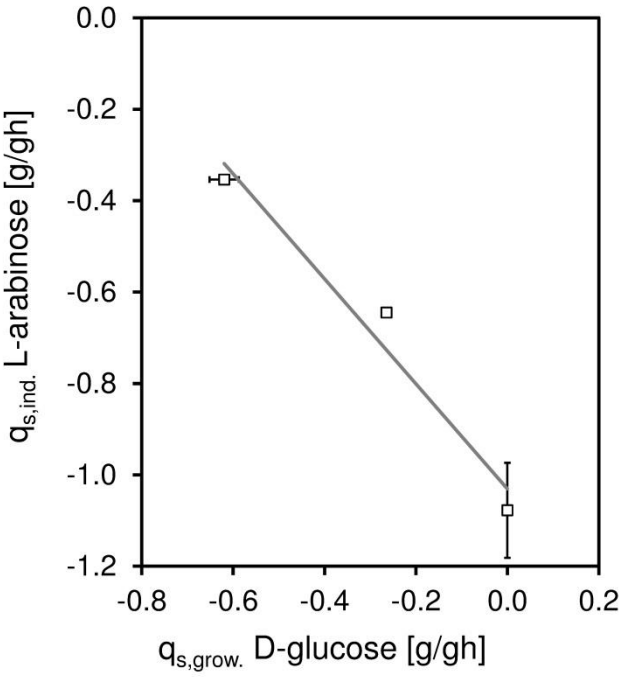

Figure 6: Specific uptake of inducing substrate as a function of specific uptake of growth substrate. Data reflects the physiological boundaries of *E. coli* C41 (pBAD) for simultaneous uptake of D-glucose as growth substrate and L-arabinose as inducing substrate at 35°C.

Catabolite repression as bioprocess design boundary for mixed feed bioprocesses

The onset of catabolite repression, hence the preferential uptake of one carbon source, constitutes a physiological boundary to the design of mixed feed bioprocesses. Violating this boundary results in accumulation of the inducing substrate while the targeted specific uptake rate of inducing substrate is not met. Hence, it is advised that mixed feed metabolic capabilities expressed as the function similar to Equation 2 are established prior to bioprocess optimization because a violation of the catabolite repression boundary impedes the evaluation of experimental plans. The maximum uptake rate of L-arabinose ($q_{s,ind.,max}$) can be considered a design criterion for mixed feed bioprocesses. Manufacturing processes should be designed to run well below the catabolite repression design criterion to avoid process failures and deviations through substrate accumulation. Using the dynamic method presented here, mixed feed metabolic capabilities can be established in an efficient manor with a minimum of fermentation runs necessary.

Onset of catabolite repression in the pBAD mixed feed system

The function depicted in Figure 6 shows that (at a temperature of 35°C) the overall total uptake of both substrates does not exceed a value of $q_{s,tot.}$ = 1.0 g/g/h. Preferential uptake of D-glucose starts once the total carbon flux through the system exceeds the threshold $q_{s,tot.}$ = 1.0 g/g/h. Approaching this threshold, acetate accumulation was observed, a metabolic regulation typically addressed to an imbalance in the TCA cycle [9]. This indicates, that genes encoding for L-arabinose transport and L-arabinose metabolism are actively down regulated once the metabolic capacity of the TCA cycle is exceeded.

Typical industrial and pharmaceutical fed-batch bioprocesses run at specific uptake rates well below 1 g/g/h. Hence, there is a high process technological freedom for the design of pBAD mixed feed bioprocesses in respect to finding optimal mixed feed ratios. Furthermore, the adaptation time to L-arabinose metabolism were determined to be very fast (<10min) compared to mixed feed bioprocesses based on the AOX promoter (typically > 5 hours) [31].

Error propagation to biomass estimation and control of specific rates

Error on biomass estimation and control of specific substrate uptake rates depend upon i) the biomass concentration at the time of the start of the soft sensor which is measured *via* OD correlation and given as an input and ii) error propagation from on-line and off-line measurements (CO_2/ O_2 quantification in off-gas, feed concentrations etc.) to real time measured rates. Error propagation to real-time measured rates is discussed elsewhere [32].

Conclusion

Mixed feed bioprocesses with assailable inducers provide novel degrees of freedom for the design of innovative new recombinant bioprocesses.

In this contribution we show that in-line FTIR measurements can be used to reliably determine the residual concentration of L-arabinose in a culture broth supernatant. Using this measurements both accumulation and consumption of L-arabinose were followed in real-time and delivered to a first-principle soft sensor which was the prerequisite for independent control of multiple specific uptake rates. To the authors' knowledge, the

chemical information obtained from in-line FTIR measurements was successfully integrated in a first-principle soft-sensor control strategy for the first time.

The presented method has proven its capability of correctly predicting the change in the biomass concentration in the system and controlling independent substrate streams even during dynamic experimentation, prerequisite for the accurate control of specific rates. .

Applying the soft-sensor assisted method for investigation of mixed feed systems allowed the fast determination of physiological boundaries with respect to maximum metabolic capabilities of (catabolite-repressed) mixed substrates and adaption behaviors. Using a combination of pulse and ramp experimentation the results were validated orthogonally and time-dependent effects were ruled out in a single fermentation run. With this method it was possible to determine the maximum specific uptake of the inducing substrate L-arabinose as a function of the specific uptake of the growth substrate D-glucose in a recombinant pBAD expression system with *E. coli* C41. This information is mandatory for mixed feed bioprocess design as well as bioprocess investigation following Design of Experiments.

The presented fast and efficient method for the investigation of physiological bioprocess boundaries can be considered generically applicable for the investigation of microbial mixed feed bioprocess and extendable to other inducing substrates. Therefore it holds potential to become a key process development tool within the emerging field of recombinant mixed feed bioprocesses. Furthermore, the demonstrated combined use of multiple advanced bioprocess development tools (first-principle for bioprocess control strategies, spectroscopic tools, and dynamic experimental strategies) sets new directions within bioprocess development, urging for the adaption and combined use of these powerful tools within routine bioprocess design and bioprocess optimization.

Outlook

The *E. coli* C41 pBAD mixed feed system (L-arabinose/D-glucose) system was identified as a promising candidate for the design of a recombinant *E. coli* mixed feed bioprocesses. The reported biochemical characteristics of the pBAD expression system allow the assumption of a relationship between the specific inducer uptake and recombinant gene expression [23–30], which would allow the design of a recombinant expression system capable of tunable recombinant protein expression. The presented physiological results describing L-arabinose/D-glucose mixed feed capabilities provide the basis for further research in this field.

Acknowledgments

This project was supported by FFG, Land Steiermark, SFG BIRD-C GmbH & Co KG, Kritzendorf, and Morphoplant GmbH, Bochum. Strains and plasmids were gratefully provided by BIRD-C GmbH & Co KG, Kritzendorf and Morphoplant GmbH, Bochum.

References of PART 1

[1] Arnau C, Casas C, Valero F. The effect of glycerol mixed substrate on the heterologous production of a Rhizopus oryzae lipase in *Pichia pastoris* system. Biochem. Eng. J. 2011;57:30–7.

[2] Hellwig S, Emde F, Raven NP, Henke M, Van Der Logt P, Fischer R. Analysis of single-chain antibody production in *Pichia pastoris* using on-line methanol control in fed-batch and mixed-feed fermentations. Biotechnol Bioeng 2001;74:344–52.

[3] Jungo C, Marison I, Von Stockar U. Mixed feeds of glycerol and methanol can improve the performance of *Pichia pastoris* cultures: A quantitative study based on concentration gradients in transient continuous cultures. J Biotechnol 2007;128:824–37.

[4] Zalai D, Dietzsch C, Herwig C, Spadiut O. A dynamic fed batch strategy for a *Pichia pastoris* mixed feed system to increase process understanding. Biotechnol Prog 2012;28:878–86.

[5] Striedner G, Cserjan-Puschmann M, Potschacher F, Bayer K. Tuning the Transcription Rate of Recombinant Protein in Strong *Escherichia coli* Expression Systems through Repressor Titration. Biotechnol Prog 2003;19:1427–32.

[6] Lendenmann U, Snozzi M, Egli T. Kinetics of the simultaneous utilization of sugar mixtures by *Escherichia coli* in continuous culture. Appl Environ Microbiol 1996;62:1493–9.

[7] Harder W, Dijkhuizen L, Postgate JR. Strategies of Mixed Substrate Utilization in Microorganisms [and Discussion]. Phil Trans R Soc Lond B 1982;297:459–80.

[8] Stumm-Zollinger E. Effects of inhibition and repression on the utilization of substrates by heterogeneous bacterial communities. Appl Microbiol 1966;14:654–64.

[9] Herwig C, Von Stockar U. A small metabolic flux model to identify transient metabolic regulations in *Saccharomyces cerevisiae*; Bioprocess. Biosyst. Eng. 2002;24:395–403.

[10] Jungo C, Schenk J, Pasquier M, Marison IW, Von Stockar U. A quantitative analysis of the benefits of mixed feeds of sorbitol and methanol for the production of recombinant avidin with *Pichia pastoris*. J Biotechnol 2007;131:57–66.

[11] Spadiut O, Rittmann S, Dietzsch C, Herwig C. Dynamic process conditions in bioprocess development. Eng. Life Sci. 2013;13:88–101.

[12] Sagmeister P, Wechselberger P, Jazini M, Meitz A, Langemann T. Soft sensor assisted dynamic bioprocess control: Efficient tools for bioprocess development. Chem. Eng. Sci. 2013; Accepted Manuscript

[13] Duboc P, Cascão-Pereira LG, Stockar U von. Identification and control of oxidative metabolism in *Saccharomyces cerevisiae* during transient growth using calorimetric measurements. Biotechnol Bioeng 1998;57:610–9.

[14] Osman JJ, Birch J, Varley J. The response of GS-NS0 myeloma cells to single and multiple pH perturbations. Biotechnol Bioeng 2002;79:398–407.

[15] Wechselberger P, Seifert A, Herwig C. PAT method to gather bioprocess parameters in real-time using simple input variables and first principle relationships. Chem. Eng. Sci. 2010;65:5734–46.

[16] Dietzsch C, Spadiut O, Herwig C. A fast approach to determine a fed batch feeding profile for recombinant *Pichia pastoris* strains. Microb. Cell Fact. 2011;10:85.

[17] Luttmann R, Bracewell DG, Cornelissen G, Gernaey KV, Glassey J, Hass VC, et al. Soft sensors in bioprocessing: A status report and recommendations. Biotechnol J 2012;7:1040–8.

[18] Jobé AM, Herwig C, Surzyn M, Walker B, Marison I, Von Stockar U. Generally applicable fed-batch culture concept based on the detection of metabolic state by on-line balancing. Biotechnol Bioeng 2003;82:627–39.

[19] Van der Heijden RTJM, Heijnen JJ, Hellinga C, Romein B, Luyben KCAM. Linear constraint relations in biochemical reaction systems: I. Classification of the calculability and the balanceability of conversion rates. Biotechnol Bioeng 1994;43:3–10.

[20] Wechselberger P, Sagmeister P, Herwig C. Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes. Bioprocess Biosyst Eng 2012.

[21] Dabros M, Amrhein M, Bonvin D, Marison IW, Von Stockar U. Data reconciliation of concentration estimates from mid-infrared and dielectric spectral measurements for improved on-line monitoring of bioprocesses. Biotechnol. Progr. 2009;25:578–88.

[22] Kornmann H, Valentinotti S, Marison I, Stockar U von. Real-time update of calibration model for better monitoring of batch processes using spectroscopy. Biotechnol Bioeng 2004;87:593–601.

[23] Greenblatt J, Schleif R. Arabinose C Protein: Regulation of the Arabinose Operin in vitro. Nature New Biology 1971;233:166–70.

[24] Schleif R. AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action. FEMS Microbiol. Rev. 2010;34:779–96.

[25] Schleif R, St NC. Regulation of the L -arabinose operon of *Escherichia coli*. Trends Genet 2000;16:559–65.

[26] Johnson CM, Schleif RF. In vivo induction kinetics of the arabinose promoters in *Escherichia coli*. J. Bacteriol. 1995;177:3438–42.

[27] Schleif R. AraC protein: a love-hate relationship. BioEssays. 2003;25:274–82.

[28] Hahn T, Dunn TM, Schleif R. Upstream repression and CRP stimulation of the *Escherichia coli* l-arabinose operon. J Mol Biol 1984;180:61–72.

[29] Miyada CG, Stoltzfus L, Wilcox G. Regulation of the araC gene of *Escherichia coli*: catabolite repression, autoregulation, and effect on araBAD expression. Proc. Nat. Acad. Sci. U.S.A 1984;81:4120–4.

[30] Lobell RB, Schleif RF. AraC-DNA looping: orientation and distance-dependent loop breaking by the cyclic AMP receptor protein. J Mol Biol 1991;218:45–54.

[31] DeLisa MP, Li J, Rao G, Weigand WA, Bentley WE. Monitoring GFP-operon fusion protein expression during high cell density cultivation of *Escherichia coli* using an on-line optical sensor. Biotechnol Bioeng 1999;65:54–64.

[31 Dietzsch C, Spadiut O, Herwig C. 2011. A dynamic method based on the specific substrate uptake rate to set up a feeding strategy for *Pichia pastoris*. Microb. Cell Fact. 10:14.

[32] Wechselberger P, Sagmeister P, Herwig C. 2013. Model-based analysis on the extractability of information from data in dynamic fed-batch experiments. Biotechnol. Prog. 29:285–296.

PART 2

The *E. coli* pBAD mixed feed platform system: Investigation of temperature on mixed feed metabolic capabilities, inclusion body purity and product titer using dynamic methods and physiological design of experiments

Magdalena Kment[1a], Patrick Sagmeister[1a], Andrea Meitz[2], Timo Langemann[2,3] and Christoph Herwig[1]*

[a] these two authors contributed to this work equivalently

* to whom the correspondence should be addressed to

1 Institute of Biochemical Engineering, Vienna University of Technology

2 Research Center of Pharmaceutical Engineering (RCPE) GmbH, Graz

3 BIRD-C GmbH, Kritzendorf, Austria

Keywords: *Christoph Herwig

Gumpendorferstrasse 1a / 166-4

A-1060 Wien, Österreich

Email to: christoph.herwig@tuwien.ac.at

Tel: +43 1 58801 166400

Fax: +43 1 58801 166980

Abstract

Background

Mixed feed bioprocessing is defined by the feeding of two or more carbon sources at a time typically aiming at increased bioprocess performance. The pBAD mixed feed platform system is characterized by the simultaneous feeding and metabolizing of D-glucose and the inducer L-arabinose and was shown to efficiently allow the tuning of recombinant protein production. However, design of pBAD mixed feed processes demand knowledge on the physiological capabilities of the system in respect to catabolite repression of L-arabinose metabolism. Cultivation temperature is one primary parameter for the optimization of recombinant bioprocesses. This contribution aims at the investigation of the impact of cultivation temperature on inclusion body purity and final product titer in the pBAD mixed feed platform system production rhBMP-2 as inclusion bodies.

Results

The physiological capabilities of L-arabinose metabolism as a function of cultivation temperature and uptake of L-glucose were studies using a dynamic pulse method. The maximum uptake of L-arabinose was found to be linearly dependent upon D-glucose uptake and temperature. The results were explained by a simple model: At a defined temperature the culture shows a maximum total specific substrate uptake. Once this is exceeded, D-glucose is consumed preferentially and L-arabinose accumulates. The knowledge obtained from the dynamic experiments was used for the design of a 3 factor design of experiments including D-glucose, L-arabinose and cultivation temperature as factors (16 experimental runs). Evaluation of the results showed a significant impact of temperature on inclusion body purity (IB) and product titer (optimal conditions at 35°C).

Conclusions

Dynamic experiments were efficiently used for the design of a multivariate study of a catabolite repressed mixed feed system and the development of a mechanistic model describing the onset of catabolite repression. For the investigated bioprocess, optimal conditions in respect to inclusion body purity and final product titer were found at 35°C.

Introduction

Escherichia coli (*E. coli*) is one of the most widely used prokaryotic expression system in the biotechnological usage for recombinant protein production [1]. The main advantages of this bacterial system are there growth on inexpensively media and that this expression system reached high growth rates, in respect of industrial biomedical market application [2].

The recombinant protein overexpression in *E. coli* often leads to form insoluble protein aggregates (inclusion bodies) in the cytoplasm which retain native-like conformation and are not active in this state [3–5]. Downstream processing (DSP) of recombinant proteins produced as inclusion bodies demands solubilization and refolding steps, as reviewed elsewhere [6]. Although downstreaming is considered as more costly and complex compared to soluble protein expression [7], inclusion body processes are still the production mode of choice in case i) toxic proteins are to be produced, ii) key factors governing process economics of inclusion body processes are a high yield of recombinant protein product as well as high inclusion body purities. With respect to the following solubilzation and refolding step at the DSP, the inclusion body purity plays an important role in order to low refolding yields which effects on the correctly folding product yield [8].

The product of interest is the recombinant human bone morphogenetic protein-2 (rhBMP-2) which belongs to the highly conserved family of transforming growth factor-β (TGF-β), these owing to its demonstrated osteoinductive capacity in clinical studies, rhBMP-2 (incorporated into a medical device) was granted FDA approval in 2002 [9].

The pBAD mixed feed expression platform

Mixed-feed strategies, that aim to optimize productivity of recombinant processes, have been successfully applied in other model organisms like *P. pastoris* [10], [11]. In this context, mixed substrate feeding, refers to the usage of two substrates that are concomitantly administered to the system, instead of using only one substrate. The *E. coli* pBAD mixed feed system was shown to efficiently allow the tuning of recombinant gene expression on cellular level by means of adjusting the specific L-arabinose uptake rate [12]. In this system, these two substrates are on the one hand D-glucose, which serves as the primary C-source, and on

the other hand L-arabinose, which serves as second C-source and also functions as specific inducer for the ectopic expression of our target protein. The main motivation in this system is the independently control of the transcription rate using a metabloizable carbon source (L-arabinose feed) while supplying sufficient energy to the cell (D-glucose feed).

However, uptake of L-arabinose is submitted to “catabolite repression”, hence the preferential uptake of L-arabinose instead of D-glucose. This well-studied phenomenon [13–19] restricts the co-utilization of D-glucose and L-arabinose in the *E. coli* pBAD mixed feed system. Recently, a dynamic method for the investigation of mixed feed capabilities in recombinant mixed feed systems was reported, that was used to study L-arabinose and D-glucose metabolic capabilities in the pBAD mixed feed expression system [20].

The pBAD mixed process for the production of rhBMP-2 inclusion bodies

As mentioned above rhBMP-2 is produced in *E. coli* as insoluble protein aggregates. Temperature during the recombinant protein expression is reported to have a strong impact on the characteristics of inclusion body formation [21]. Hence, an impact of temperature on inclusion body yields and purity can be expected. It was shown, that at higher temperature the inclusion bodies are denser packed at lower temperature whereby their better protected against proteases [21] and showed higher inclusion body purity yields. In this case the purity of the IB´s is the primary aim and should be high after a homogenization and solubilization step to reach a high percentage of pure inclusion bodies to yield high refolding yields and correct folded active protein [3], [4], [6]. Furthermore, cellular pathways influenced by process temperature are oxidative stress response, energy metabolism, protein folding, RNA/ribosomal biogenesis and amino acid metabolism. Oxidative stress response, energy metabolism and protein folding decrease at lower temperature, while components from RNA/ribosomal biogenesis and amino acid metabolism are indirectly linked to temperature [22].

Design of Experiments

Design of experiment (DoE) is a defined systematic procedure to discover unknown effects, to test or establish a hypothesis, or to illustrate known effects [23]. However, information on physiological process boundaries is necessary for the definition of DoE ranges, hence to span the space which is to be systematically investigated. This is especially challenging in mixed feed processes, since information on mixed feed metabolic capabilities are necessary [20]. Although induced mixed feed metabolic capabilities were already investigated previously [20], induced mixed feed metabolic capabilities as a function of temperature were not investigated so far. However, this information is necessary in case cultivation temperature is included in a multivariate study.

Goals of this study

The primary goal of this contribution is the demonstration of mixed feed experimental design on the basis of the developed method, while the determination of the maximum specific uptake rate of the pBAD inducer and second C-source, L-arabinose at different temperatures, to characterize physiological boundaries of this expression system and to optimize the feeding profile of this biotechnological process utilizing dynamic experimentation. Based on these results, afterwards a design of experiments is implemented to determine the impact of process parameters on inclusion body purity and product titer.

The novelty approaches for the investigated experiments are the i) first demonstration of pBAD mixed feed DoE and process design on the basis of dynamic methods and the ii) first investigation of the impact of temperature on inclusion body purity and product titer in the pBAD mixed feed platform.

It is suggested that the experimental approach presented herein can easily be adapted to investigate the physiological capabilities of other recombinant mixed feed systems (glucose/lactose using the PET-expression system, glycerol/methanol using the AOX promoter), providing the basis for the design and optimization of recombinant mixed feed bioprocesses.

Materials and Methods

Strain

For the biotechnological production of recombinant human bone morphogenetic protein 2 (rhBMP-2) a prokaryotic systems was applied using *E. coli* C41 (F– ompT hsdSB (rB- mB-) gal dcm (DE3); Lucigene, Middleton, WI, USA) strains. These *E. coli* cells carry the plasmid pBK-BMP (23.03.2012 c1), which originates from the plasmid pBAD24 (provided by BIRD-C, Vienna, Austria). The applied *E. coli* C41 strains had an intact L-arabinose metabolism. The overexpression of the target protein was controlled *via* the p_{BAD} promotor, which is specific for its inducer L-arabinose (also acts as secondary carbon source). Additionally, the used expression vector also featured a kanamycin antibiotic resistance cassette.

Media

For cultivation a defined minimal medium as described in DeLisa et al. [24], which was additionally complemented with the amino-glycoside antibiotic kanamycin (0.02 g/l) was used. D-glucose was used as main carbon source (batch medium D-glucose concentration: 20 g/l; fed-batch medium D-glucose concentration: 400 g/l). L-arabinose was added as a pulse (L-arabinose pulse concentration: 2.5 g/l) and also as a mixed feed (different dues of D-glucose and L-arabinose) with a final feed concentration of 400 g/l.

Bioreactor setup

The optimized (with DoE) experiments were carried out with a parallel bioreactor system (DASGIP, Eppendorf, Hamburg Germany) for microbial process development, which consists of four glass bioreactors with a maximum volume of 3 l and a working volume of 2 l . They were equipped with baffles and a three disk stainless steel Rushton impellers stirrer. The DASGIP control software v4.5 revision 230 was used to adjust the process parameters: pH and pO2 (pH probe (Hamilton, Reno, USA) pO2 probe (Mettler Toledo, Greifensee, Switzerland) module DASGIP PH4PO4), temperature and stirrer speed (module DASGIP TC4SC4), aeration (module DASGIP MX4/4). The pH was controlled using the high precision speed controlled miniature peristaltic pump module DASGIP MP8 and 12.5 % NH4OH as

base, feed was added using the same pump module. The reactors were sterilized in an autoclave at 121°C for 20 min.

Fermentation parameters

Dissolved oxygen levels (DO2) were maintained above 40% saturation and the culture was aerated with 1.5 vvm dried air. The pH was kept constant at 7.2 by adding 12.5% NH4OH, which also served as nitrogen source and the stirring speed was kept constant at 1400 rpm. Temperature and feeding profiles were varied and controlled as described in the mixed feed section.

Mixed-feed platform based on dynamic pulse experiments

The investigation of mixed-feed metabolic capabilities was conducted using a dynamic pulse method described in Sagmeister et al. [20]. In short the method is based on a dynamic L-arabinose pulse experiment with alternating temperature shifts in the range of 25, 30 and 35 degree. In order to investigate the simultaneous uptake of D-glucose and the catabolite repressed secondary carbon source L-arabinose at different temperatures, a previously reported dynamic method was employed as described in detail elsewhere [20]. In brief, multiple pulse experiments were combined within one fermentation run. Throughout the process, the specific D-glucose uptake rate is controlled at a defined setpoint *via* a first principle rate based soft-sensor accounting for accumulation of secondary carbon source (L-arabinose) through online-FTIR measurements.

The mixed feed bioprocess was designed as described by Sagmeister et al. [12].

In short, the mixed feed recombinant protein production process based on the above mentioned dynamic method and is composed of three phases: Phase 1: biomass is accumulated within a non-induced D-glucose batch. Phase 2: Fed-batch phase with D-glucose as sole substrate is carried out for biomass accumulation (μ of 0.2 h-1) until a final biomass concentration of 12 g/l was reached. Then, a temperature shift to the temperature set point (25, 30 or 35 °C) is done by following an L-arabinose pulse (concentration of 2.5 g/l) to adapt the *E. coli* cells to L-arabinose metabolism. Phase 3: The L-arabinose pulse is

depleted as detected by the decrease in the off-gas signal, a mixed feed of L-arabinose and D-glucose (in a range of $q_{s\ ara}$. and $q_{s\ gluc}$. 0.05 to 0.2 g/g/h) is started.

Determination of inclusion body purity *via* densitometry evaluation and quantification of rhBMP-2 product titer *via* RP-HPLC

Quality and quantity of the intracellular inclusion body product, rhBMP-2, was analyzed after homogenization and solubilization. Homogenized pellets were analyzed using Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) to assess the purity of the product in respect to host cell proteins. After solubilization of homogenized pellets, reverse-phase high pressure liquid chromatography (RP-HPLC) was used for quantification of rhBMP-2. Both methods, SDS-Page and RP-HPLC, were assessed for their reproducibility and the obtained errors in measurement (for detailed information see appendix).

Results and Discussion

Dynamic investigation of onset of catabolite repression as a function of the D-glucose uptake rate and cultivation temperature

In order to investigate the onset of catabolite repression (preferential consumption of D-glucose instead of L-arabinose) as a function of temperature, multiple dynamic pulse experiments at different setpoints of D-glucose uptake and temperature were conducted (Table 1). On the basis of these results a multivariate MLR model, displaying the maximum specific uptake rate of L-arabinose as a function of the D-glucose uptake rate and cultivation temperature, was established, displayed in contour plot (Figure 4). Model statistics are as follows $R^2 = 0.819$ and $Q^2 = 0.721$ which indicating a significant model.

Table 1: Summary of rates over all pulses of the experiment

Temperature [°C]	$q_{s\ gluc.}$ [g/g/h]	$q_{s\ ara.\ max}$ [g/g/h]	$q_{s\ total}$ [g/g/h]
35	0.26	0.65	0.91
35	0.65	0.35	1.00
35	0.58	0.35	0.93
35	0.63	0.36	0.99
35	0	1.16	1.16
35	0	0.93	0.93
35	0	1.08	1.08
35	0	1.14	1.14
35	0.58	0.58	1.16
35	0.59	0.20	0.79
35	0.63	0.55	1.18
30	0.14	0.33	0.47
30	0.16	0.41	0.57
25	0.05	0.54	0.59
25	0.20	0.26	0.46

The presented results (Table 1) showed a significant impact of the temperature on the total specific uptake rate. Figure 1 displays that the maximum specific L-arabinose uptake rate is a function of temperature and of the specific D-glucose uptake rate. Modeling the maximum specific uptake rate, higher temperature with lower specific D-glucose uptake rate has a positive effect on $q_{s\ ara,max.}$ and reached the maximum at 35 degree and $q_{s\ gluc.} < 0.1$ [g/g/h].

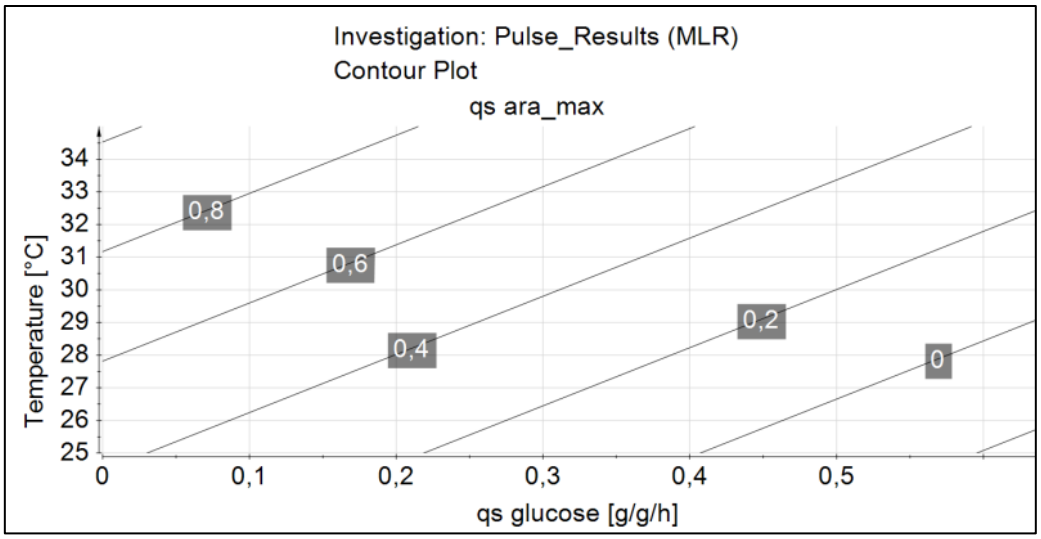

Figure 1: Contour plot: The vertical axis shows the independent variable temperature [°C] in the range 25 to 35 degree. The horizontal axis shows the second independent variable $q_{s\ gluc}$. [g/g/h] in the range 0 to 0.65 [g/g/h]. The lines indicate the iso-response values of $q_{s\ ara\ max}$. [g/g/h].

Summarizing the results from this pulse experiment, it can be concluded that $q_{s\ ara,max}$. is a function of $q_{s\ gluc}$. and cultivation temperature. The response $q_{s\ ara,max}$. reached the maximum at 35 degree and at low $q_{s\ gluc}$.

DoE design

Based on the results of the physiological boundaries of the *E. coli* strain C41 a three factor DoE ($q_{s\ ara}$; $q_{s\ gluc}$ and temperature) was carried out with two responses (inclusion body purity and rhBMP-2 product titer).

The specific L-arabinose uptake rate and specific D-glucose uptake rate, which were controlled using a feed-forward approach. Additionally, induction phase temperature was included in the DoE. On the basis of the pulse results DoE ranges were defined as specific D-glucose and L-arabinose uptake rates of 0.05, 0.125 and 0.20 g/g/h in a temperature range of 25 to 35 °C. In this range no catabolite repression is expected according to the pulse experiments.

Each single fermentation run followed the process design as described in section 2 (Mixed-feed platform based on dynamic pulse experiments).

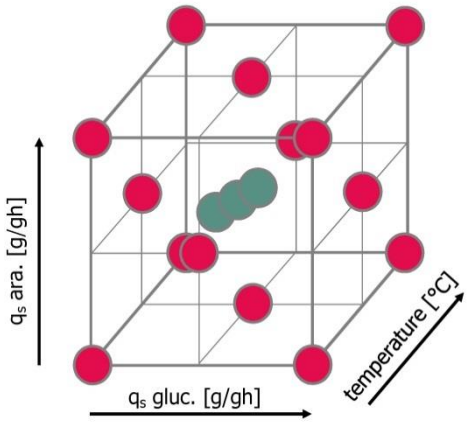

Figure 2: This figure shows an orthogonal factorial design with three factors. The x-axis represents the specific D-glucose uptake rate ($q_{s\ gluc.}$) in the range 0.050; 0.125 and 0.200 [g/g/h], the y-axis represents the specific L-arabinose uptake rate.

Impact of cultivation temperature on inclusion body purity

Quality of the intracellular inclusion body product, rhBMP-2, were analyzed after homogenization and solubilization. Homogenized pellets were analyzed *via* densitometry as described in appendix to assess the purity of the product in respect to host cell proteins as described in the appendix.

The goal was the identification of the interactions between the parameters $q_{s\ ara.}$; $q_{s\ gluc.}$ and temperature on the inclusion body purity. The results of DoE were statistically evaluated using multiple linear regressions (MLR). With the aid of this multivariate statistic data analysis it is possible to discover the impacts on the inclusion body purity. It is shown that inclusion body purity is a function of $q_{s\ gluc.}$ and temperature. The results of the inclusion body purity are visualized in Figure 3 where the impact of the temperature and $q_{s\ gluc.}$ is shown. Model statistics are as follows R^2 = 0.782 and Q^2 = 0.566 which indicating a significant model. The purity optimum is found at 35 degree (Figure 3 A) and Figure 3 B shows the influence of $q_{s\ gluc.}$ at the same temperature.

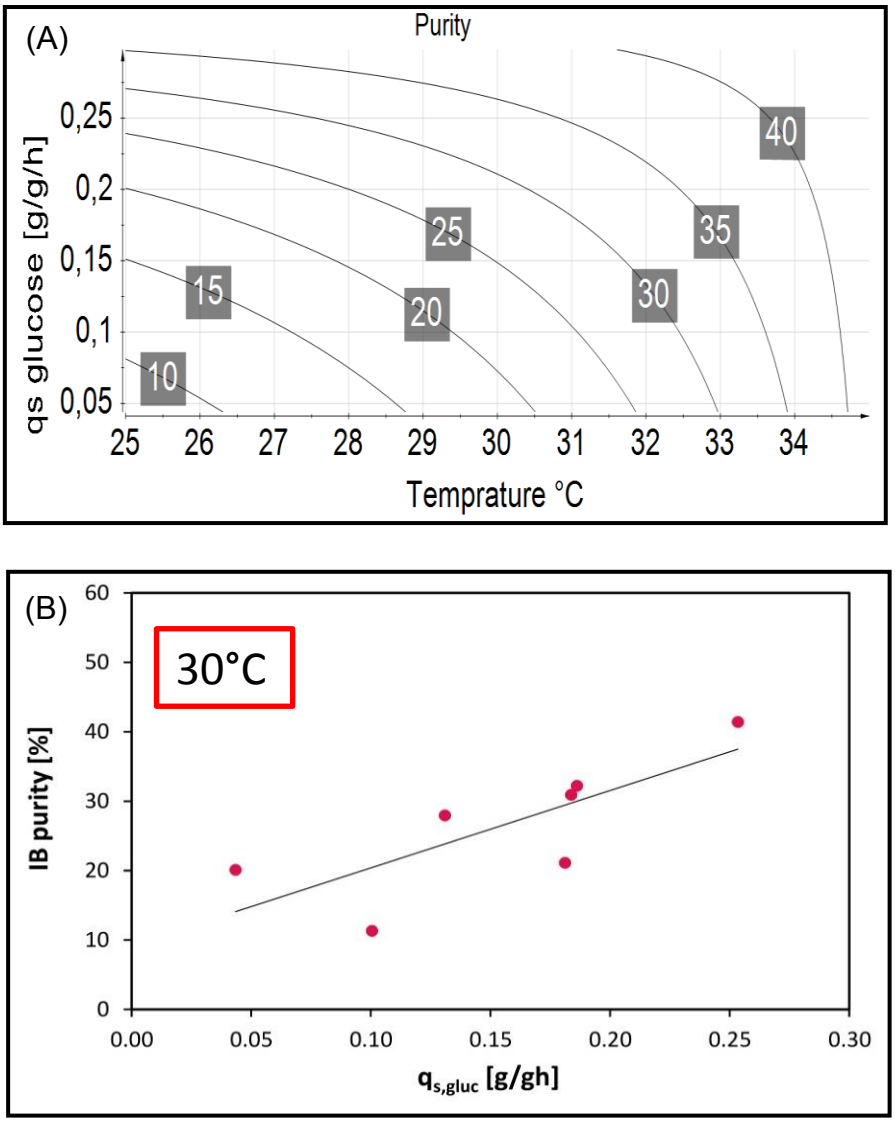

Figure 3 (A-B): Figure A Contour plot: The vertical axis shows the independent variable $q_{s\ gluc.}$ [g/g/h] at a range 0.050 to 0.200 [g/g/h]. The horizontal axis shows the second independent variable temperature [°C] in the range 25 to 35 degree. The lines indicate the iso-response values of purity [%]. This plot demonstrates that the purity is a function of $q_{s\ gluc.}$; temperature and $q_{s\ gluc.}$ * temperature. The response purity reached their optimum at 35 degree. Figure 3 B: the vertical axis represents the inclusion purity [%] and the horizontal axis depicts the specific D-glucose uptake rate ($q_{s\ gluc.}$) [g/g/h] at 30 degree. It is obvious that the higher the $q_{s\ gluc.}$ the higher the inclusion body purity at the same temperature. Model statistics are as follows R^2 = 0.782 and Q^2 = 0.566 which indicating a significant model.

Figure 3 depicts the deviation between the fermentation runs at the same temperature and the impact of $q_{s\ gluc.}$ on the IB purity. The experiment with the highest purity (41.41 %) is a fermentation run with the highest $q_{s\ gluc.}$ (0.254 g/g/h) and also the highest $qs_{\ total}$ (0.409 g/g/h).

In general, it can be said that at the same temperature with increasing $q_{s\ gluc.}$ the inclusion body purity reached higher values. Generally it could be indicated that the temperature and $q_{s\ gluc.}$ are not the only effects on IB purity which could be a reason of different physiological states of the *E. coli* cells, which means that the cells have different expression rates for example. The higher the temperature the higher the expression rates with higher aggregation (IB formation) and higher solubility could be also a reason for higher IB purity.

It can be concluded that both process parameters (temperature and $q_{s\ gluc.}$) and the physiology have an effect on the IB purity.

Temperature is reported to have an effect on the transcription rate , therefore it is likely to assume that IB consistency depends an bioprocess conditions such as modification of the expression rate by inducing regime and temperature [7], [25], [26].

Probably the temperature effect on purity can also be explained by rising to higher refolding yields as reported elsewhere [6].

Impact of cultivation temperature on final product titer

In order to investigate the impact of the investigated process parameters on product titer, RP-HPLC was used to identify and quantify rhBMP-2 in the cell debris pellet (centrifuged fermentation broth).

Furthermore the results of DoE were statistically evaluated using multiple linear regressions (MLR). With this statistically method it is possible to discover the impacts on the rhBMP-2 product titer. It is figured out, that only the temperature has a significant influence on the productivity (R^2 0.765 and Q^2 0.654, indicates that the model is valid). According to Sagmeister et al. (submitted) [12] the product titer also depends on the specific L-arabinose

uptake rate. May a reason that only the temperature is significant toward the presented product titer data could be the short process times and low biomass concentrations.

The black arrowhead in Figure 4 depicts the experiment with the highest rhBMP-2 titer (0.0603 g/g with the experimental settings: $q_{s\ ara}$ 0.286 g/g/h, $q_{s\ gluc}$. 0.071 g/g/h at 35 degree).

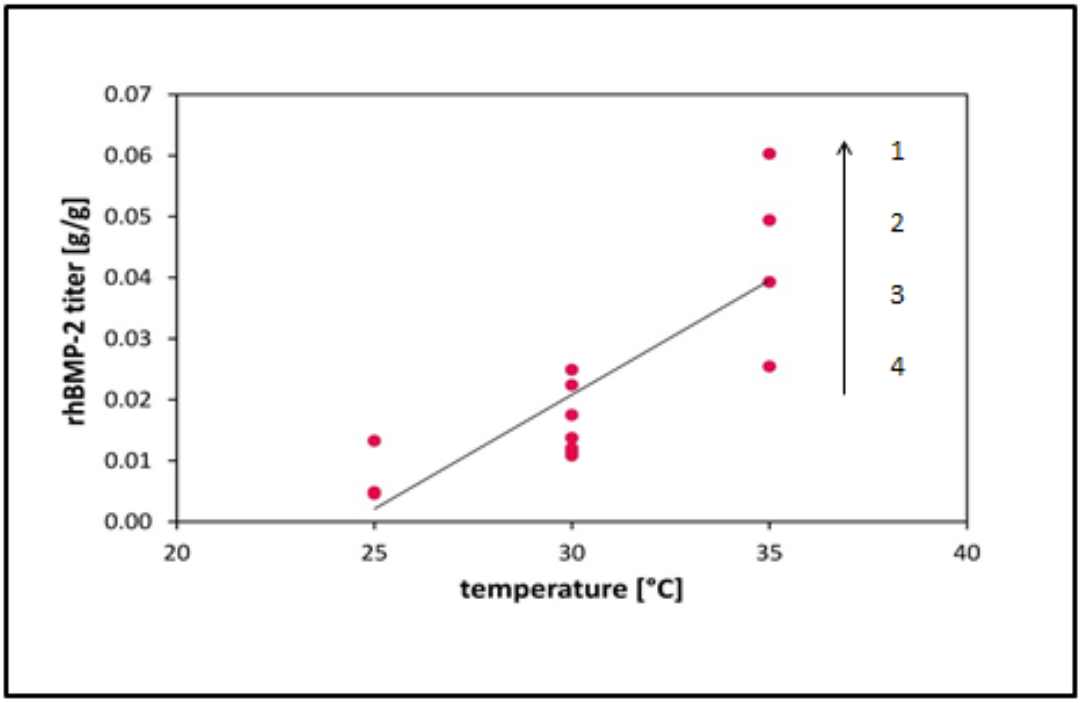

Figure 4: The y-axis depicts the rhBMP-2 product titer [g/g] and the x-axis shows the temperature [°C]. Each single fermentation run with different $q_{s\ ara}$. and $q_{s\ gluc}$. are shown as circles. This plot shows that the higher the temperature the higher the rhBMP-2 titer. The arrow indicates the difference between qs total of the single fermentation runs. Circle 1: $q_{s\ tot.}$ = 0.357 g/g/h (rhBMP-2 titer = 0.0603 g/g); 2: $q_{s\ tot.}$ = 0.372 g/g/h (rhBMP-2 titer = 0.0494 g/g); 3: $q_{s\ tot.}$ = 0.302 g/g/h (rhBMP-2 titer = 0.0393 g/g); 4: $q_{s\ tot.}$ = 0.562 g/g/h (rhBMP-2 titer = 0.0254 g/g).

The major finding of the temperature depending productivity is an output which is also reported elsewhere [27]. It is also pronounced that at reduced growth rates the productivity is higher according to the findings that the expression rate follows a Monod-Kinetic.

As illustrated at Figure 4 the temperature influence the yield of rhBMP-2 titer. The observed deviations of the titer at the same temperature can most probably be explained through the different due to the percentage of different q_s values. This is clearly shown at 35 degree where the experiment with the highest titer has not the highest $q_{s\ total}$, which could be a reason of other impacts on the productivity like q_s L-arabinose.

The experiment with the highest $q_{s\ total}$ (0.562 g/g/h, experimental settings: $q_{s\ ara}$. 0.281 g/g/h and $q_{s\ gluc}$. 0.281 g/g/h) at 35 degree is the circle at the beginning of the arrow (Figure 3) which is in agreement that at high growth rates the protein expression is limited and causes in to the metabolic load [28], [29].

Process results

As an illustration one single fermentation process at 30°C induction temperature and an exponential induction feed profile is given in Figure 5.

In order to validate the analyses, a C-balance was carried out. A C-balance of 100% denotes a closed system. During the induction phase of this experiment a C-balance of 96 % was determined. The observed errors can most probably be explained through an error in the biomass determination.

Furthermore, we determined the respiratory quotient (RQ) of our system. RQ provides information on the metabolic state of the cell by relating the amount of produced CO_2 to the amount of consumed O_2. In our experiments RQs of 0.99 was measured, close to the expected RQ on glucose [1]. With respect to the analysis of process-specific yields, a “yield coefficient” (Yx/s) of 0.44 Cmol/Cmol was determined during the induction phase.

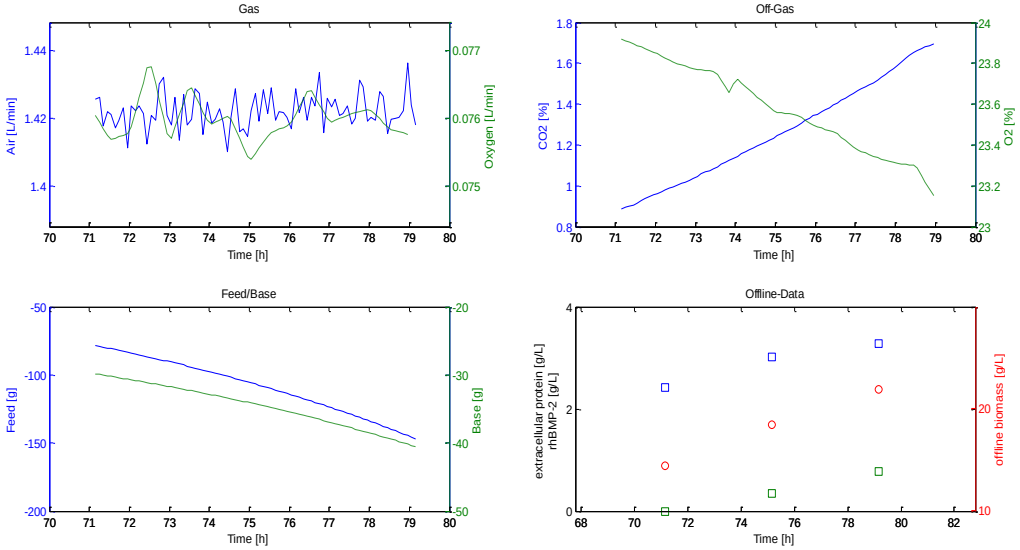

Figure 5 (A-D): Fermentation at 30 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle)

Conclusion

Identification of the physiological boundaries for the pBAD mixed-feed system:

- Maximum specific uptake rate of the pBAD inducer and second C-source, L-arabinose detected at $q_{s\ total}$ 1.01 [g/g/h] at high temperature
- Detection of maximum arabinose uptake rates as a function of q_s and temperature

  Understanding of the interaction between process parameters ($q_{s\ ara}$, $q_{s\ gluc.}$ and temperature):

- higher purity [%] with higher $q_{s\ gluc.}$ [g/g/h] at the same temperature

  → Optimum found at 35 °C

- significant impact of cultivation temperature on product titer

  → higher titers at high cultivation temperatures (35°C)

References of PART 2

[1] M. Papagianni, “Recent advances in engineering the central carbon metabolism of industrially important bacteria.,” Microbial cell factories, vol. 11, p. 50, Jan. 2012.

[2] G. Hannig and S. C. Makrides, “Strategies for optimizing heterologous protein expression in *Escherichia coli*.,” Trends in biotechnology, vol. 16, no. 2, pp. 54–60, Feb. 1998.

[3] A. Mitraki, “protein folding intermediates and inclusion body formation,” Nature Biotechnology, vol. Vol 7, 1989.

[4] L. Wang, “Towards revealing the structure of bacterial inclusion bodies.,” Prion, vol. 3, no. 3, pp. 139–45, 2009.

[5] P. Gatti-Lafranconi, A. Natalello, D. Ami, S. M. Doglia, and M. Lotti, “Concepts and tools to exploit the potential of bacterial inclusion bodies in protein science and biotechnology.,” The FEBS journal, vol. 278, no. 14, pp. 2408–18, Jul. 2011.

[6] B. Batas, C. Schiraldi, and J. B. Chaudhuri, “Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography.,” Journal of biotechnology, vol. 68, no. 2–3, pp. 149–58, Feb. 1999.

[7] S. Ventura and A. Villaverde, “Protein quality in bacterial inclusion bodies.,” Trends in Biotechnology, vol. 24, no. 4, pp. 179–85, Apr. 2006.

[8] A. M. Sandén, I. Prytz, I. Tubulekas, C. Förberg, H. Le, A. Hektor, P. Neubauer, Z. Pragai, C. Harwood, A. Ward, A. Picon, J. T. De Mattos, P. Postma, A. Farewell, T. Nyström, S. Reeh, S. Pedersen, and G. Larsson, “Limiting factors in *Escherichia coli* fed-batch production of recombinant proteins.,” Biotechnology and bioengineering, vol. 81, no. 2, pp. 158–66, Jan. 2003.

[9] “http://www.accessdata.fda.gov/cdrh_docs/pdf/P000058b.pdf.” .

[10] M. C. Anjou and A. J. Daugulis, “Mixed-feed exponential feeding for fed-batch culture of recombinant methylotrophic yeast,” Biotechnology Letters, pp. 341–346, 2000.

[11] D. Zalai, C. Dietzsch, C. Herwig, and O. Spadiut, “A dynamic fed batch strategy for a *Pichia pastoris* mixed feed system to increase process understanding.,” Biotechnology progress, vol. 28, no. 3, pp. 878–86, May 2012.

[12] P. Sagmeister, C. Schimek, A. Meitz, C. Herwig, and O. Spadiut, “Design and multivariate investigation of an E . coli pBAD mixed feed fed-batch expression platform capable of tunable recombinant protein expression,” pp. 1–20.

[13] J. Greenblatt and R. Schleif, “Arabinose C Protein: Regulation of the Arabinose Operin in vitro,” Nature New Biology, vol. 233, pp. 166–170, 1971.

[14] R. Schleif, “AraC protein, regulation of the l-arabinose operon in *Escherichia coli*, and the light switch mechanism of AraC action.,” FEMS Microbiology Reviews, vol. 34, no. 5, pp. 779–96, Sep. 2010.

[15] R. Schleif and N. C. St, “Regulation of the L -arabinose operon of *Escherichia coli* Over forty years of research on the L -arabinose operon of *Escherichia coli* have provided insights into the,” Trends in Genetics, vol. 16, no. 2, pp. 559–565, 2000.

[16] R. Schleif, “AraC protein: a love-hate relationship.,” BioEssays : news and reviews in molecular, cellular and developmental biology, vol. 25, no. 3, pp. 274–82, Mar. 2003.

[17] C. M. Johnson and R. F. Schleif, “In vivo induction kinetics of the arabinose promoters in *Escherichia coli*.,” Journal of Bacteriology, vol. 177, no. 12, pp. 3438–42, Jun. 1995.

[18] T. M. Dunn and R. Schleif, “Deletion analysis of the *Escherichia coli* ara PC and P_{BAD} promoters,” J. Mol. Biol., vol. 180, no. 1, pp. 201–204, 1984.

[19] D. H. Lee and R. Schleif, “Repression of the araBAD promoter from araO1,” J. Mol. Biol., vol. 224, no. 4, pp. 335–341, 1992.

[20] P. Sagmeister, M. Kment, P. Wechselberger, A. Meitz, T. Langemann, and C. Herwig, “Soft-sensor assisted dynamic investigation of mixed feed bioprocesses,” pp. 1–26.

[21] N. S. de Groot and S. Ventura, “Effect of temperature on protein quality in bacterial inclusion bodies.,” FEBS letters, vol. 580, no. 27, pp. 6471–6, Nov. 2006.

[22] M. Carnicer, A. Ten Pierick, J. van Dam, J. J. Heijnen, J. Albiol, W. van Gulik, and P. Ferrer, “Quantitative metabolomics analysis of amino acid metabolism in recombinant *Pichia pastoris* under different oxygen availability conditions.,” Microbial cell factories, vol. 11, no. 1, p. 83, Jan. 2012.

[23] M. Shivhare and G. Mccreath, “Practical Considerations for DoE Implementation in Quality By Design,” BioProcess International, pp. 22–30, 2010.

[24] M. P. DeLisa, J. Li, G. Rao, W. Weigand, and W. E. Bentley, “Monitoring GFP-operon fusion protein expression during high cell density cultivation of *Escherichia coli* using an on-line optical sensor.,” Biotechnology and Bioengineering, vol. 65, no. 1, pp. 54–64, Oct. 1999.

[25] D. Groot and S. Ventura, “Effect of temperature on protein quality in bacterial inclusion bodies ´,” vol. 580, pp. 6471–6476, 2006.

[26] A. Villaverde and M. M. Carrió, “Protein aggregation in recombinant bacteria: biological role of inclusion bodies.,” Biotechnology letters, vol. 25, no. 17, pp. 1385–95, Sep. 2003.

[27] J. Rodriguez, M. Spearman, T. Tharmalingam, K. Sunley, C. Lodewyks, N. Huzel, and M. Butler, “High productivity of human recombinant beta-interferon from a low-temperature perfusion culture.,” Journal of biotechnology, vol. 150, no. 4, pp. 509–18, Dec. 2010.

[28] D. Korz, E. A. Sanders, and A. Ror, “High cell density cultivation of *Escherichia coli* at controlled specific growth rate,” vol. 20, 1991.

[29] B. Adwn and B. R. Glick, “Pergamon,” vol. 13, no. 2, pp. 247–261, 1995.

Conclusion and Outlook

Summary of main findings:

In order to identify physiological boundaries of the pBAD mixed-feed system expressing rhBMP-2, dynamic experiments (pulse and ramps) and a DoE (3-factor CCF design) were conducted. The results from the pulse and ramp experiments served as data-driven inputs for the consecutive planning and design of DoE experiments. Specifically, the experiments were designed below $q_{s\ ara\ max.}$ and the adaptation time (< 10 mins.) were considered.

The main findings from these two approaches can be summarized as follows:

- Dynamic experimentation allowed the detection of the maximum L-arabinose uptake rates as a function of q_s D-glucose and temperature.

In order to understand the interaction between physiological process parameters ($q_{s\ ara}$, $q_{s\ gluc}$. and temperature), the application of DoE highlighted that:

- At 30°C temperature IB purity is positively correlated with $q_{s\ gluc}$. The higher $q_{s\ gluc.}$ (maximum: 0.25 [g/g/h]), the higher the IB purity [%] (maximum: 41.4 %).
- The highest IB purity (48.4 %) was found at a temperature of 35 °C
- Furthermore, the product titer was also found to be positively associated with the cultivation temperature. Higher product titers (maximum: 1.65 g/L) were achieved at higher temperatures (35°C).

Limitations of the current work and recommendations for further improvement

The underlying thesis provides important contributions to the upstream processing for the production of rhBMP-2 in an *E. coli* mixed-feed system. However, despite the encouraging results, the current work also suffers from some experimental limitations that offer place for improvement in future work.

- **Longer process times to reach higher productivity and higher biomass yields during longer fed-batch processes**

The current work aimed at elucidating the influence of the physiological process parameters ($q_{s\ gluc}$, $q_{s\ ara}$, temperature) on product quality (i.e. IB purity and product titer). The importance of the cultivation temperature and $q_{s\ gluc}$ was established. However, no association between the product parameters and $q_{s\ ara}$ could be found (it must be noted that others found an impact of $q_{s\ ara}$ on product titer [16] [88]). A potential reason for this, might be that the process times in the underlying fermentation experiments were in general too short to detect the influence of $q_{s\ ara}$ on product titer. Hence, future experiments might consider the use of longer process times.

- **Additional validation experiment for fine-tuning of $q_{s\ gluc}$ and temperature**

In this thesis the positive association of $q_{s\ gluc}$ and cultivation temperature was elaborated. While DoE was a very helpful tool, because it significantly reduced the number of experiments, consumption of time and costs during formulation development. In order to confirm the relationship between these physiological process parameters in more detail, additional a validation experiment (of the optimum space) to elucidate the identified design space of these parameters in more detail could be conducted.

- **DoE: feed-forward feeding strategy vs. soft-sensors**

In the current work, the QbD approached DoE-guided fermentations were monitored using a feed-forward feeding strategy. Soft-sensor assisted control of the process was applied during the dynamic experiments but not for DoE. Since, the results from dynamic experiments outlined the usability of soft-sensors, the application of these process analytical tools (PAT) should also be considered for DoE experiments. Mainly because of the fact that soft-sensors harbor the potential to adapt feeding rates to different metabolic states.

References

[1] S. Lawrence, “Billion dollar babies,” *Nature Biotechnology*, vol. 25, no. 4, pp. 380–382, 2008.

[2] G. Walsh, “Biopharmaceutical Benchmarks 2010,” *Nature Biotechnology*, vol. 28, no. 9, pp. 917–24, Sep. 2010.

[3] S. a Berkowitz, J. R. Engen, J. R. Mazzeo, and G. B. Jones, “Analytical tools for characterizing biopharmaceuticals and the implications for biosimilars,” *Nature Reviews Drug Discovery*, vol. 11, no. 7, pp. 527–40, Jan. 2012.

[4] ICH Q8, “Pharmaceutical Development Q8 (R2),” vol. 8, no. August. 2009.

[5] U. von Stockar, S. Valentinotti, I. Marison, C. Cannizzaro, and C. Herwig, “Know-how and know-why in biochemical engineering,” *Biotechnology Advances*, vol. 21, no. 5, pp. 417–430, Aug. 2003.

[6] P. Wechselberger, A. Seifert, and C. Herwig, “PAT method to gather bioprocess parameters in real-time using simple input variables and first principle relationships,” *Chemical Engineering Science*, vol. 65, no. 21, pp. 5734–5746, Nov. 2010.

[7] A. S. Rathore and H. Winkle, “Quality by design for Biopharmaceuticals,” *Nature Biotechnology*, vol. 27, no. 1, pp. 26–34, 2009.

[8] A. S. Rathore, “Roadmap for implementation of quality by design (QbD) for biotechnology products.,” *Trends in Biotechnology*, vol. 27, no. 9, pp. 546–53, Sep. 2009.

[9] E. K. Read, R. B. Shah, B. S. Riley, J. T. Park, K. a Brorson, and a S. Rathore, “Process analytical technology (PAT) for biopharmaceutical products: Part I. concepts and applications.,” *Biotechnology and Bioengineering*, vol. 105, no. 2, pp. 285–95, Feb. 2010.

[10] T. Makino, G. Skretas, and G. Georgiou, “Strain engineering for improved expression of recombinant proteins in bacteria.,” *Microbial Cell Factories*, vol. 10, no. 1, p. 32, Jan. 2011.

[11] L. Wang, “Towards revealing the structure of bacterial inclusion bodies.,” *Prion*, vol. 3, no. 3, pp. 139–45, 2009.

[12] B. Batas, C. Schiraldi, and J. B. Chaudhuri, “Inclusion body purification and protein refolding using microfiltration and size exclusion chromatography.,” *Journal of Biotechnology*, vol. 68, no. 2–3, pp. 149–58, Feb. 1999.

[13] D. Gangadharan, K. M. Nampoothiri, and A. Pandey, “a-Amylase Production by Bacillus amyloliquefaciens Using Agro Wastes as Feed Stock,” *Food Technol. Biotechnol.*, vol. 49, no. 3, pp. 336–340, 2011.

[14] O. Spadiut, S. Rittmann, C. Dietzsch, and C. Herwig, “Dynamic process conditions in bioprocess development,” *Engineering in Life Sciences*, vol. 13, no. 1, pp. 88–101, Jan. 2013.

[15] C. Dietzsch, O. Spadiut, and C. Herwig, “A dynamic method based on the specific substrate uptake rate to set up a feeding strategy for Pichia pastoris.,” *Microbial Cell Factories*, vol. 10, no. 1, p. 14, Jan. 2011.

[16] D. Zalai, C. Dietzsch, C. Herwig, and O. Spadiut, “A dynamic fed batch strategy for a Pichia pastoris mixed feed system to increase process understanding.,” *Biotechnology Progress*, vol. 28, no. 3, pp. 878–86, May 2012.

[17] P. Gatti-Lafranconi, A. Natalello, D. Ami, S. M. Doglia, and M. Lotti, “Concepts and tools to exploit the potential of bacterial inclusion bodies in protein science and biotechnology.,” *The FEBS journal*, vol. 278, no. 14, pp. 2408–18, Jul. 2011.

[18] D. P. Clark and N. J. Pazdernik, *Molekulare Biotechnologie*. Heidelberg: Spektrum Akademischer Verlag, 2009.

[19] H. E. Kubitschek, “Cell volume increase in Escherichia coli after shifts to richer media,” *Journal of Bacteriology*, vol. 172, no. 1, pp. 94–101, Jan. 1990.

[20] G. Margreiter, P. Messner, K. D. Caldwell, and K. Bayer, “Size characterization of inclusion bodies by sedimentation field-flow fractionation.,” *Journal of Biotechnology*, vol. 138, no. 3–4, pp. 67–73, Nov. 2008.

[21] M. Papagianni, “Recent advances in engineering the central carbon metabolism of industrially important bacteria.,” *Microbial Cell Factories*, vol. 11, p. 50, Jan. 2012.

[22] G. W. Luli and W. R. Strohl, “Comparison of growth, acetate production, and acetate inhibition of Escherichia coli strains in batch and fed-batch fermentations.,” *Applied and Environmental Microbiology*, vol. 56, no. 4, pp. 1004–11, Apr. 1990.

[23] M. a Eiteman and E. Altman, “Overcoming acetate in Escherichia coli recombinant protein fermentations.,” *Trends in Biotechnology*, vol. 24, no. 11, pp. 530–6, Nov. 2006.

[24] M. Ignatova and V. Lyubenova, “Control of class bioprocess using on-line information of intermediate metabolite production and consumption rates,” *Food Technology*, vol. XI, no. 1, pp. 3–16, 2007.

[25] S. Y. Lee, “High cell-density culture of Escherichia coli.,” *Trends in Biotechnology*, vol. 14, no. 3, pp. 98–105, Mar. 1996.

[26] R. Schleif, “AraC protein, regulation of the l-arabinose operon in Escherichia coli, and the light switch mechanism of AraC action.,” *FEMS Microbiology Reviews*, vol. 34, no. 5, pp. 779–96, Sep. 2010.

[27] L. Gielow, M. Largen, and E. Englesberg, “Initiator constitutive mutants of the L-arabinose operon (OIBAD) of Escherichia coli B/r,” *Genetics*, vol. 69, pp. 289–302, 1971.

[28] J. Colomé, G. Wilcox, and E. Englesberg, “Constitutive mutations in the controlling site region of the araBAD operon of Escherichia coli B/r that decrease sensitivity to catabolite repression.,” *Journal of Bacteriology*, vol. 129, no. 2, pp. 948–58, Mar. 1977.

[29] D. Selection and O. F. L. Negative, “Direct selection of l-arabinose negative,” *Genetics*, vol. 47, no. 2, 1962.

[30] D. E. Sheppard and E. Englesberg, “Further evidence for positive control of the l-arabinose system by gene araC,” *J. Mol. Biol.*, vol. 25, no. 3, pp. 443–454, 1967.

[31] R. B. Helling and R. Weinberg, “Complement studies of arabinose genes in Escherichia Coli,” *Genetics*, vol. 487, no. October, pp. 1397–1410, 1963.

[32] J. Greenblatt and R. Schleif, “Arabinose C Protein: Regulation of the Arabinose Operon in vitro,” *Nature New Biology*, vol. 233, pp. 166–170, 1971.

[33] R. Schleif and N. C. St, “Regulation of the L -arabinose operon of Escherichia coli,” *Trends in Genetics*, vol. 16, no. 2, pp. 559–565, 2000.

[34] C. M. Johnson and R. F. Schleif, “In vivo induction kinetics of the arabinose promoters in Escherichia coli.,” *Journal of Bacteriology*, vol. 177, no. 12, pp. 3438–42, Jun. 1995.

[35] R. Schleif, “AraC protein: a love-hate relationship.,” *BioEssays : news and reviews in molecular, cellular and developmental biology*, vol. 25, no. 3, pp. 274–82, Mar. 2003.

[36] T. M. Dunn and R. Schleif, “Deletion analysis of the Escherichia coli ara PC and PBAD promoters,” *J. Mol. Biol.*, vol. 180, no. 1, pp. 201–204, 1984.

[37] D. H. Lee and R. Schleif, “Repression of the araBAD promoter from araO1,” *J. Mol. Biol.*, vol. 224, no. 4, pp. 335–341, 1992.

[38] R. B. Lobell and R. F. Schleif, “AraC-DNA looping: orientation and distance-dependent loop breaking by the cyclic AMP receptor protein.,” *J. Mol. Biol.*, vol. 218, no. 1, pp. 45–54, Mar. 1991.

[39] T. Hahn, T. M. Dunn, and R. Schleif, “Upstream repression and CRP stimulation of the Escherichia coli l-arabinose operon,” *J. Mol. Biol.*, vol. 180, no. 1, pp. 61–72, 1984.

[40] C. G. Miyada, L. Stoltzfus, and G. Wilcox, “Regulation of the araC gene of Escherichia coli: catabolite repression, autoregulation, and effect on araBAD expression.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 81, no. 13, pp. 4120–4, Jul. 1984.

[41] A. Mitraki, “protein folding intermediates and inclusion body formation,” *Nature Biotechnology*, vol. Vol 7, 1989.

[42] V. R. Agashe, S. Guha, H.-C. Chang, P. Genevaux, M. Hayer-Hartl, M. Stemp, C. Georgopoulos, F. U. Hartl, and J. M. Barral, “Function of trigger factor and DnaK in multidomain protein folding: increase in yield at the expense of folding speed.,” *Cell*, vol. 117, no. 2, pp. 199–209, Apr. 2004.

[43] V. P. Pigiet and B. J. Schuster, “Thioredoxin-catalyzed refolding of disulfide-containing proteins.,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 20, pp. 7643–7, Oct. 1986.

[44] E. García-Fruitós, “Inclusion bodies: a new concept.,” *Microbial Cell Factories*, vol. 9, no. 1, p. 80, Jan. 2010.

[45] D. Groot and S. Ventura, “Effect of temperature on protein quality in bacterial inclusion bodies,” *FEBS Letters published by Elsevier*, vol. 580, pp. 6471–6476, 2006.

[46] Š. Peternel, V. Gaberc-porekar, and R. Komel, “Bacterial Growth Conditions Affect Quality of GFP Expressed Inside Inclusion Bodies,” *Acta Chim. Slov*, vol. 56, pp. 860–867, 2009.

[47] M. Liovic, M. Ozir, A. B. Zavec, S. Peternel, R. Komel, and T. Zupancic, “Inclusion bodies as potential vehicles for recombinant protein delivery into epithelial cells.,” *Microbial Cell Factories*, vol. 11, no. 1, p. 67, Jan. 2012.

[48] E. García-Fruitós, N. González-Montalbán, M. Morell, A. Vera, R. M. Ferraz, A. Arís, S. Ventura, and A. Villaverde, “Aggregation as bacterial inclusion bodies does not imply inactivation of enzymes and fluorescent proteins.,” *Microbial Cell Factories*, vol. 4, p. 27, Sep. 2005.

[49] B. R. . and Deutscher.M.P, “Guide to Protein Purification,” *Methodes in Enzymology*, vol. 463, 2009.

[50] I. Vigentini, A. Merico, M. L. Tutino, C. Compagno, and G. Marino, “Optimization of recombinant human nerve growth factor production in the psychrophilic Pseudoalteromonas haloplanktis.,” *Journal of bBotechnology*, vol. 127, no. 1, pp. 141–50, Dec. 2006.

[51] P. Li, A. Anumanthan, X.-G. Gao, K. Ilangovan, V. V. Suzara, N. Düzgüneş, and V. Renugopalakrishnan, “Expression of Recombinant Proteins in Pichia Pastoris,” *Applied Biochemistry and Biotechnology*, vol. 142, no. 2, pp. 105–124, Apr. 2007.

[52] C. Noguère, A. M. Larsson, J. Guyot, and C. Bignon, “Fractional factorial approach combining 4 Escherichia coli strains , 3 culture media , 3 expression temperatures and 5 N-terminal fusion tags for screening the soluble expression of recombinant proteins,” *Protein Expression and Purification by Elsevier*, vol. 84, pp. 204–213, 2012.

[53] J. M. Song, Y. J. An, M. H. Kang, Y.-H. Lee, and S.-S. Cha, “Cultivation at 6-10°C is an effective strategy to overcome the insolubility of recombinant proteins in Escherichia coli.,” *Protein Expression and Purification by Elsevier*, vol. 82, no. 2, pp. 297–301, Apr. 2012.

[54] P. Bird, “Production of recombinant serpins in Escherichia coli,” *Methods by Elsevier*, vol. 32, no. 2, pp. 169–176, Feb. 2004.

[55] G. Georgiou and P. Valaxt, “Expression of correctly folded proteins in Escherichia coil,” *Current Opinion in Biotechnology*, vol. 7, pp. 190–197, 1996.

[56] H. J. Ihm, S.-J. Yang, J.-W. Huh, S. Y. Choi, and S.-W. Cho, “Soluble expression and purification of synthetic human bone morphogenetic protein-2 in Escherichia coli.,” *BMB reports*, vol. 41, no. 5, pp. 404–7, May 2008.

[57] J. J. Greene, “Recombinant Gene Expression,” *Methodes in Molecular Biology*, vol. 267.

[58] M. Dragosits, J. Stadlmann, J. Albiol, K. Baumann, M. Maurer, B. Gasser, M. Sauer, F. Altmann, P. Ferrer, D. Mattanovich, and B. Cerdanyola, “The Effect of Temperature on the Proteome of Recombinant Pichia pastoris research articles,” *Journal of Proteome Research*, vol. 8, pp. 1380–1392, 2009.

[59] E. Rodríguez-Carmona, O. Cano-Garrido, M. Dragosits, M. Maurer, A. Mader, R. Kunert, D. Mattanovich, A. Villaverde, and F. Vázquez, “Recombinant Fab expression and secretion in Escherichia coli continuous culture at medium cell densities: Influence of temperature,” *Process Biochemistry*, vol. 47, no. 3, pp. 446–452, Mar. 2012.

[60] C. S. Shin, M. S. Hong, C. S. Bae, and J. Lee, “Enhanced production of human mini-proinsulin in fed-batch cultures at high cell density of Escherichia coli BL21(DE3)[pET-3aT2M2].,” *Biotechnology Progress*, vol. 13, no. 3, pp. 249–57, 1997.

[61] A. K. Patra, R. Mukhopadhyay, R. Mukhija, A. Krishnan, L. C. Garg, and a K. Panda, “Optimization of inclusion body solubilization and renaturation of recombinant human growth hormone from Escherichia coli.,” *Protein Expression and Purification*, vol. 18, no. 2, pp. 182–92, Mar. 2000.

[62] B. Bragdon, O. Moseychuk, S. Saldanha, D. King, J. Julian, and A. Nohe, “Bone morphogenetic proteins: a critical review.,” *Cellular Signalling*, vol. 23, no. 4, pp. 609–20, Apr. 2011.

[63] D. Chen, M. Zhao, and G. R. Mundy, “Bone morphogenetic proteins.,” *Growth Factors*, vol. 22, no. 4, pp. 233–41, Dec. 2004.

[64] C. Scheufler, W. Sebald, and M. Hu, “Crystal Structure of Human Bone Morphogenetic Ê Resolution Protein-2 at 2 . 7 A,” *J. Mol. Biol.*, vol. 287, pp. 103–115, 1999.

[65] S. von Einem, E. Schwarz, and R. Rudolph, “A novel TWO-STEP renaturation procedure for efficient production of recombinant BMP-2.,” *Protein Expression and Purification*, vol. 73, no. 1, pp. 65–9, Sep. 2010.

[66] “http://www.accessdata.fda.gov/cdrh_docs/pdf/P000058b.pdf.” .

[67] M. C. Anjou and A. J. Daugulis, “Mixed-feed exponential feeding for fed-batch culture of recombinant methylotrophic yeast,” *Biotechnology Letters*, pp. 341–346, 2000.

[68] C. Jungo, J. Schenk, M. Pasquier, I. W. Marison, and U. von Stockar, “A quantitative analysis of the benefits of mixed feeds of sorbitol and methanol for the production of recombinant avidin with Pichia pastoris,” *Journal of Biotechnology*, vol. 131, no. 1, pp. 57–66, Aug. 2007.

[69] C. Jungo, I. Marison, and U. von Stockar, “Mixed feeds of glycerol and methanol can improve the performance of Pichia pastoris cultures: A quantitative study based on concentration gradients in transient continuous cultures,” *Journal of Biotechnology*, vol. 128, no. 4, pp. 824–837, Mar. 2007.

[70] E. Bäcklund and M. Sc, “Impact of glucose uptake rate on recombinant protein production in Escherichia coli,” School of Biotechnology, 2011.

[71] G. Striedner, M. Cserjan-Puschmann, F. Pötschacher, K. Bayer, and F. Potschacher, “Tuning the transcription rate of recombinant protein in strong Escherichia coli expression systems through repressor titration.,” *Biotechnology Progress*, vol. 19, no. 5, pp. 1427–32, 2003.

[72] Z. Amribt, H. Niu, and P. Bogaerts, “Macroscopic modelling of overflow metabolism and model based optimization of hybridoma cell fed-batch cultures,” *Biochemical Engineering Journal*, vol. 70, pp. 196–209, Jan. 2013.

[73] B. Sonnleitner and O. Käpelli, “Growth of Saccharomyces cerevisiae Is Controlled by Its Limited Respiratory Capacity: Formulation and Verification of a Hypothesis,” in *Biotechnology and Bioengineering*, 1986, pp. 927–937.

[74] C. Herwig, I. Marison, and U. von Stockar, “On-line stoichiometry and identification of metabolic state under dynamic process conditions.,” *Biotechnology and Bioengineering*, vol. 75, no. 3, pp. 345–54, Nov. 2001.

[75] P. Wechselberger, P. Sagmeister, H. Engelking, T. Schmidt, J. Wenger, and C. Herwig, “Efficient feeding profile optimization for recombinant protein production using physiological information.,” *Bioprocess and Biosystems Engineering*, Jun. 2012.

[76] N. K. Khatri and F. Hoffmann, “Oxygen-limited control of methanol uptake for improved production of a single-chain antibody fragment with recombinant Pichia pastoris.,” *Applied Microbiology and Biotechnology*, vol. 72, no. 3, pp. 492–8, Oct. 2006.

[77] N. K. Khatri and F. Hoffmann, “Impact of methanol concentration on secreted protein production in oxygen-limited cultures of recombinant Pichia pastoris.,” *Biotechnology and Bioengineering*, vol. 93, no. 5, pp. 871–9, May 2006.

[78] K. Kiviharju, K. Salonen, U. Moilanen, and T. Eerikäinen, “Biomass measurement online: the performance of in situ measurements and software sensors.,” *Journal of Industrial Microbiology & Biotechnology*, vol. 35, no. 7, pp. 657–65, Jul. 2008.

[79] P. Sagmeister, M. Kment, P. Wechselberger, A. Meitz, T. Langemann, and C. Herwig, “Soft-sensor assisted dynamic investigation of mixed feed bioprocesses,” *Process Biochemistry - currently in review*, pp. 1–26, 2013.

[80] T. L. and C. H. Patrick Sagmeister, Patrick Wechselberger, Andrea Meitz, “Soft-Sensors for Bioprocess Monitoring and Control -manuscript in preparation,” *unpublished*, 2012.

[81] G. Liu, S. Yu, C. Mei, and Y. Ding, “A novel soft sensor model based on artificial neural network in the fermentation process,” *African Journal of Biotechnology*, vol. 10, no. 85, pp. 19780–19787, Dec. 2011.

[82] R. Luttmann, D. G. Bracewell, G. Cornelissen, K. V Gernaey, J. Glassey, V. C. Hass, C. Kaiser, C. Preusse, G. Striedner, and C.-F. Mandenius, “Soft sensors in bioprocessing: A status report and recommendations,” *Biotechnology Journal*, vol. 7, no. 8, pp. 1040–1048, Aug. 2012.

[83] FDA, “Guidance for Industry PAT — A Framework for Innovative Pharmaceutical,” no. September. 2004.

[84] H. Tye, “Application of statistical ‘design of experiments’ methods in drug discovery.,” *Drug Discovery Today*, vol. 9, no. 11, pp. 485–91, Jul. 2004.

[85] C.-F. Mandenius and A. Brundin, “Bioprocess Optimization using Design-of-Experiments Methodology,” *Biotechnol. Prog.*, vol. 24, pp. 1191–1203, 2008.

[86] A. S. Rathore, R. Bhambure, and V. Ghare, “Process analytical technology (PAT) for biopharmaceutical products.,” *Analytical and Bioanalytical Chemistry*, vol. 398, no. 1, pp. 137–54, Sep. 2010.

[87] C. Herwig, “Prozess Analytische Technologie in der Biotechnologie,” *Chemie Ingenieur Technik*, vol. 82, no. 4, pp. 405–414, Apr. 2010.

[88] P. Sagmeister, C. Schimek, A. Meitz, C. Herwig, and O. Spadiut, “Design and multivariate investigation of an E . coli pBAD mixed feed fed-batch expression platform capable of tunable recombinant protein expression,” *unpublished*.

Appendix

Processed data

Fermentation N2 (DASGIP 11)

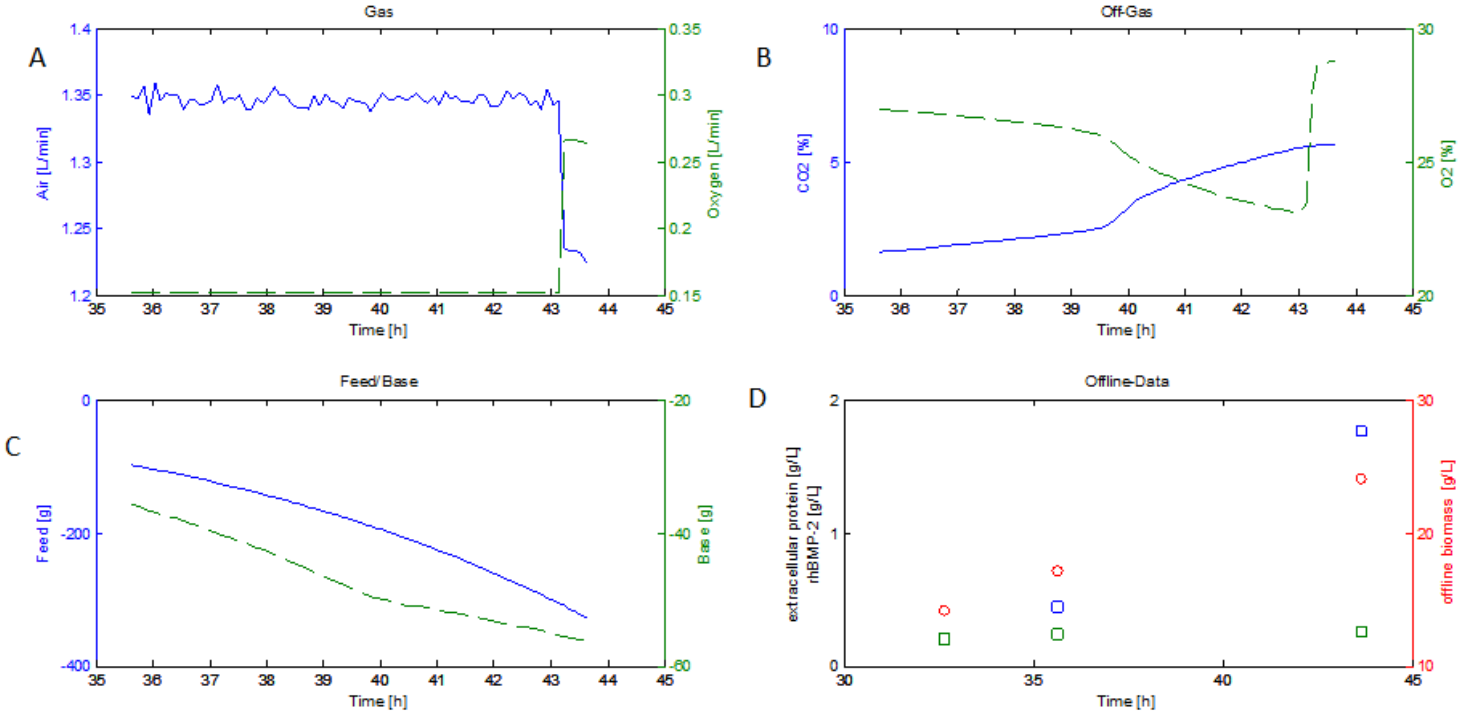

Figure 10 (A-D): Fermentation at 25 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A – C.

Fermentation N4 (DASGIP 11)

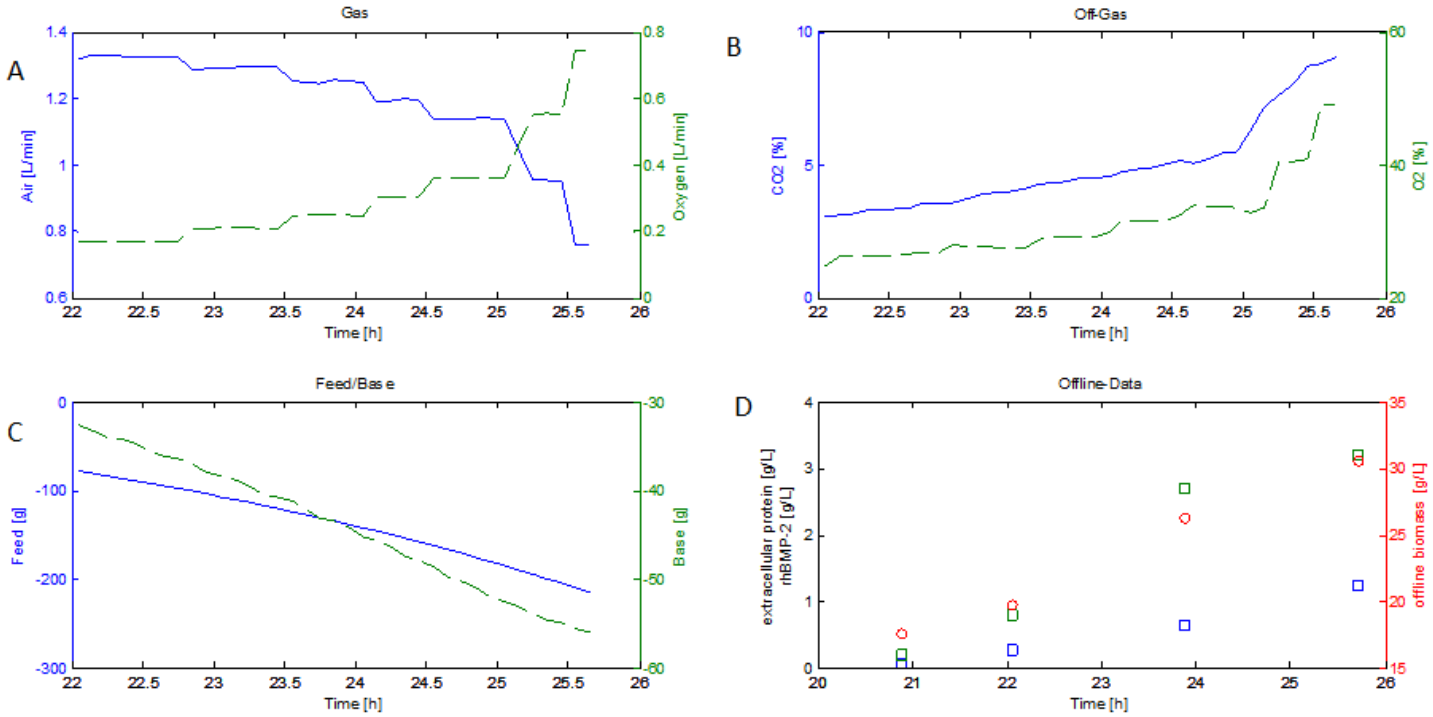

Figure 11 (A-D): Fermentation at 25 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A – D.

Fermentation N13 (DASGIP 11)

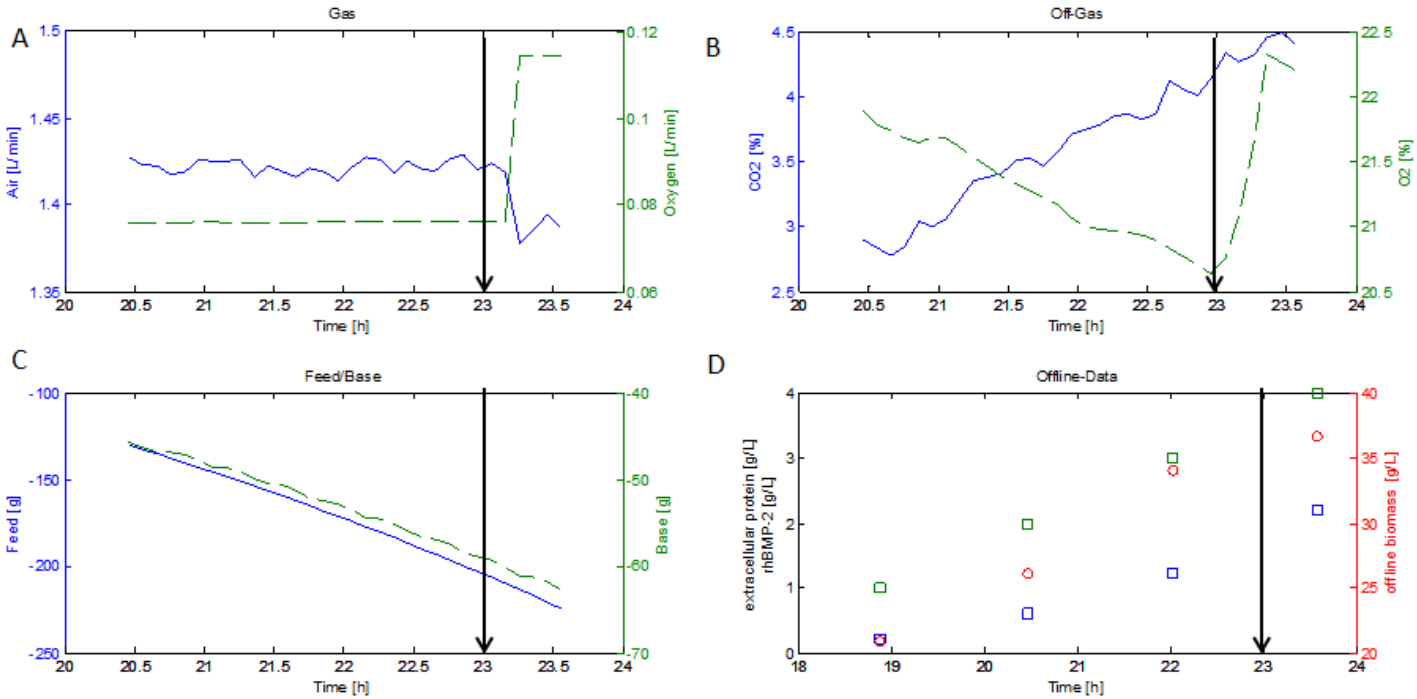

Figure 12 (A-D): Fermentation at 35 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –C. The black arrows denote that from this point the data are not included in the data evaluation.

Fermentation N10 (DASGIP 14)

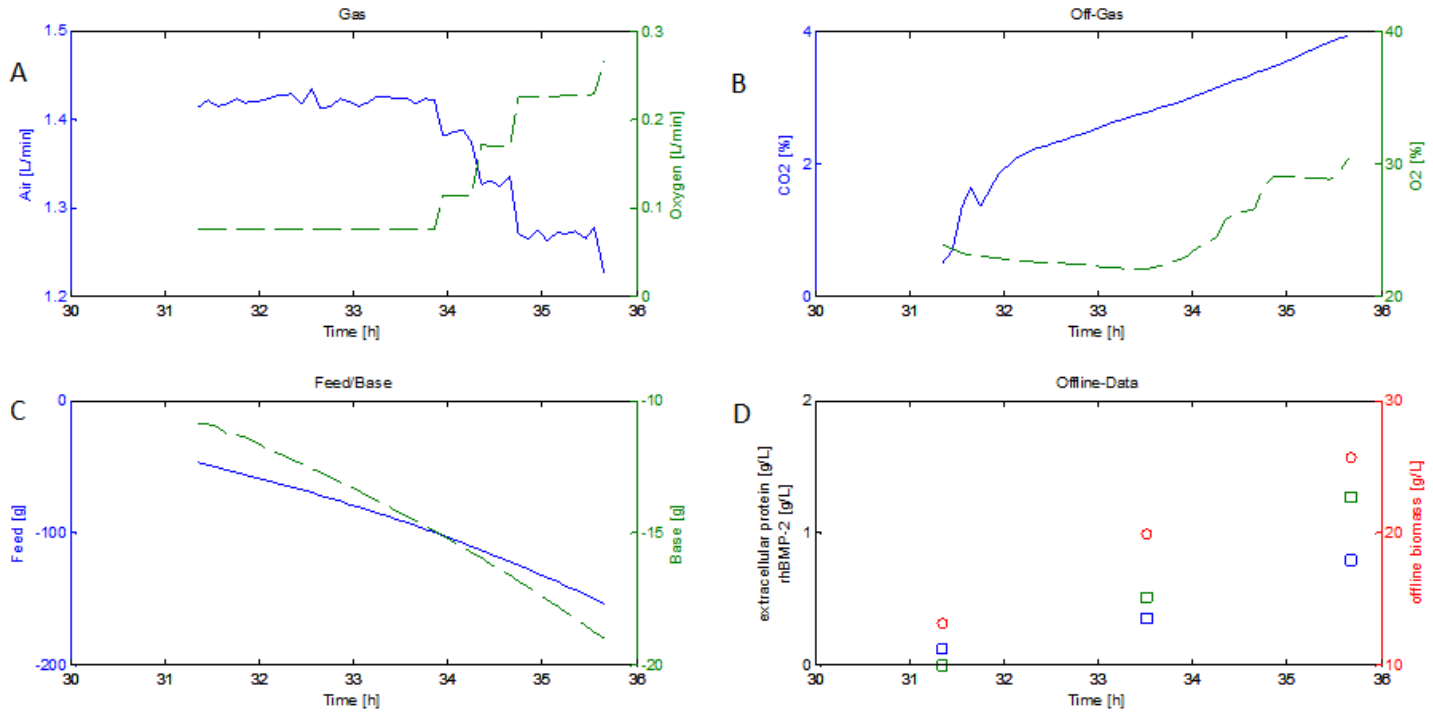

Figure 13 (A-D): Fermentation at 30 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –C.

Fermentation N14 (DASGIP 14)

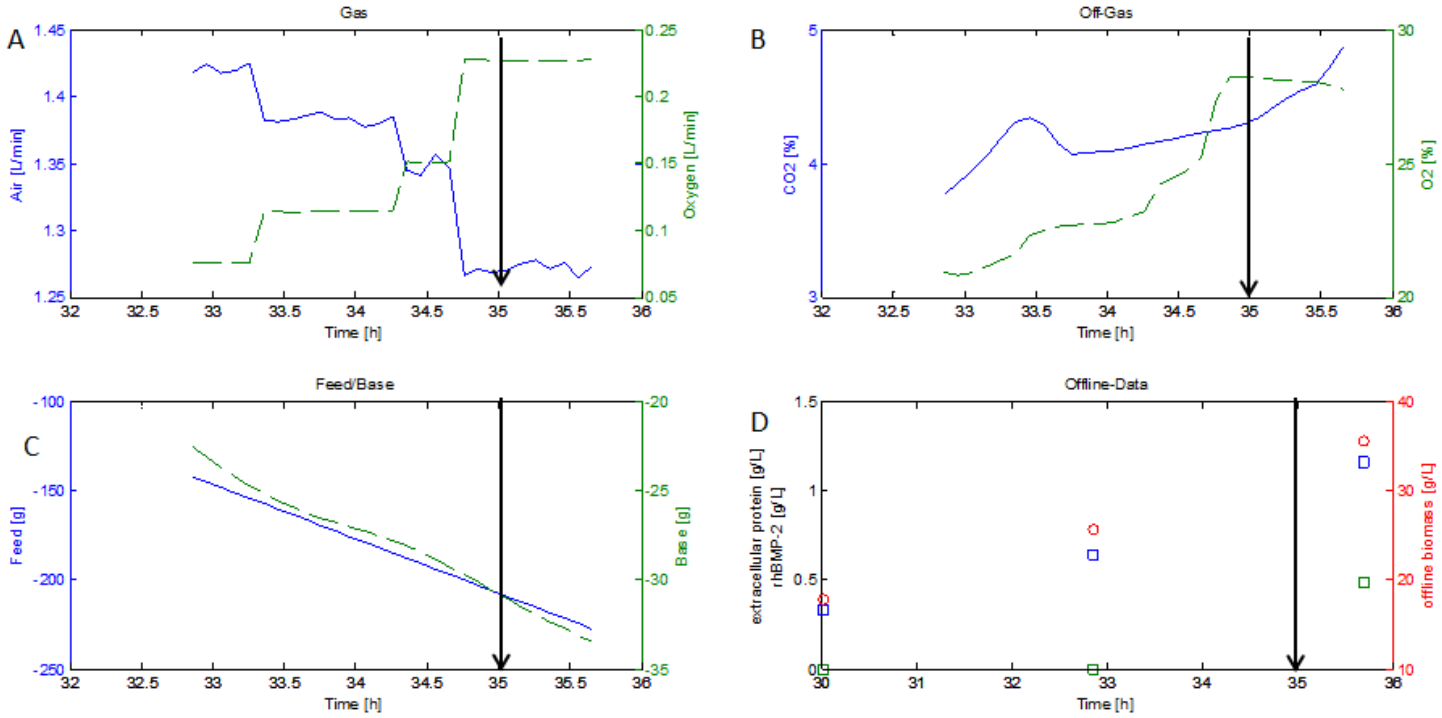

Figure 14 (A-D): Fermentation at 25 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –B. The black arrows denote that from this point the data are not included in the data evaluation.

Fermentation N7 (DASGIP 14)

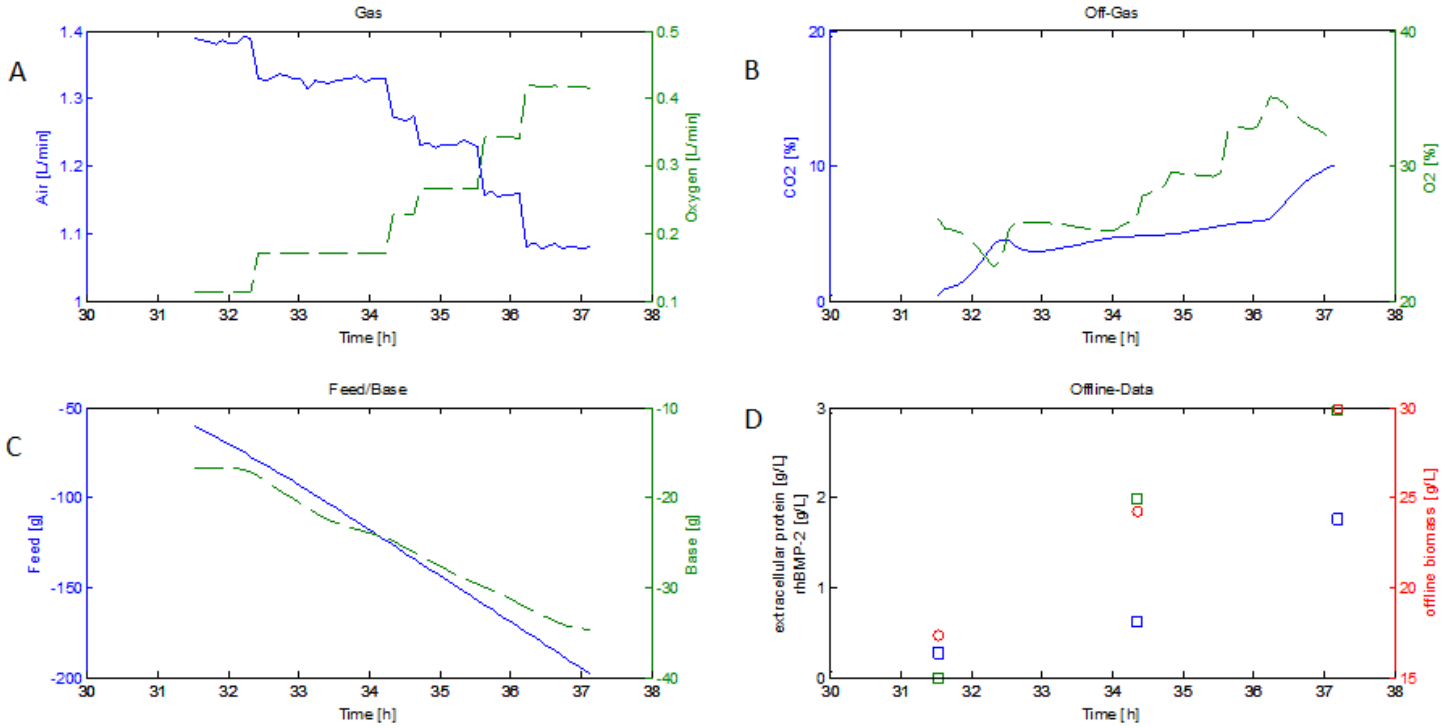

Figure 15 (A-D): Fermentation at 35 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –C.

Fermentation N6 (DASGIP14)

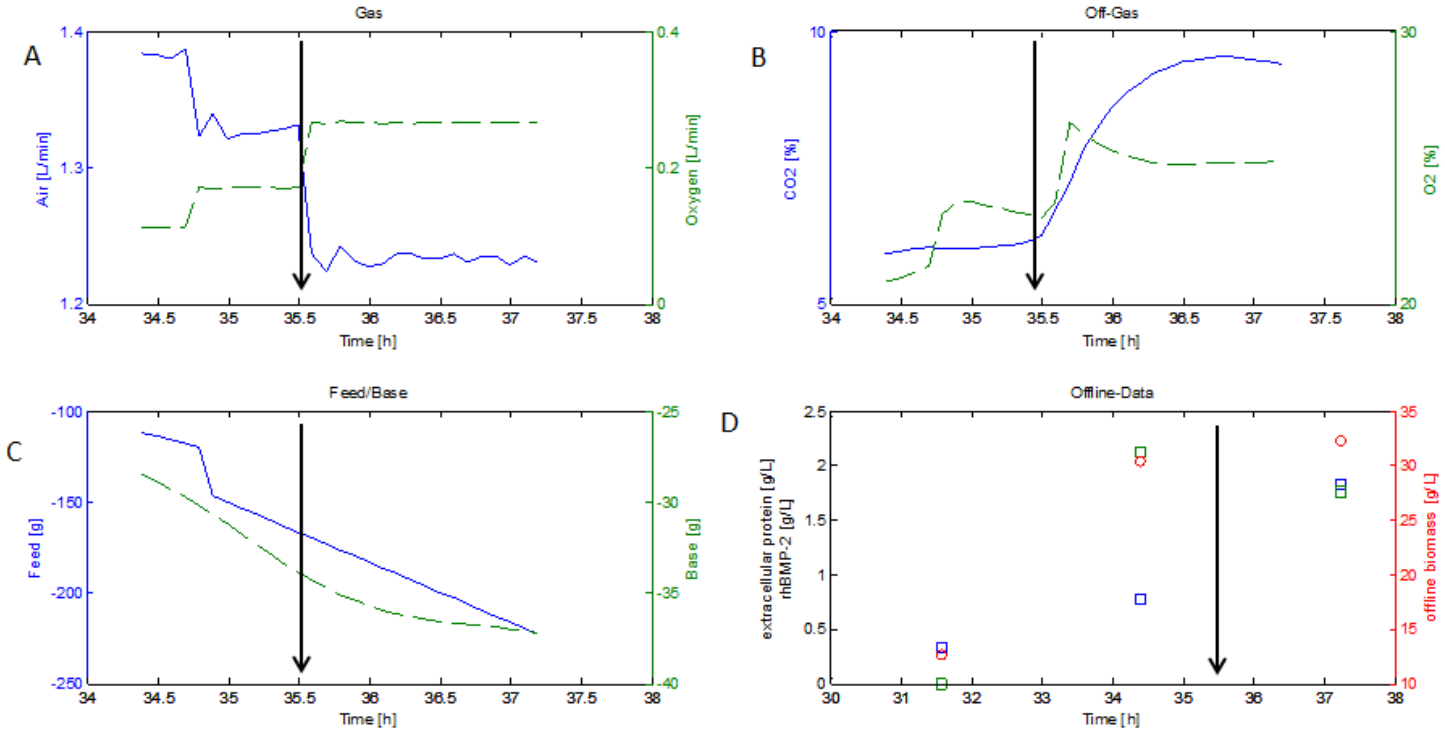

Figure 16 (A-D): Fermentation at 35 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –B. The black arrows denote that from this point the data are not included in the data evaluation.

Fermentation N12 (DASGIP 15)

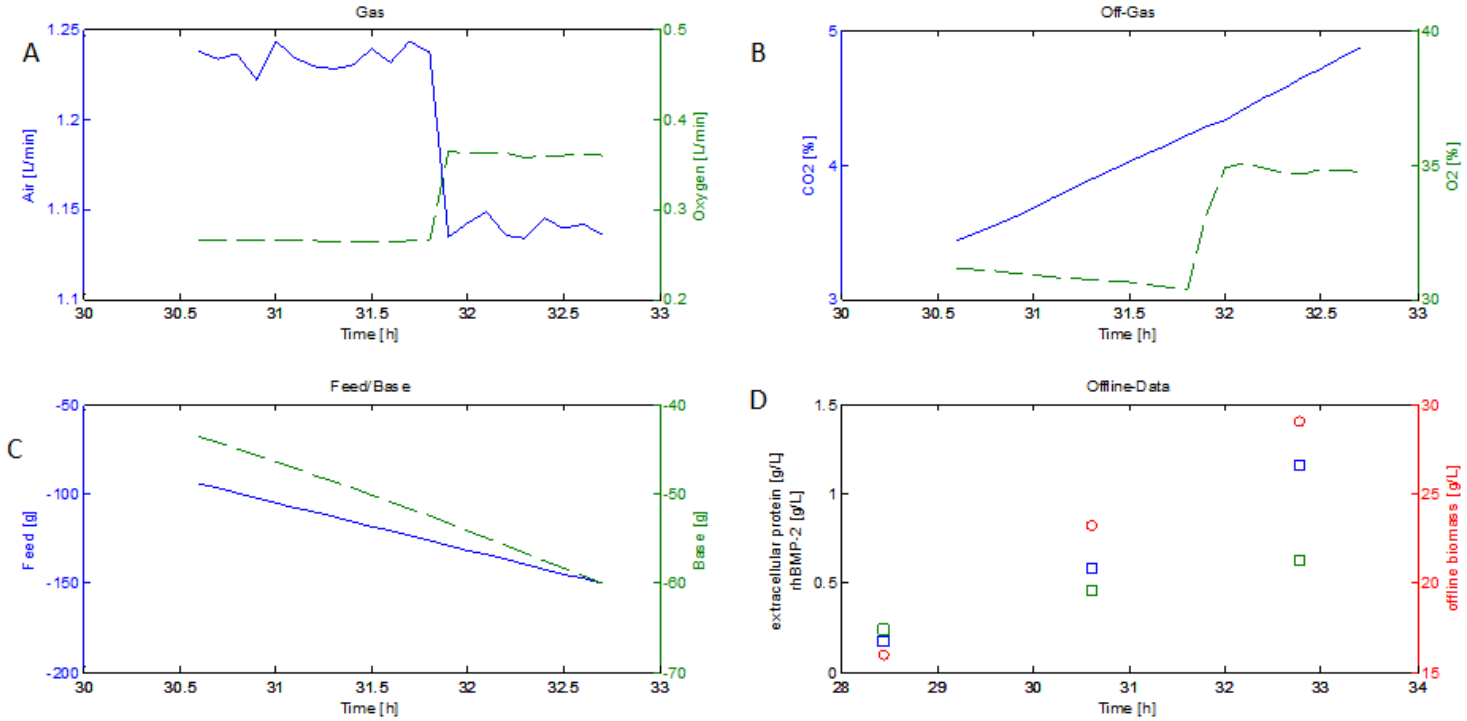

Figure 17 (A-D): Fermentation at 30 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –C.

Fermentation N16 (DASGIP 15)

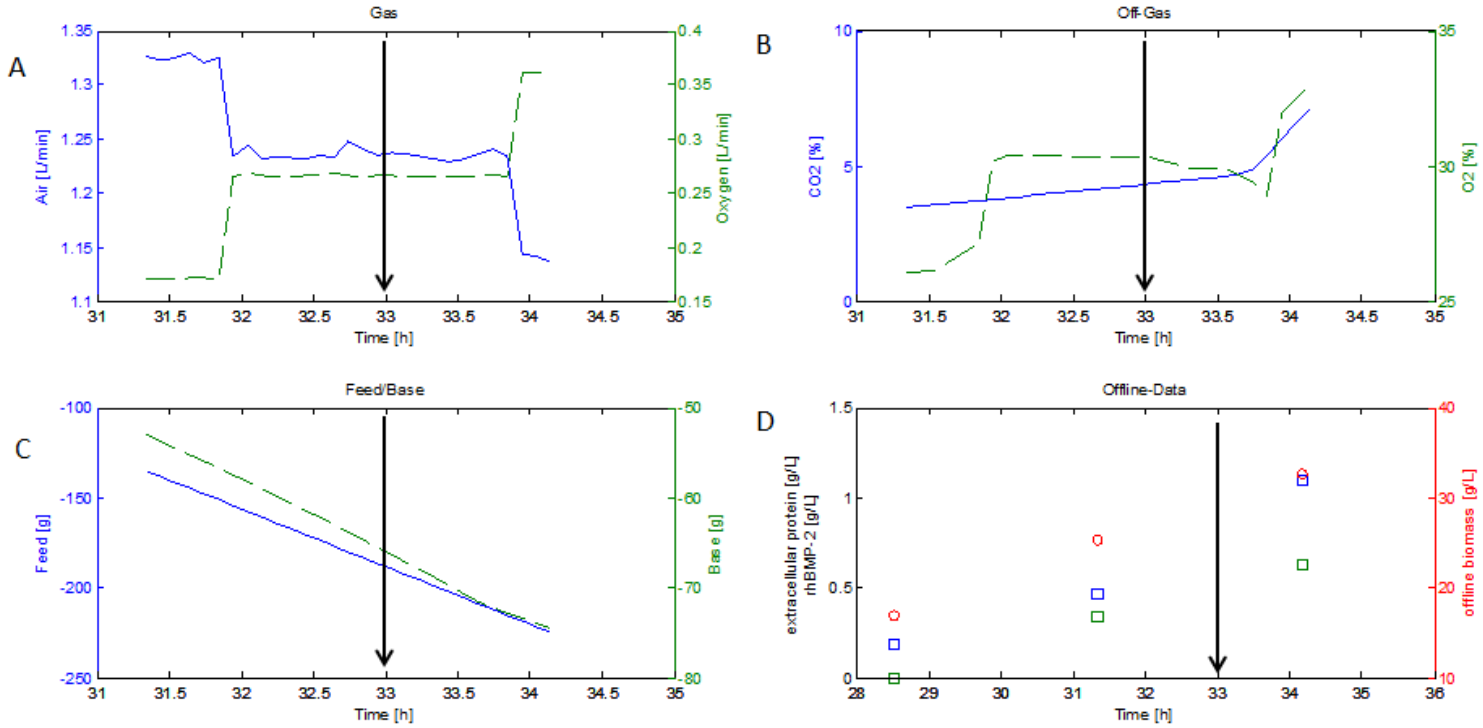

Figure 18 (A-D): Fermentation at 30 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –B. The black arrows denote that from this point the data are not included in the data evaluation.

Fermentation N8 (DASGIP 15)

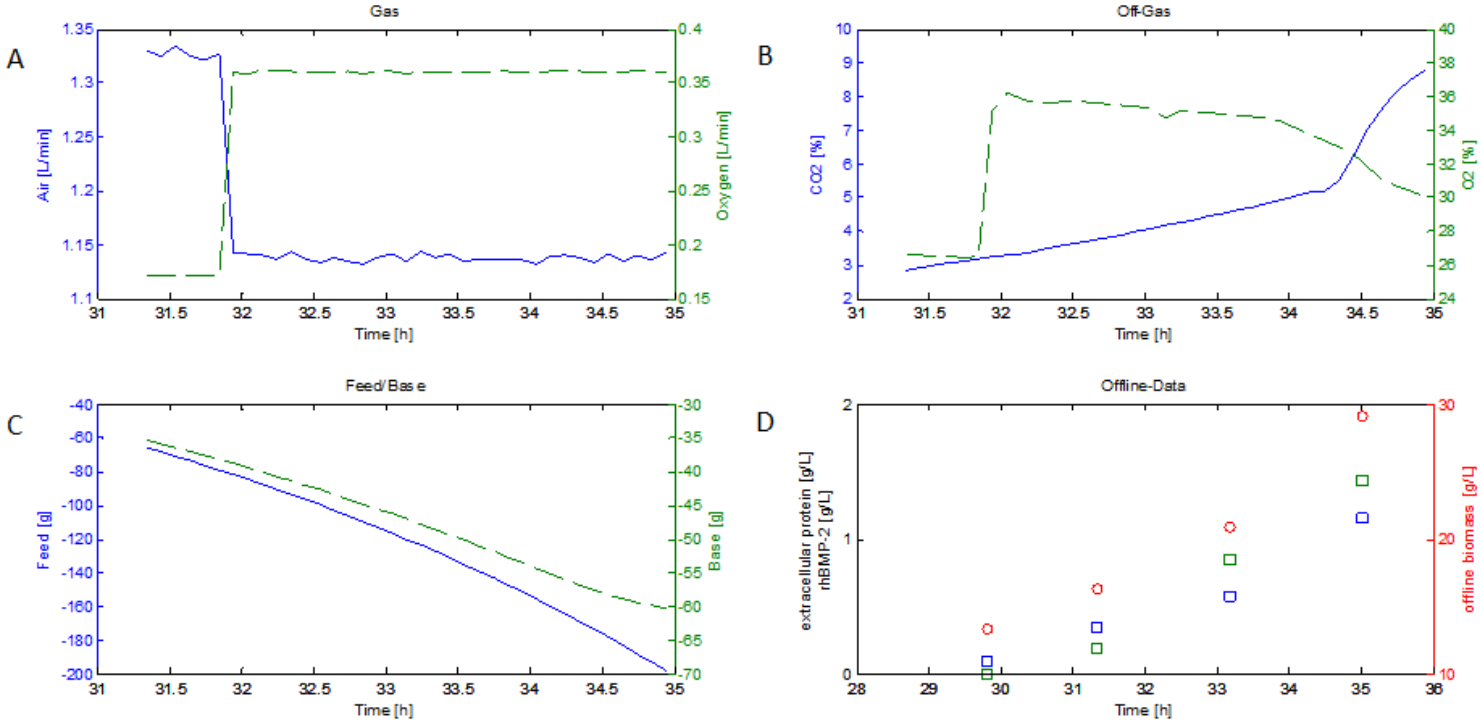

Figure 19 (A-D): Fermentation at 35 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –D.

Fermentation N15 (DASGIP 15)

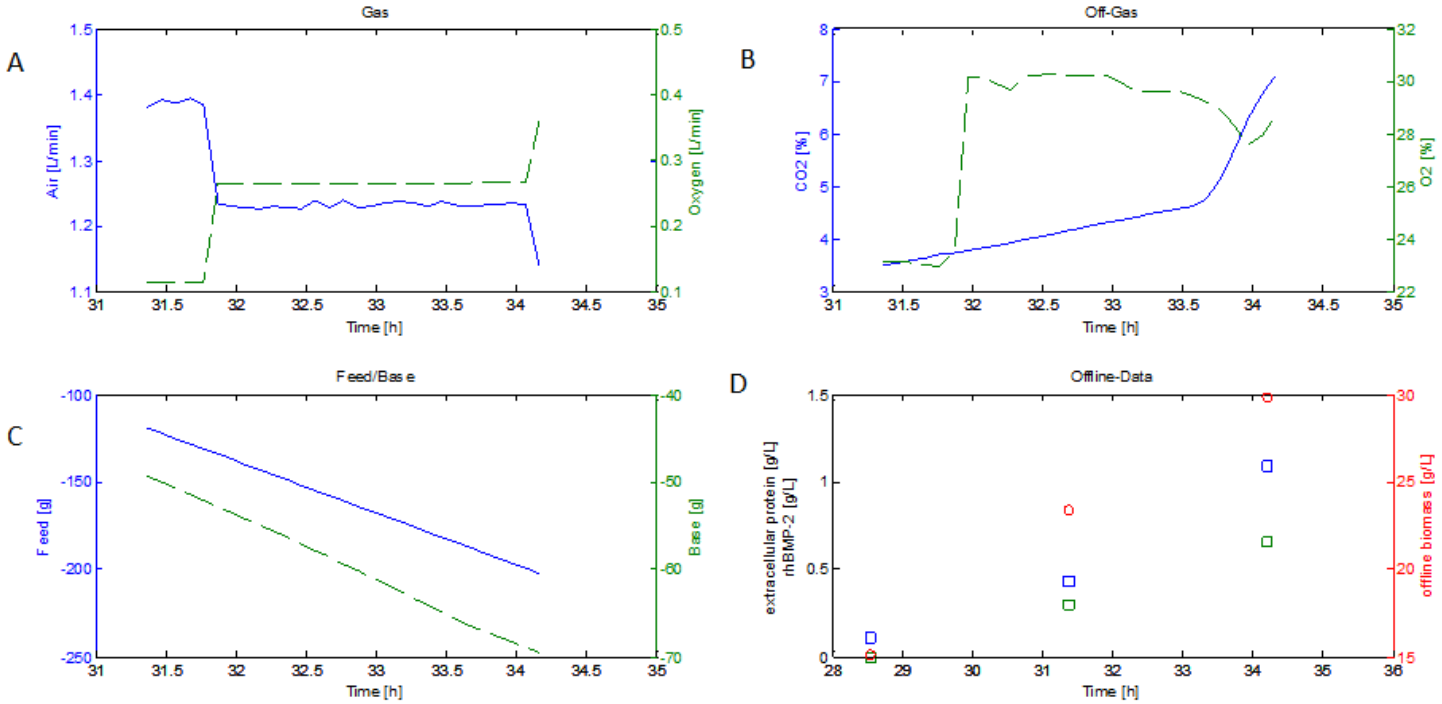

Figure 20 (A-D): Fermentation at 30 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –C.

Fermentation N3 (DASGIP 16)

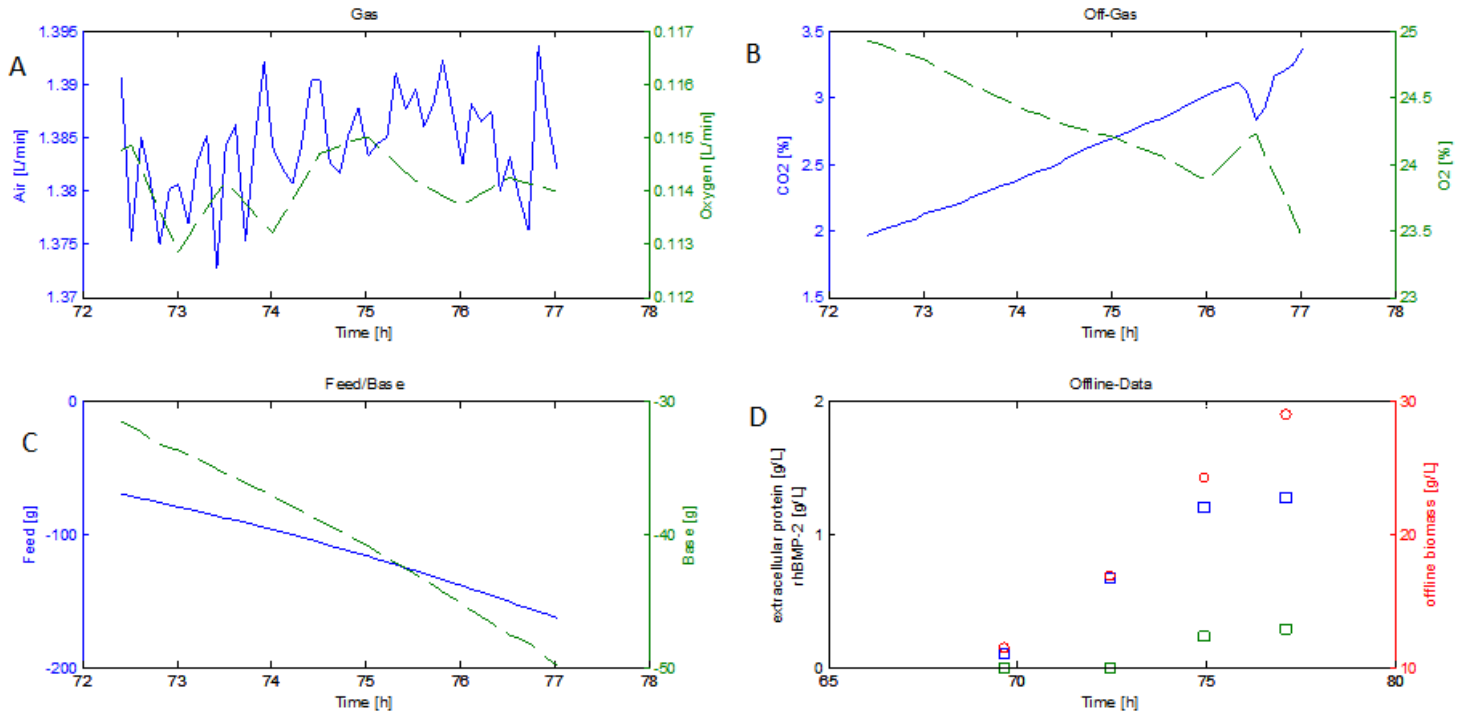

Figure 21 (A-D): Fermentation at 25 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –D.

Fermentation N17 (DASGIP 16)

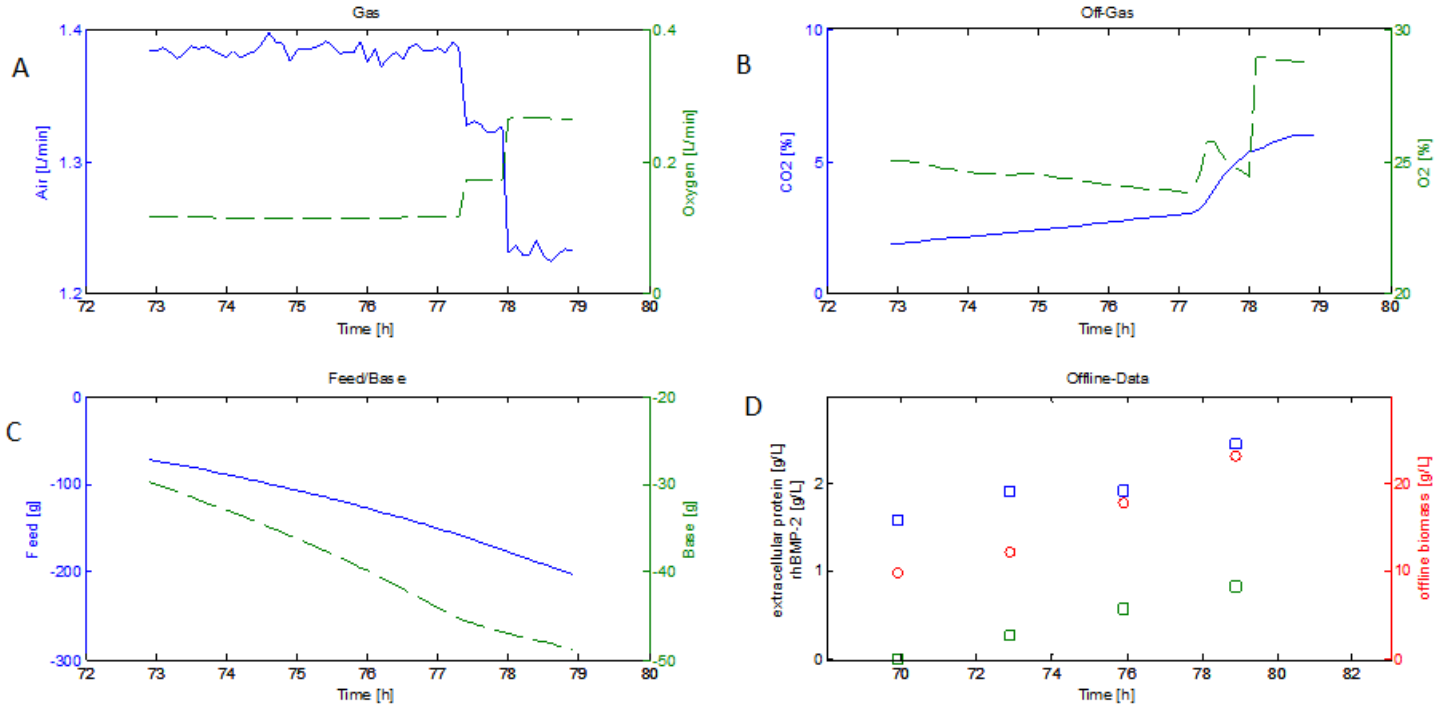

Figure 22 (A-D): Fermentation at 30 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples A –D.

Fermentation N9 (DASGIP 16)

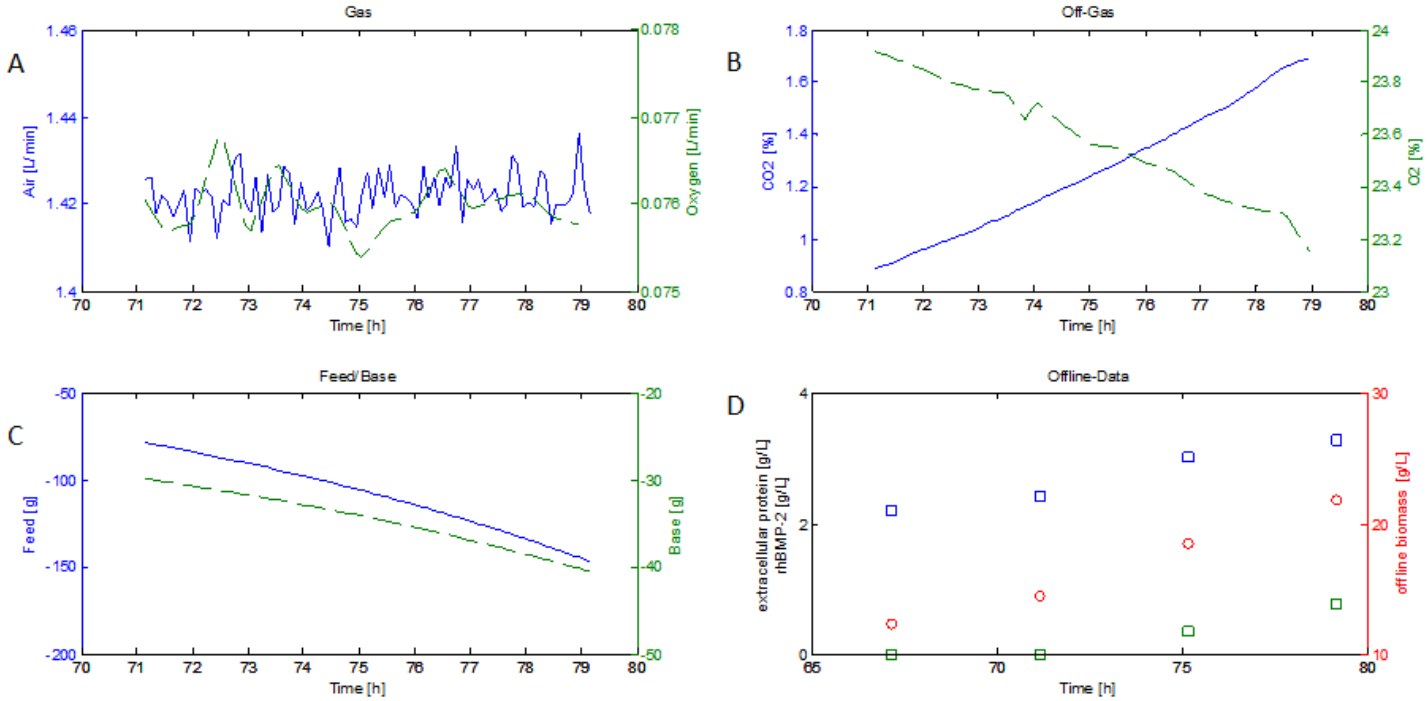

Figure 23 (A-D): Fermentation at 30 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples B –E.

Fermentation N11 (DASGIP 16)

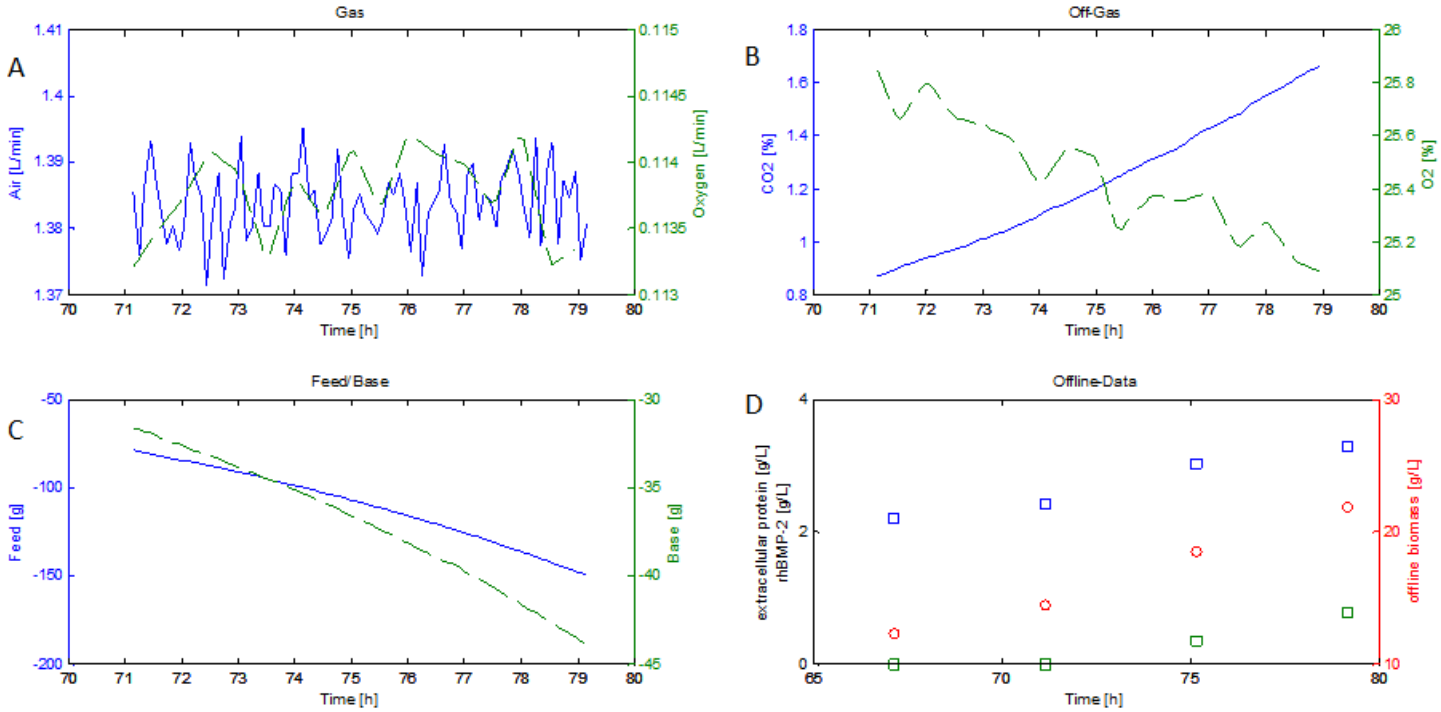

Figure 24 (A-D): Fermentation at 30 °C applying an exponential mixed-feed (L-arabinose and D-glucose) A) Gas inlet flow of air (blue) and oxygen (green) over process time. B) off-gas measurement of CO_2 (blue) and O_2 (green). C) Signals record from the feed (blue) and base (green) balance. D) extracellular protein concentration (blue square), rhBMP-2 concentration (green square) and offline biomass (red circle) biomass samples B – E.

SDS-PAGE evaluation

Experiments N2; N4; N13 (DASGIP 11)

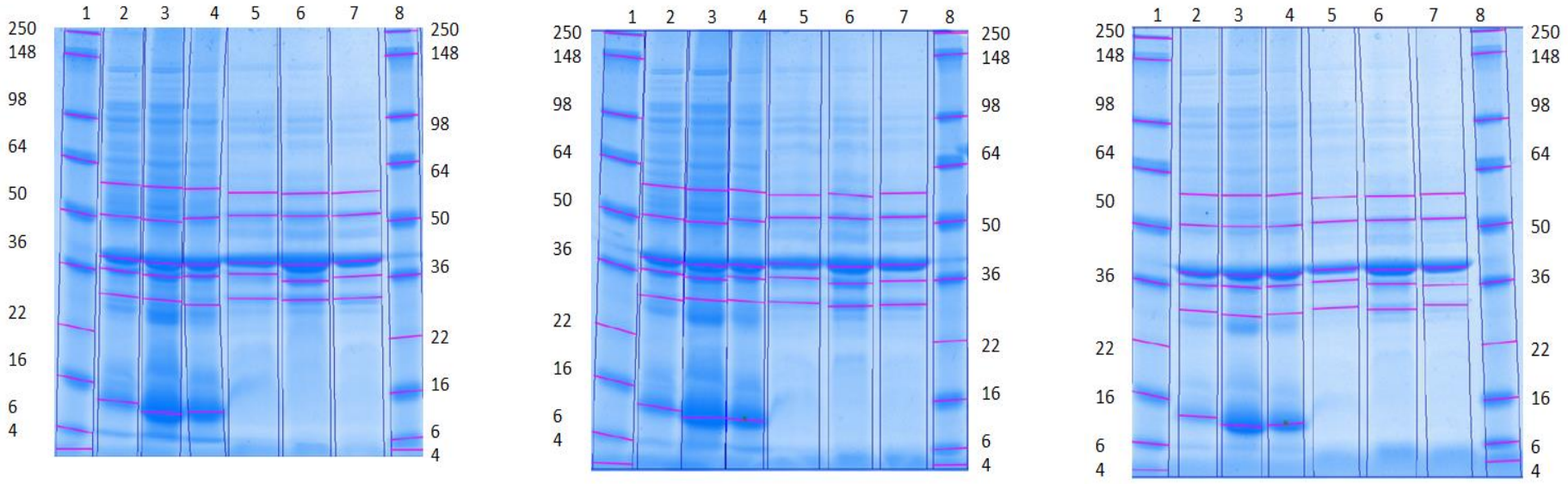

Figure 25 (SDS-PAGE gel analysis: all three gels represent identical triplicates, the lane composition is also identical): Lane 2 (N2 sample C at 25°C), 3 (N4 sample D at 25°C) and Lane 4 (N13 sample C at 35°C) represent the homogenization pellet sample; Lane 5 (N2 sample C at 25°C), 6 (N4 sample D at 25°C) and 7 (N13 sample C at 35°C) represent the fermentation supernatant (no soluble rhBMP-2 is detected), Lane 1, Land 8 = ladder SeeBlue® Plus2 Pre-Stained Standard (4-250 kDa). RhBMP-2 band detect at approximately ~ 14 kDa (depicts at lane 2,3 and 4).

Experiments N10; N14; N7; N6 (DASGIP 14)

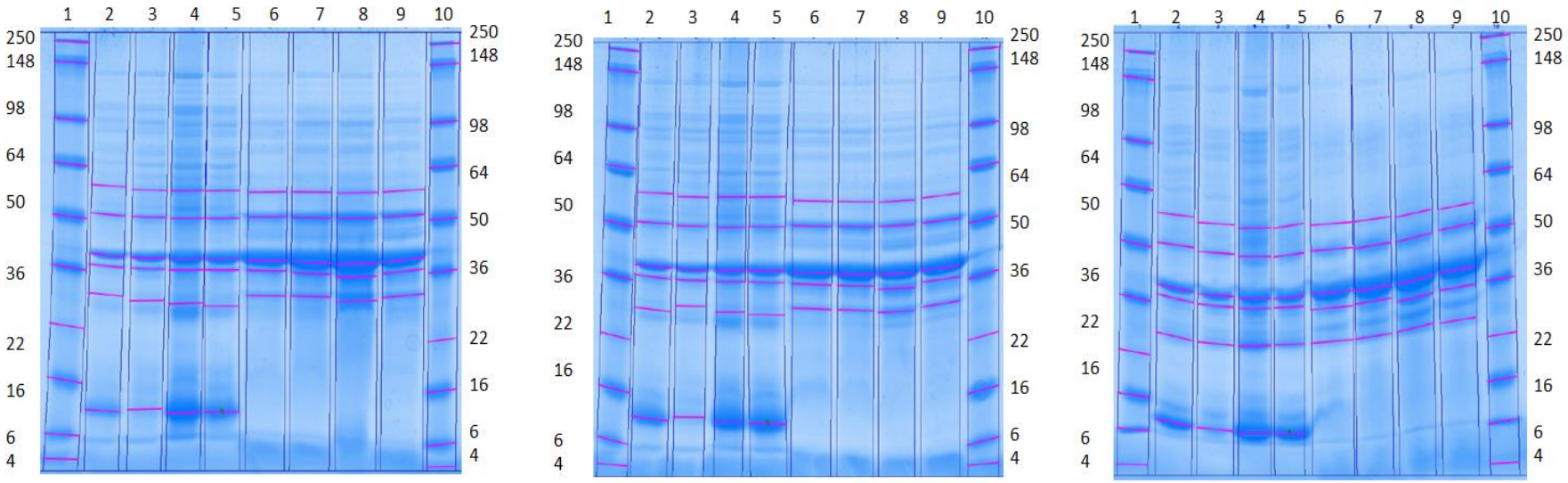

Figure 26 (SDS-PAGE gel analysis: all three gels represent identical triplicates, the lane composition is also identical): Lane 2 (N10 sample C at 30°C), 3 (N14 sample B at 25°C), Lane 4 (N7 sample C at 35°C) and Lane 5 (N6 sample B at 35°C) represent the homogenization pellet sample; Lane 6 (N10 sample C at 30°C), 7 (N14 sample B at 25°C) Lane 8 (N7 sample C at 35°C) and Lane 9 (N6 sample B at 35°C) represent the fermentation supernatant (no soluble rhBMP-2 is detected), Lane 1, Land 10 = ladder SeeBlue® Plus2 Pre-Stained Standard (4-250 kDa). RhBMP-2 band detect at approximately ~ 14 kDa (depicts at lane 2, 3, 4 and 5).

Experiments N12; N16; N8; N15 (DASGIP 15)

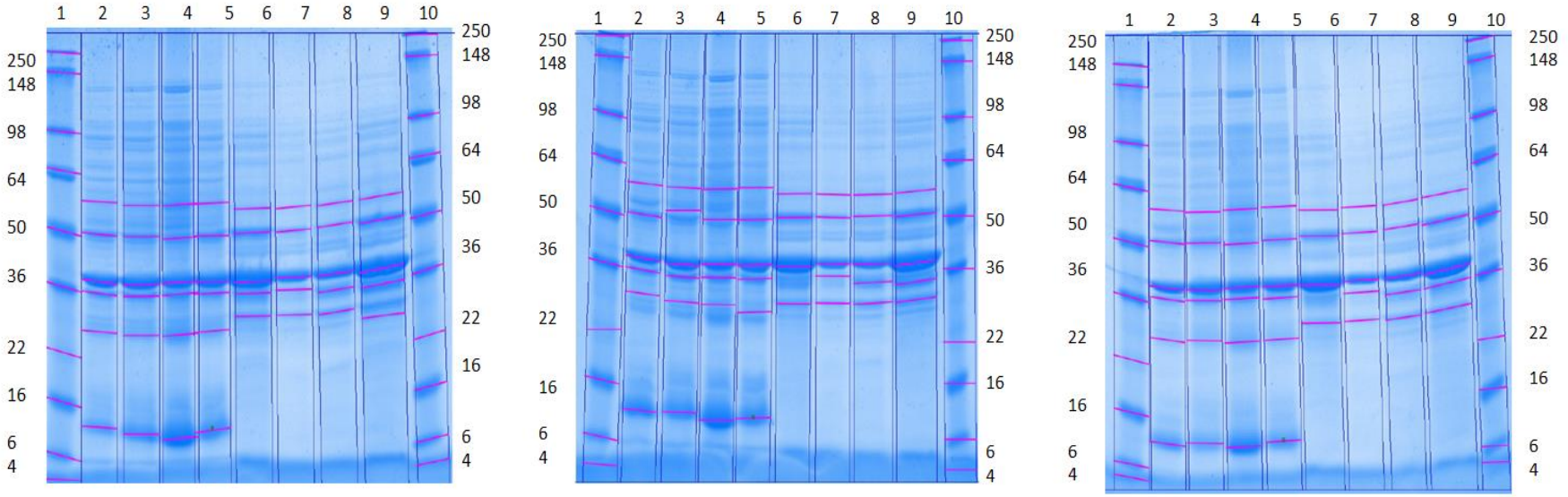

Figure 27 (SDS-PAGE gel analysis: all three gels represent identical triplicates, the lane composition is also identical): Lane 2 (N12 sample C at 30°C), 3 (N16 sample B at 30°C), Lane 4 (N8 sample D at 35°C) and Lane 5 (N15 sample C at 30°C) represent the homogenization pellet sample; Lane 6 (N12 sample C at 30°C), 7 (N16 sample B at 30°C) Lane 8 (N8 sample D at 35°C) and Lane 9 (N15 sample C at 30°C) represent the fermentation supernatant (no soluble rhBMP-2 is detected), Lane 1, Land 10 = ladder SeeBlue® Plus2 Pre-Stained Standard (4-250 kDa). RhBMP-2 band detect at approximately ~ 14 kDa (depicts at lane 2, 3, 4 and 5).

Experiments N3; N17; N9; N11 (DASGIP 16)

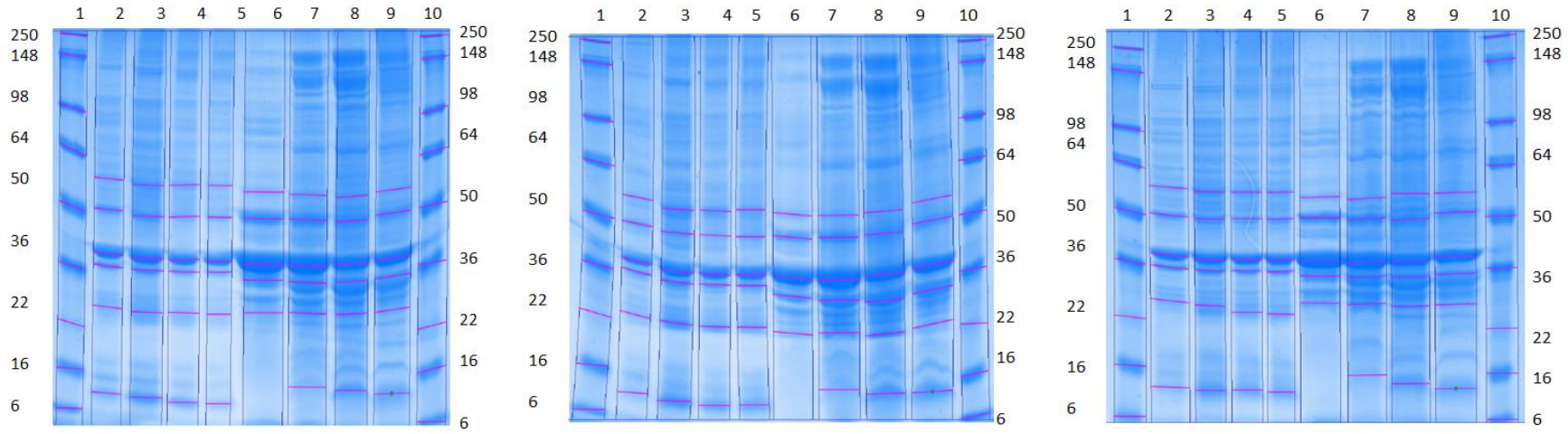

Figure 28 (SDS-PAGE gel analysis: all three gels represent identical triplicates, the lane composition is also identical): Lane 2 (N3 sample D at 25°C), 3 (N17 sample D at 30°C), Lane 4 (N9 sample E at 30°C) and Lane 5 (N11 sample E at 30°C) represent the homogenization pellet sample; Lane 6 (N3 sample D at 25°C), 7 (N17 sample D at 30°C) Lane 8 (N9 sample E at 30°C) and Lane 9 (N11 sample E at 30°C) represent the fermentation supernatant (soluble rhBMP-2 is detected approximately ~ 14 kDa (Lane 7,8 and 9)), Lane 1, Land 10 = ladder SeeBlue® Plus2 Pre-Stained Standard (4-250 kDa). RhBMP-2 band detect at approximately ~ 14 kDa (depicts at lane 2, 3, 4 and 5).

Fermentation Analytics for rhBMP-2 Fermentation Processes

Principal Investigator & Head of Laboratory:

Prof. Dr. Christoph Herwig

Supervisor:

DI Patrick Sagmeister

Master Student:

Bakk. tech. Magdalena Kment

Faculty:

Institute of Chemical Engineering
Technical University of Vienna
Getreidemarkt 9
1060 Wien

Associated Departments:

RCPE – Research Center Pharmaceutical Engineering
Inffeldgasse 21a/II
A-8010 Graz

BIRD-C Biotech Innovation Research Development & Consulting
Erne-Seder-Gasse 4/ Stiege 2/ Lokal 3
1030 Vienna, Austria

Morphoplant GmbH
Universitätsstrasse 136
44799 Bochum

REPORT WP8 2/2013

project:	A2.29
project title:	QbD Ghosts for GF
duration:	
project leader:	Prof. Christoph Herwig
	Dr. Stefan Leitgeb
key researcher:	Prof. Christoph Herwig
researcher:	Timo Langemann
	Andrea Meitz
	Patrick Sagmeister
project partners:	BIRD-C GmbH & Co KG
	Morphoplant GmbH

	Name	Date	Signature
author:	Magdalena Kment		
reviewed by :	Patrick Sagmeister		
approved by:			

distribution:

Abstract / Executive summary

Bioanalytical evaluation of fermentation products is an integral process of the production and development of recombinant biopharmaceuticals. Furthermore, the characterization of a biotechnological product by state-of-the-art analytical techniques is necessary to allow relevant product specifications to be established. Herein, an analytical scheme for the bioanalytical assessment of rhBMP-2 fermentation processes. RhBMP-2 is expressed under the control of a L-arabinose specific p_{BAD}-promoter using *E.Coli* C41 cells, is described. The analytical methods encompass the assessment of extracellular and also intracellular analytes.

Presence of extracellular protein indicates cell lysis. Extracellular total protein in the fermentation supernatant was quantified using BCA (Bichionic Acid) following TCA (Trichloracetic Acid) precipitation to remove interfering substances.

Quality and quantity of the intracellular inclusion body product, rhBMP-2, is analyzed after homogenization and solubilization. Homogenized pellets are analyzed using Sodium Dodecyl Sulphate Polyacrylamide Gel Electrophoresis (SDS-PAGE) to assess the purity of the product in respect to host cell proteins. After solubilization of homogenized pellets, reverse-phase high pressure liquid chromatography (RP-HPLC) is used for quantification of rhBMP-2. Both methods, SDS-Page and RP-HPLC, are assessed for their reproducibility and the obtained errors in measurement.

Objectives

Primary Objective

- The primary objective of this evaluation is to establish an analytical scheme for bioanalysis of rhBMP-2 fermentation processes.

Secondary Objectives

- The following secondary objectives are envisaged:
    - assessment of the reproducibility of analytical methods
    - assessment of measurement errors of analytical methods

- assessment of the comparability/feasibility of RP-HPLC and SDS-Page for the quantification/ quality assessment of rhBMP-2

Materials and Methods

Extracellular Analytes

Two times 2ml of the cell suspension (from the fermentation broth) were centrifuged (RZB 5171, 10min) and the cell-free supernatant (two times 1ml) frozen (-20°C) and stored for further analytics.

Extracellular total protein quantification

- **TCA (Trichloroacetic acid) precipitation:**

  Media components were detected to interfere with BCA total protein quantification. Hence, TCA precipitation was used to separate extracellular protein from the fermentation supernatant from interfering substances before protein quantification. Therefore, the cell-free supernatant (two times 0.5 ml) was mixed with 0.5 ml 10% w/v TCA solution and incubated for 10 min on 4°C for precipitation. After centrifugation (13000 g, 10 min, 4°C) the supernatant was discarded and the pellet of precipitate was washed with 1 ml -20°C acetone. After a further centrifugation (13000 g, 10 min, 4°C) the supernatant (containing acetone) was discarded and the cap of the Eppendorf tubes was left open for a few minutes to allow evaporation of remaining acetone.

- **Protein content quantification using the BCA protein assay (incl. sample preparation):**

  The sodium salt of Bicinchoninic acid (BCA) in complex with copper ions (Cu1+) is able to react with protein in an alkaline environment according to the principles of the biuret reaction [see Smith PK et al, Anal Biochem, 1985]. The BCA method is more reliable than the method proposed by Lowry [Andrew Wong et al.Application Note – Industrial BioDevelopment Laboratory (www.ibdl.ca)] Before protein content was

measured in a 20 GENESYS SPECTROPHOTOMETER (Thermo Spectronic) at 562nm, the samples and the necessary working reagent were prepared as follows:

Sample preparation and preparation of the working reagent:

The pellet of the protein precipitate was dissolved in an adequate volume (e.g. 1 ml) of 0.1 M / 1% NaOH/SDS buffer overnight to allow full solubilization.

For the preparation of the working reagent, 50 parts of BCA reagent A (Bichionic Acid Solution) and one part of BCA reagent B (Cooper II Solution) were mixed.

Calibration was done using 1 mg BSA/mL as standard (as diluent, the same buffer (0.1 M /1% NaOH/SDS buffer) as in the sample was used). The samples should show concentrations in the calibration range.

Table 3: Calibration using BSA solution as standard

conc [µg/ml]	Standard [µl]	Diluent [µL] (0.1 M / 1% NaOH/SDS buffer)
50	10	190
100	20	180
200	40	160
400	80	120
600	120	80

BCA measurement:

The measurement was carried out by mixing 50 µl of protein sample with 1 ml BCA working reagent. Afterwards, the mixture of protein sample and BCA working reagent was incubated. For incubation, two different protocols were applied:

- 60 °C using the water bath for 15 minutes
- 37 °C using the water bath for 30 minutes

After incubation the absorbance of standards and prepared samples was measured at 562 nm using spectrophotometer.

Intracellular Analytes

2ml of the cell suspension (from the fermentation broth) were centrifuged (RZB 5171, 10min), washed once with distilled water and the pellets were stored at -20°C until further analysis.

Homogenization for cell rupture

In order to disrupt cell membranes of fermented *E. coli* cells, samples were homogenized as follows: Samples were re-suspended in 20ml of 50mM Tris buffer supplemented with 1mM EDTA pH 8. Agglomerates of cells can hamper the function of the homogenizer. In case there were agglomerates present in the sample, samples were pretreated with an Ultra-Turax® (IKA® T10-basic) for 1 min. at level 6. The obtained slurry was pumped six times at 1500 bar through a high-pressure homogenizer (Avestin EmulsiFlex©C3, Ottawa). For six passages it took 3 minutes and 23 seconds, one passage for 20 ml took 32.3 seconds. The homogenized sample was divided into six aliquots. Then the aliquots were transferred into six 1ml Eppendorf tubes and centrifuged for 10 min at 13000 x g at room temperature. The residual homogenized sample was stored at -20°C until further analysis.

Solubilization of inclusion bodies

The biologically active conformation of rhBMP-2 is constituted *via* a disulfide-bond between two monomers (which form the active homodimer) and also *via* additional intrachain disulfide bonds. However, *E. coli* expression system usually produces ectopically expressed proteins containing S-S-bonds in form of insoluble inclusion bodies. In order to free included rhBMP-2, inclusion bodies were solubilized. For solubilization, guanidinium HCL and 2-Mercaptoethanol were used. Guanidinium HCL exhibits chaotropic properties, while 2-mercaptoethanol reduces oxidized cysteine residues in disulfide bridges. Solubilization was done according to routines previously described by Morphoplant GmBH. Briefly, a concentration of 5g/L of homogenized pellet was resuspended in solubilization buffer (10 mM Tris, 6 M Guanidinium Hydrochlorid, 50 mM 2- Mercaptoethanol, 10 mM Iodacetamid,

pH 7.6) and gently shaked overnight using an orbital shaker (IKA® - VXR- Basic Vibrax, typ VX-7 (Jenk & Kungel)) and centrifuged for 10 min at 13000 x g at room temperature.

RhBMP-2 quantification via RP-HPLC

RP-HPLC was used to identify and quantify rhBMP-2 in the cell debris pellet (centrifuged fermentation broth).

Separation from other components and quantification of rhBMP-2 by reversed phase high performance liquid chromatography was performed on a ProSwift® Reversed –Phase Monolithic Column RP-1S (4.6 x 50mm, Stainless Steel, Dionex).
The standard run used a water-acetonitril gradient, as illustrated in Figure 29 . Therefore, two different buffer solutions (Buffer A, Buffer B) were prepared. Buffer A was composed of 95% Milli-Q water and 5% AcN plus 0.1% TFA (trifluoroacetic acid). Buffer B contained the same ingredients but in a different composition (Buffer B: 95% AcN, 5% Milli-Q water, 0.1% TFA).

100% Buffer A was used for equilibration for 5 mins. Afterwards, Buffer B was used for the gradient elution (duration 10 mins; 0% to 100%). Then, another 2 mins. using 100% Buffer B. Afterwards, the concentration of Buffer B was reduced from 100% to 0% within 1 min. Finally, 100% Buffer A was used again for 2 mins of equilibration.
At a flow rate of 0.5 ml/min, an injection volume of 80 µl was analyzed within 20 minutes. Detection of rhBMP-2 was performed at 214 and 280 nm [Duggirala 1996]. The analysis was carried out on a Thermo Scientific Dionex UltiMate 3000 –Serie with a DA (Diodearray) Detector at a wavelength range between 190–800 nm at 1 nm increment (Thermo Scientific Dionex, Germany). The system was operated with Chromeleon 7 software (Thermo Scientific Dionex).

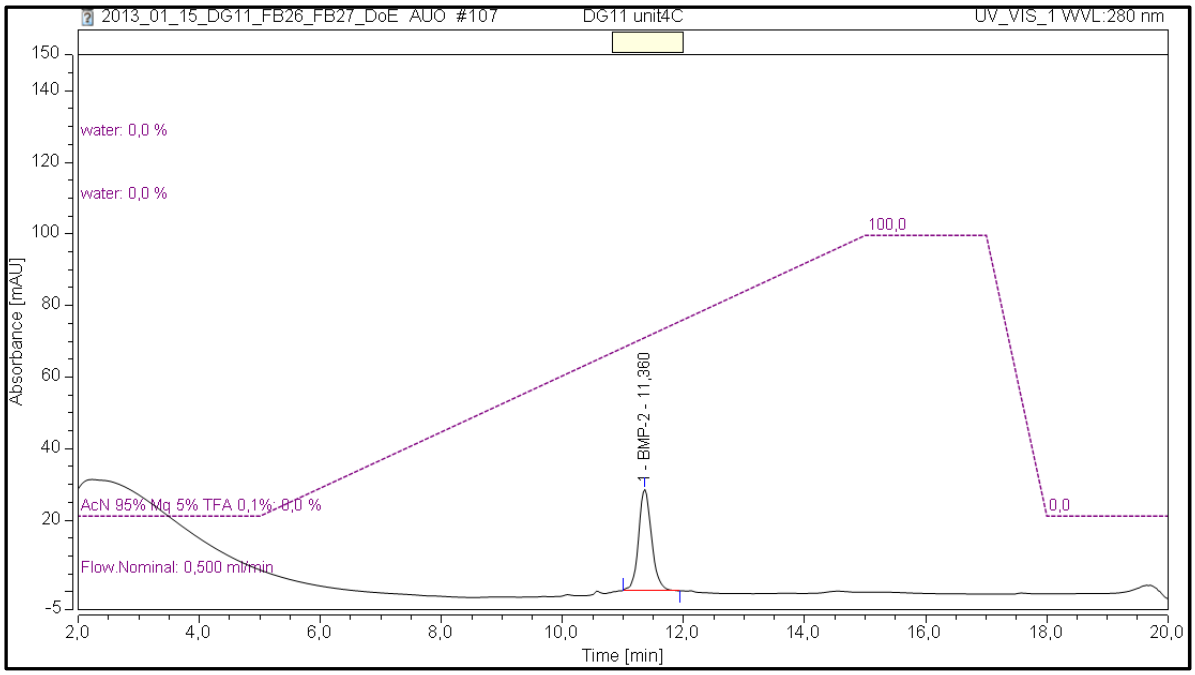

Figure 29: Equilibration profile of rhBMP-2: pink line = : 95% AcN, 5% Milli-Q water, 0.1% TFA

SDS-Page

Gel electrophoresis was carried out in an Amersham ™ ECL ™ Gel Box (GE Healthcare Life Sciences, Sweden) using Amersham ™ ECL ™ pre-cast gels (8 - 16 %, 10 wells, 1.4 mm) and 1x SDS running buffer (25 mM Tris base, 192 mM glycine, 0.1% SDS).

The homogenized protein pellet was dissolved in 200 µl of 1 x Laemmli sample buffer (For 100 ml 1x buffer: 62.5 mM Tris (w/v); 2 % SDS (w/v); 10 % Glycerin (w/v); 0.002 % Bromphenolblau (v/v) and fill up to 100 ml with high-purity-water, pH: 6.8, add ß-Mercaptoethanol for end concentration 5 %).

After heating for 10 minutes at approx. 95°C and centrifugation, 20 µl - typically containing an estimated amount of 15-20 µg rhBMP-2 - were loaded per lane. The loading volume of the BenchMark ™ prestained protein ladder (Invitrogen) was 8 µl. Gels were run with a constant current of 140 volt per gel for 1.3 hours.

After a washing step in deionised water the Sensitive Coomassie staining was carried out in one step. The gel was processed with staining mixture (0.02 % (w/v) Coomassie Brilliant Blue G 250, 5 % (w/v) Aluminium Sulfat-(14-18)-Hydrat, 10 % (v/v) Ethanol, 2 % (v/v)

Orthophosphoric Acid) for 3 to 12 hours on an orbital shaker (IKA® - VXR- Basic Vibrax, typ VX-7 (Jenk & Kungel)) followed by a washing step in deionised water for a few minutes. After drying, analysis of the stained gels was performed by densitometry, using Gel Doc™ with System Image Lab™ software (Bio-Rad).

Image Lab software can be used to annotate and document images, analyze molecular weights (or base pairs, when imaging nucleic acid gels), and to determine accurate quantitation and purity of samples. A standard filter was used for colorimetric (white light) applications. Gel Doc XR+ Imager: Following were the basic steps to acquiring, analyzing, and archiving an image using the Gel Doc XR+ system and Image Lab software:

1. Select an customize protocol : Protein gels (with Coomassie Blue)
2. Position the sample to be imaged.
3. Run a selected protocol.(Filter: Standard Filter 1; Light: White Trans Illumination)
4. View the displayed results.
5. Optimize the analysis.

Perform lane-based background subtraction by selecting Enable Subtraction in the Background Subtraction field. Use the Lane Profile view to see the subtracted lane background. To compare the relative quantities of bands, first select Quantity Tools (Relative tab) to select a reference band from an image and quantitate all other bands relative to the reference band. The relative quantity is the ratio of the band volume divided by the reference band volume.

6. Generate a report.
7. Save or export the results.

Table 4: Chemicals

Chemicals	Company
Trichloroacetic acid	Roth
Acetone	Loba Chemie
BCA A	Sigma
BCA B	Sigma
Albumin Fraktion V (BSA)	Roth
Orthophosphoretic Acid	Merk
Ethanol	VWR
Coomassie briliant Blue G-250 (Merck)	VWR
Aluminiumsulfat Tetradecahydrat (Merck)	VWR
TRIS	CarlRoth
Glycerin	Roth
2-ME	Carlroth
SDS	Sigma
Bromophenolblue	Fluka
Glycine	Merck
Tris base	Roth
EDTA	Carlroth
Gdn-HCl	Sigma
Iodoacetamid	Sigma

Calculation of SNR, LOD, LOQ

The following formulas for the calculation of the Signal-to-noise Ratio (SNR), the Limit of Detection (LOD) and Limit of Quantification (LOQ) were used:

$$SNR = \frac{X}{SD}$$

LOD=3*SD

LOQ=10*SD

Results & Discussion

Setup of Fermentation runs

For the analytical methods described herein, samples of three different fermentation runs were used. These three fermentation runs differed with respect to q_s and temperature.

The following table gives a brief overview on the four different fermentation runs:

Table 5: Setup of the different fermentation runs that served as basis for the analytical evaluation

Name of fermentation run	$q_{s\ glucose}$	$q_{s\_arabinose}$	Temp.
DG011_Unit1	0.2 g/g/h	0.05 g/g/h	25°C
DG011_Unit3	0.2 g/g/h	0.2 g/g/h	25°C
DG011_Unit4	0.125 g/g/h	0.125 g/g/h	35°C

Comparative assessment of different preparation methods for analysis of rhBMP-2 via SDS-PAGE

At first, the last two samples drawn during fermentation were analyzed and compared using SDS-PAGE with respect to the following items:

- Is there extracellular rhBMP-2 in the supernatant?

- Is there a difference between rhBMP-2 concentration with or without TCA precipitation?
- Is there an influence of pellet dilution?

The results of these assessments are shown in Figure 30. The 10 Lanes represented in Figure 29 contain the following arrangement:

- Lane 1: Ladder (Benchmark pre-stained Invitrogen)
- Lane 2: DG011_last sample_1D_homogenized supernatant
- Lane 3: DG011_last sample_3G_homogenize supernatant
- Lane 4: DG011_last sample_1C_homogenized pellet
- Lane 5: DG011_last sample_3F_homogenized pellet
- Lane 6: DG011_last sample_1C_homogenized pellet (1:2)
- Lane 7: DG011_last sample_3F_homogenized pellet (1:2)
- Lane 8: DG011_last sample_1D_homogenized pellet (TCA precipitation)
- Lane 9: DG011_last sample_3G_homogenized pellet (TCA precipitation)
- Lane 10: Ladder

When comparing Lane 2 and Lane 3 of Figure 30, which represent the supernatants of the two samples, to those lanes containing the pellet (Lanes 2,3 und 8,9) it can be seen that no-rhBMP-2 is present in the supernatant at a molecular weight of 13kDa.

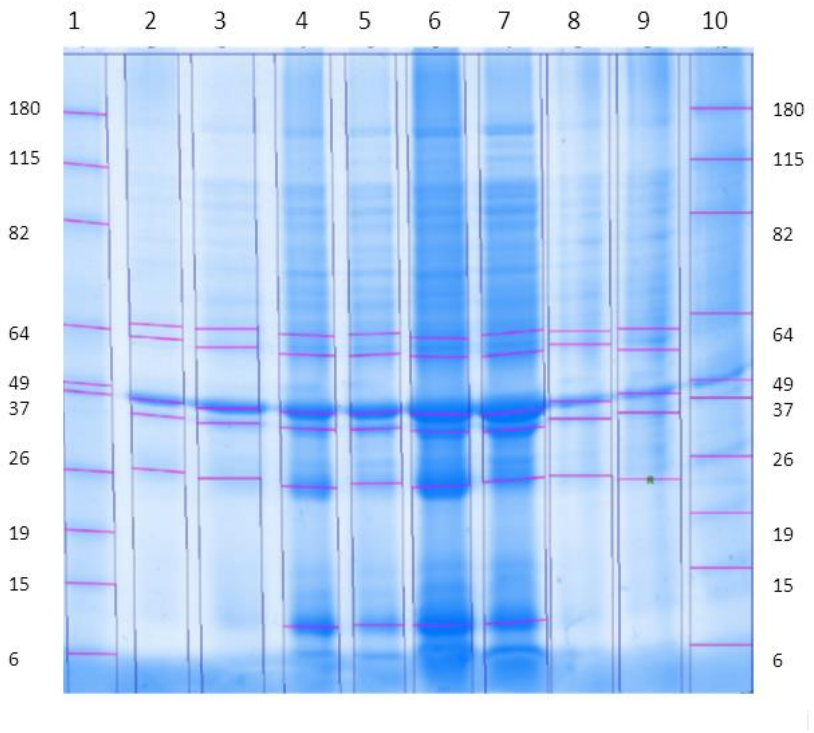

Figure 30: SDS-PAGE gel analysis: Comparative assessment of different preparation methods; Lane 2 and Lane 3 represent the homogenization supernatant; Lane 4, 6 , 8 = Sample 1 (Lane 4: homogenization pellet, Lane 6: homogenization pellet – dilution: 1:2, Lane 8: Pellet – TCA precipitation); Lane 5,7,9 = Sample 2 (Lane 5: homogenization pellet, Lane 7: homogenization pellet – dilution: 1:2, Lane 9: Pellet – TCA precipitation); Lane 1, Land 10 = ladder (6-180 kDa)

Furthermore, the influence of TCA precipitation on product purity prior to SDS-PAGE was analysed. Lanes 8 and 9 shown in Figure 30 represent the two TCA precipitated samples. It can be seen that these two lanes do not show the rhBMP-2 band at 13kDa, which indicates that the product was lost within TCA precipitation. Since TCA precipitation did not reveal positive results, protein quantification using the BCA assay was not possible. Since the exact concentration of the product of interest was unknown, the two samples were also analysed at different dilutions. A comparison of the diluted samples (factor=1:2), which are shown in Lane 6 and Lane 7 for the two samples, to the non-diluted samples (shown in Lane 4 and Lane 5), showed that all four lanes contain rhBMP-2. However, Lanes 6 and 7 show a slight smear. This can be most likely attributed to a disruption of the pellet due to a pipetting error.

In order to quantify purity and the signal-to-noise ratio (SNR) of the three samples drawn again from fermentation, further SDS-PAGE analysis of directly dissolved homogenization pellet samples was done in triplicates. The results of these three gels are visualized in Figure 31, Figure 32 and Figure 33. Since these three gels represent identical triplicates, the lane composition is also identical. The supernatant of each sample is shown in Lane 2, 4 and 6, whereas the pellets of the three samples are displayed in Lane 3, 5 and 7. Hence, Lane 3 represents the sample from DG011_1D, Lane 5 represents the sample from DG011_3G and Lane 7 represents the sample from DG011_4F.

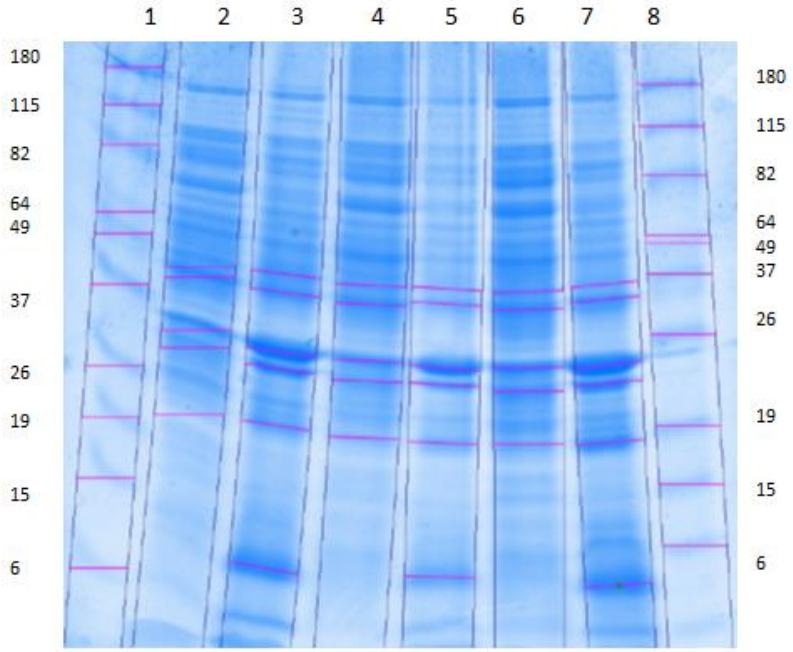

Figure 31: SDS-PAGE gel analysis – Gel 1; Sample 1: supernatant(homogenized) =Lane2, pellet(homogenized)=Lane3 ; Sample 2: supernatant(homogenized) =Lane4, pellet(homogenized)=Lane5; Sample 3: supernatant(homogenized) =Lane6, pellet(homogenized)=Lane7; Lane 1, Land 8 = ladder

When comparing Figure 31, Figure 32 and Figure 33 it can be seen that independent of the sample, the supernatants do not carry any rhBMP-2 (band at 13kDA). Contrary to that, rhBMP-2 is found in the three pellets of the samples (Lane 3, Lane 5 and Lane 7).

Besides the characteristic rhBMP-2 band (at 13 kDa) shown in Lane 3, 5 and 7, other bands are also visible in the three gels. These bands appear in all samples (independent if supernatant or pellet) and show proteins with a higher molecular weight than rhBMP-2. These bands represent so-called host-cell proteins (i.e. proteins that are constitutively expressed by *E.coli*) and are further referred to as host cell protein (HCP) contaminants.

The purity of rhBMP-2 inclusion bodies in respect to host cell protein (HCP) contaminants was assessed *via* densitometry evaluation of the SDS gels.

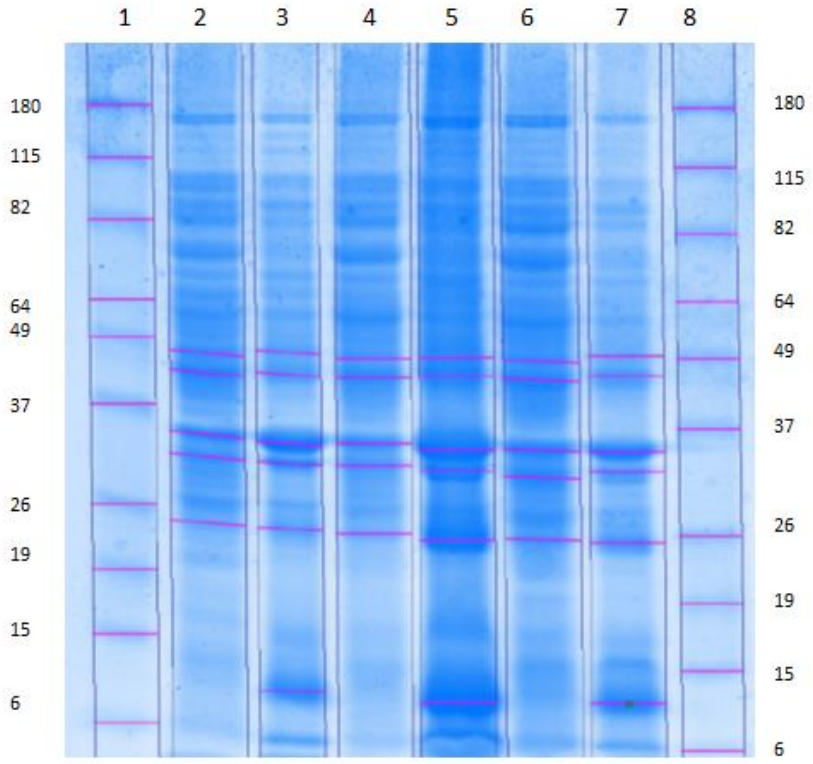

Figure 32: SDS-PAGE gel analysis – Gel 2; Sample 1: supernatant(homogenized) =Lane2, pellet(homogenized)=Lane3 ; Sample 2: supernatant(homogenized) =Lane4, pellet(homogenized)=Lane5; Sample 3: supernatant(homogenized) =Lane6, pellet(homogenized)=Lane7; Lane 1, Land 8 = ladder

It needs to be noted that the three pellets (Lane 3, 5, 7) origin from different fermentation runs that differ with respect to q_{s_total} and the temperature (compare Section 4.1). However,

when comparing the three gels shown in Figure 31, Figure 32 and Figure 33 it can be seen there is no visible influence of q_{s_total}.

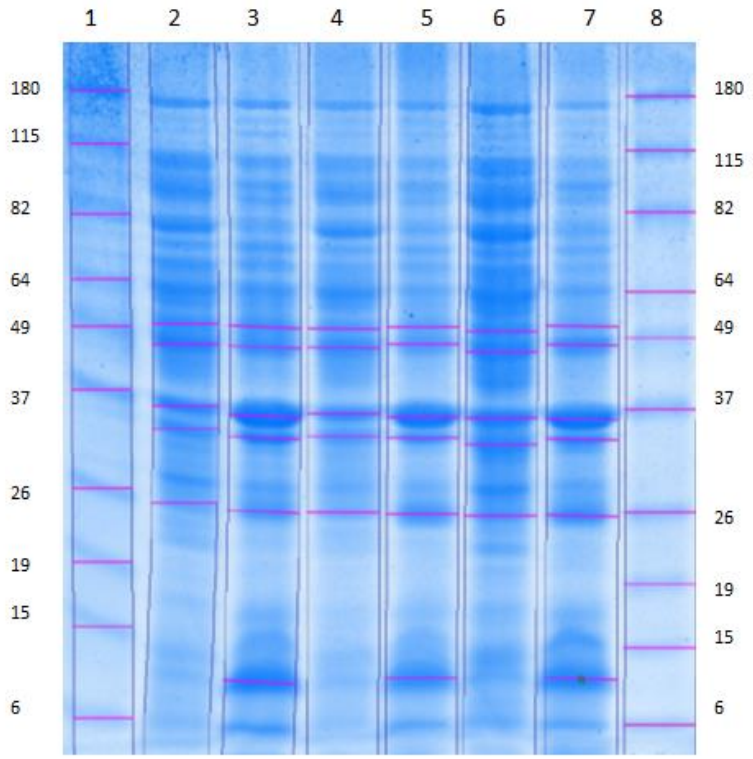

Figure 33: SDS-PAGE gel analysis – Gel 3; Sample 1: supernatant(homogenized) =Lane2, pellet(homogenized)=Lane3 ; Sample 2: supernatant(homogenized) =Lane4, pellet(homogenized)=Lane5; Sample 3: supernatant(homogenized) =Lane6, pellet(homogenized)=Lane7; Lane 1, Land 8 = ladder

However, quantification of purity revealed that product obtained from fermentation run: DG011_Unit4 shows higher purity (mean of the three gels: 42.09%) than the other two fermentation runs (see Figure 34). This suggests that temperature has a positive influence on product purity. This finding is in contrast to previously published literature, which suggests that decreasing temperature during protein production has been shown to be an effective way to raise solubility of proteins [Li P et al.; *Chinese journal of virology*]and thereby prevent inclusion body

formation [Song JM et. al, *Protein expression and purification* 2012]. Furthermore, DeGroot et al. suppose that IB formation at higher temperature results in resistance to digestion and is also associated with increased stability against chemicals [de Groot NS, *FEBS letters* 2006].

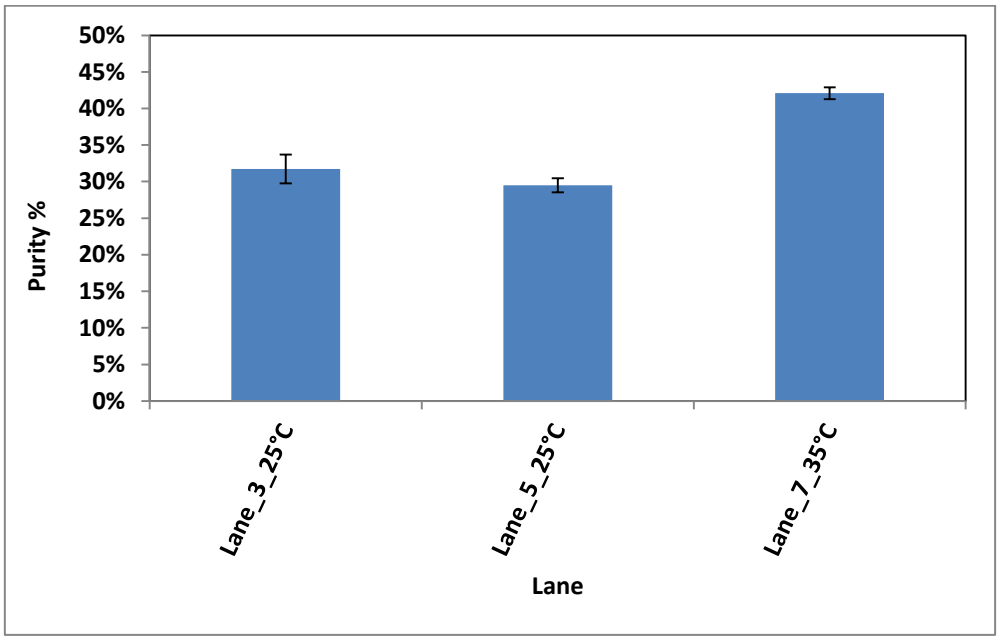

Figure 34: Purity of homogenized pellets at different q_s values and temperatures. Data are represented as means of % purity of the three gels (error bars indicate ±SEM)

Table 6: rhBMP-2 Purity of the different gels. Columns represent the homogenized pellets of the three gels. SEM=standard error of the mean; SNR-signal-to-noise ratio; LOD=limit of detection; LOQ limit of quantification.

	Lane 3	Lane 5	Lane 7
GEL_1	30.28%	27.78%	42.90%
GEL_2	35.42%	31.12%	41.29%
GEL_3	28.93%	29.58%	ND
Mean (purity)	31.72%	29.49%	42.09%
SEM	0.02	0.01	0.14
low_bound	29.75%	28.53%	28.06%
upp_bound	33.70%	30.46%	56.13%

Determination of the Signal-to-noise Ratio (SNR) and the Limit of Detection (LOD) and Limit of Quantification (LOQ)

Finally, the SNR, LOD and LOQ were determined for the three gels. The results are summarized in the table below. A change of product quality above 6.14 % can be detected and a change of quality above 20.48 % can be quantified.

Table 7: Determination of SNR, LOD, LOQ for the method "purity assessment *via* SDS gels"

mean_SNR	21.39
mean_LOD	6.14%
mean_LOQ	20.48%

Conclusion

In summary, the following conclusions can be drawn from the fermentation analytics for rhBMP-2:

- An analytical scheme for the bioanalysis of rhBMP-2 has been developed (see also Appendix)
- TCA precipitation and consecutive quantification of product *via* the BCA assay from homogenized pellets did not work, because the measurements were found to be outside of the ranges of calibration.
- Analysis *via* SDS-PAGE revealed that no protein is found in the supernatant. RhBMP-2 can only be detected from homogenized pellets.

Appendix

- Analytical scheme