Diplomarbeit

High throughput inclusion body sizing and closed loop physiological process control

ausgeführt zum Zwecke der Erlangung des akademischen Grades eines Diplom-Ingenieurs (Dipl.-Ing.)

eingereicht an der TU Wien, Fakultät für Technische Chemie

von

Andreas KAINEDER

Mat.Nr.: 0525258

unter der Leitung von

Univ.Prof. Dipl.-Ing. Dr. Christoph Herwig

Institut für Verfahrenstechnik, Umwelttechnik und Technische Biowissenschaften (E166)

begutachtet von

Univ.Prof. Dipl.-Ing.
Dr. Christoph Herwig
Institut für Verfahrenstechnik,
Umwelttechnik und Technische
Biowissenschaften (E166)
Getreidemarkt 9, 1060 Wien, Österreich

Associate Prof. Dipl.-Ing.
Dr.nat.techn.Oliver Spadiut
Institut für Verfahrenstechnik,
Umwelttechnik und Technische
Biowissenschaften (E166)
Getreidemarkt 9, 1060 Wien, Österreich

Eidesstattliche Erklärung

Ich nehme zur Kenntnis, dass ich zur Drucklegung meiner Arbeit unter der Bezeichnung Diplomarbeit nur mit Bewilligung der Prüfungskommission berechtigt bin.

Ich erkläre an Eides statt, dass die vorliegende Arbeit nach den anerkannten Grundsätzen für wissenschaftliche Abhandlungen von mir selbstständig erstellt wurde. Alle verwendeten Hilfsmittel, insbesondere die zugrunde gelegte Literatur, sind in dieser Arbeit genannt und aufgelistet. Die aus den Quellen wörtlich entnommenen Stellen, sind als solche kenntlich gemacht.

Das Thema dieser Arbeit wurde von mir bisher weder im In- noch Ausland einer Beurteilerin/einem Beurteiler zur Begutachtung in irgendeiner Form als Prüfungsarbeit vorgelegt. Diese Arbeit stimmt mit der von den Begutachterinnen/Begutachtern beurteilten Arbeit überein.

Wien, September 2018

Andreas Kaineder

ACKNOWLEDGEMENTS

Foremost I would like to express my honest gratitude to my supervisors Prof. Christoph Herwig and Dr. Wieland Reichelt for giving me the opportunity to carry out my thesis in such a great working environment. Without their guidance, inspiration and challenging as well as promoting supervision this thesis would not have been possible.

Special thanks to my colleagues Markus Brillmann, Peter Thurrold, Anna Bankuti and Banu Yildiz who always supported me with a helping hand and words of advice. It is furthermore important to acknowledge the help and support of many team members of the working group Herwig, in particular Jens Fricke, Oliver Spadiut, Paul Kroll and David Wurm.

Last but not least I want to express my deepest gratitude to my wife and partner in life Kristin. Thank you for inspiring me to resume my studies and showing me the light at the end of the tunnel. Without you, I would not have finished.

ABSTRACT

Quality by design (QbD) became an indispensable part of pharmaceutical quality. To increase product quality, understanding how formulation and manufacturing process variables influence product quality is essential.

In that holistic context, this work attempted to develop transferable methods to help increase process understanding, control and robustness. For the purpose of these goals a recombinant protein production process with *E. coli* as a host was used. The product was formed as inclusion bodies (IB). A high throughput method for IB sizing using nano-particle tracking analysis (NTA) was developed and established. The effect of temperature oscillations on IB solubility could not be investigated as planned due to insufficient heating and cooling performance of the reactor setup. A softsensor, previously developed in the working group, has been further developed and was used for real-time estimation of the specific substrate uptake rate q_S. Therefore, the obligatory real-time biomass estimation was decoupled from the carbon balance and realized using a weighted average approach, enabling biomass estimation robust against substrate accumulation. Subsequently, the sensor was used to establish a closed loop control of q_S. Because of the declining physiological capacity of the host to metabolize substrate (q_{Scrit}) during induction phase, a logical query was developed to detect reaching q_{Scrit} in real-time.

The physiological closed loop control was used to successfully avoid substrate and metabolite accumulation throughout induction phase of an industrial relevant production process.

DEUTSCHE KURZFASSUNG

Quality by Design (QbD) ist ein unverzichtbarer Bestandteil der pharmazeutischen Qualität geworden. Um die Produktqualität zu erhöhen, ist es wichtig zu verstehen, wie Prozessvariablen die Produktqualität beeinflussen.

In diesem ganzheitlichen Kontext wurde versucht, übertragbare Methoden zu entwickeln, die das Verständnis, die Kontrolle und die Robustheit von Prozessen verbessern. Für die Zwecke dieser Ziele wurde ein rekombinanter Proteinproduktionsprozess mit E. coli als Wirt verwendet. Das Produkt wurde als sogenannte inclusion bodies (IB) gebildet. Es wurde eine Hochdurchsatzmethode für IB-Größenbestimmung unter Verwendung von Nano-Partikel-Tracking-Analyse (NTA) entwickelt und etabliert. Die Auswirkung von Temperaturschwankungen auf die IB-Löslichkeit konnte aufgrund unzureichender Heiz- und Kühlleistung des Reaktoraufbaus nicht wie geplant untersucht werden. Ein zuvor in der Arbeitsgruppe entwickelter Softsensor wurde weiterentwickelt und zur Echtzeitschätzung der spezifischen Substrataufnahmerate q_S verwendet. Dafür wurde die obligatorische Echtzeit-Biomasse-Schätzung von der Kohlenstoffbilanz entkoppelt und mit einem gewichteten Mittelwert realisiert, was eine robuste Biomasse-Schätzung gegen Substratakkumulation ermöglicht. Anschließend wurde der Sensor verwendet, um einen geschlossenen Regelkreis für die Regelung von q_S zu realisieren. Aufgrund der abnehmenden physiologischen Kapazität des Wirts, während der Induktionsphase Substrat zu metabolisieren (q_{Scrit}), wurde eine logische Abfrage entwickelt, um das Erreichen von q_{Scrit} in Echtzeit zu erfassen.

Die physiologische Regelung wurde angewendet, um die Akkumulation von Substrat und Metaboliten während der Induktionsphase eines industriell relevanten Produktionsprozesses erfolgreich zu vermeiden.

TABLE OF CONTENTS

1 INTRODUCTION

1.1 MOTIVATION

In the bioprocess industry, as in many other industries too, the pursuit for optimization has led to increased productivity over the past years. While this boost is largely due to improvements in upstream processing (USP), downstream processing (DSP) has become a bottleneck in bioprocesses (Gottschalk 2008). Apart from optimizing DSP itself, controlling product properties in USP could also lead to debottlenecking of DSP. Although the approach of separately improving USP and DSP is simpler, a holistic approach originates from the concepts of Quality by Design (QbD) and is likely to increase over the next years (Calo-Fernández and Martínez-Hurtado 2012).

QbD is a systematic approach to pharmaceutical development that begins with predefined objectives and emphasizes product and process understanding and process control, based on sound science and quality risk management (ICH 2009). That means a QbD development process includes identifying critical process parameters (CPP), which need to be controlled to achieve critical material attributes of the final product (Lionberger, Lee et al. 2008).

1.2 BACKGROUND

1.2.1 PRODUCTION PLATFORM

E. coli remains one of the most important production platforms for recombinant protein production overall and the single most common nonmammalian-based production cell type, although the percentage use of *E. coli* as an expression system declined over the last years due to steadily increasing mammalian expression systems. From 2010 to 2014 almost a third (29 %) of all product approvals used *E. coli* as an expression system (Walsh 2014).

E. coli, as other microbial cells, lacks the ability to correctly fold complex molecular structures of eukaryotic derived recombinant proteins. This can cause the formation of inclusion bodies (IB), aggregates of intracellular proteins not properly folded (Williams, Van Frank et al. 1982). Nevertheless, the expression system comes with advantages prevailing the formation of IBs, like high growth rates, higher overall product yields, low media costs compared to mammalian cell lines and being the most studied microorganism (Swartz 2001), what facilitates genetic modifications. Another advantage is the N^{pro} technology, where IB formation is used deliberately to facilitate the expression of toxic proteins (Dürauer, Ahrer et al. 2010). In general IB formation comes with the advantage of less purification but with the downside of solubilisation and refolding, which impacts the overall product yield (Pan, Zelger et al. 2014). Although literature provides insight in the study of IBs itself (Ami, Natalello et al. 2006, Margreiter, Messner et al. 2008, Margreiter, Schwanninger et al. 2008, Upadhyay, Murmu et al. 2012), investigations of controlling IB quality attributes (IB-QA) like e.g. size, purity or solubility in USP to positively influence DSP are scarce.

1.2.2 CONTROL STRATEGY

As stoichiometric restrictions apply to IB formation, control of IB formation and its attributes should be possible using physiological process control. As shown by Wechselberg et al. (Wechselberger, Sagmeister et al. 2012) the specific substrate uptake rate q_S (see Equation 1), which can be controlled via the feed rate, is the most important process parameter for optimization. Several other publications also show that the feed rate is one of the most promising factors to increase productivity (Levisauskas, Galvanauskas et al. 2003, Sandén, Prytz et al. 2003, Ramalingam, Gautam et al. 2007, Babaeipour, Shojaosadati et al. 2008, Kavanagh and Barton 2008, Sagmeister, Schimek et al. 2014).

$$q_S = \frac{g_S}{g_X \cdot h}$$

Equation 1: Specific substrate uptake rate q_S

For q_S control a so-called first principle softsensor can be used for obligatory real-time biomass estimation (Wechselberger, Sagmeister et al. 2013). Implying no substrate accumulation, the softsensor uses the substrate feed rate and the carbon dioxide evolution rate CER to estimate the biomass formation rate r_X by taking the biomass composition into account. Knowing the initial biomass, q_S can be simply controlled via the substrate feed rate. By definition, this is an open loop (or feed forward) control, since the current state of the controlled variable is not measured or estimated.

The substrate feed rate plays a key role, because it largely impacts the physiological state of the cell. In case of overfeeding the physiological capacity to metabolize substrate (q_{Scrit}) unwanted overflow occurs, substrate accumulates and productivity is impaired (Jensen and Carlsen 1990). The fact that q_{Scrit} is not a constant (Schaepe, Kuprijanov et al. 2014, Reichelt, Brillmann et al. 2017) suggests closed loop (or feedback) control of q_S. While closed loop control is widely used to control technological variables like temperature or pH, the biggest challenge with physiological variables is to estimate their current state, since it is commonly not possible to directly measure it.

1.3 GOALS

The goal of this master thesis is to take one step towards custom made inclusion bodies and guarantee robustness in DSP through robust USP.

1. Inclusion body quality attributes
   a) Investigate effects of combined q_S and temperature oscillations during induction phase on inclusion body solubility in *E. coli*
   b) Establish novel high throughput method for IB sizing using nano-particle tracking analysis (NTA)
2. Physiological process control
   a) Real-time q_S estimation
   b) Closed loop control of q_S
   c) Real-time detection of reaching q_{Scrit}

1.4 HYPOTHESES

- Temperature oscillations during induction phase influence inclusion body solubility.
- IB sizing can be measured with NTA
- Closed loop control of q_s can be achieved by decoupling the biomass estimation from the federate to estimate a process value of q_s in real-time.
- Reaching of $q_s$crit during induction phase can be detected in real-time with a simple logical query.

2 MATERIALS & METHODS

2.1 LIST OF EXPERIMENTS

Table 1: Performed fermentation experiments and author's contribution

Experiment		Focus	Author's contribution		
Nr.	Code		Ferm.	Analytics	Data analysis
1	WR27A-D	qs control	-	-	X
2	WR35A-D	qs control	-	-	X
3	WR39A-D	Temp. osc.	X	X	X
4	WR40A-D	IB sizing	X	X	X
5	WR42A-D	IB sizing	X	X	X
6	WR43A-D	IB sizing	X	X	X
7	WR44A-D	closed loop control	X	X	X
8	WR45A-D	closed loop control.	X	X	X

2.2 STRAIN

A recombinant Escherichia coli strain (BL21 DE3) with a C-molar biomass stoichiometry $C_1H_{1.78}N_{0.25}S_{0.01}O_{0.45}$ and ash 5% was provided by an industrial partner and used for the fermentations. It produces an intracellular N^{pro}-fusion protein as inclusion bodies after a 1 mM IPTG induction.

2.3 REACTOR SETUP

The fermentations were conducted in a master slave reactor system. The master reactor (Sartorius BIOSTAT® C plus, Sartorius, Germany) had 10 l working volume and was used for the batch phase. It was equipped with a triple port for base, acid (not used) and feed addition, pH probe (Mettler Toledo USA), Pt_{100} temperature an optical pO_2 sensor (Mettler Toledo USA), a double jacket for temperature regulation and an off-gas cooler. Mixing was done with a triple rushton turbine stirrer and four additional baffles inside of the reactor. Process control was done using Lucullus PIMS (Seucurecell Switzerland). The four slave reactors (DASGIP® Parallel Bioreactor System, Eppendorf, Germany) had 2 l working volume and were used for the fed-batch and induction phase. The reactors were equipped with Pt_{100} temperature sensors, pH probes (Mettler Toledo USA), optical pO_2 sensors (Mettler Toledo USA, module DASGIP PH4PO4), triple rushton turbine stirrers and four

baffles for mixing (module DASGIP TC4SC4), heating pads and cooling fingers as well as cooling jacket for temperature control. Further they featured a triple port for base/feed addition and sampling. The DASGIP control software v4.5 revision 230 was used for data logging and control: pH, pO_2, temperature, stirrer speed and aeration (module DASGIP MX4/4). CO_2, O_2 concentrations in the off-gas were quantified by a gas analyser (module DASGIP GA4) using the non-dispersive infrared and zircon dioxide detection principle, respectively. Reactor contents were monitored using four scales (Sartorius, Germany).

2.4 PRECULTURE

The preculture was obtained from shake flask cultures (3*100 ml in 1 l Erlenmeyer flasks) inoculated with frozen cryo-stocks (-80 °C) using a chemically defined media (recipe not shown due to confidentiality constrains). After autoclaving the prefilled flasks, ampicillin (selective pressure), sterilised glucose glucose solution and trace element solution were added to the media before inoculation. Incubation was done in a shaker at 230 rpm and 30 °C. After reaching an OD600 of 1.5-3 (approx. 17h) the preculture was used for batch inoculation.

2.5 BATCH

The batch phase was performed in the master reactor. 10 l batch media (recipe not shown) was sterilized in situ. After autoclavation the pH of 6.7 was set with NH_4OH (12.5 % w/w) solution and sterile glucose solution was added to the media. 250 ml of the preculture was added through a septum using sterile syringes to the batch media. Temperature and pH were held constant at 30°C and 6.7 respectively. Aeration was done using pressurised air at a flow rate of 1.4 vvm. Stirrer speed was 400 rpm. Final biomass concentration was approximately 2.4 g/l. Batch end was indicated by a drop of the CO_2 off-gas signal to 0% and a sudden increase of the pO_2 signal.

2.6 FED-BATCH AND INDUCTION PHASE

At the end of the batch phase 1 l of fermentation broth was transferred via an autoclaved tube to each of the pre-sterilized slave reactors. Glucose feed solution was added using an exponential feeding profile corresponding to a specific growth rate μ of 0.17-0.22 h^{-1}, which is significantly lower than the maximal specific growth rate μ_{Max}. This leads to a self-controlling behavior of the cellular growth, resulting in a stable biomass concentration X and specific growth rate at induction time as shown by (Jenzsch, Gnoth et al. 2006). pH and temperature were kept at 7 and 30°C respectively, aeration was done using a flow rate of 1.4 vvm. Stirrer speed was kept at 1400 rpm. pO_2 was controlled to stay above 30 % by mixing pure oxygen to the in-gas using a step controller.

At a biomass concentration of 20-30 g/l (depending on the conducted experiment) an adaption phase of 30 minutes was started. The biomass concentration was calculated based on a feed forward profile with a constant biomass yield $Y_{x/s}$ of 0.4 g/g. During adaption phase feed rate to meet the intended q_S mean, pH and temperature were set for induction phase. After adaption phase sterile IPTG solution was added (1 mM final concentration) and the control script was started.

2.6.1 TEMPERATURE AND q_S OSCILLATION

The q_S setpoint was control by using a first principle softsensor as described in detail elsewhere (Sagmeister, Wechselberger et al. 2013). The softsensor K2S1 performs biomass estimation following a cumulative estimation approach by calculating metabolic rates in real-time based on an over determined equation system using the carbon as well as the Degree of Reduction (DoR) balance. In addition to the calculation itself it allows a consistency check of the observed equation system, which is given by the h-value as described by (Jobe, Herwig et al. 2003). The elemental biomass composition, the substrate concentration in the feed as well as an initial biomass concentration are obligatory input parameters. According to biomass growth the feed supply is adapted in order to maintain the q_S of interest. The flow diagram of the process is shown in Figure 1.

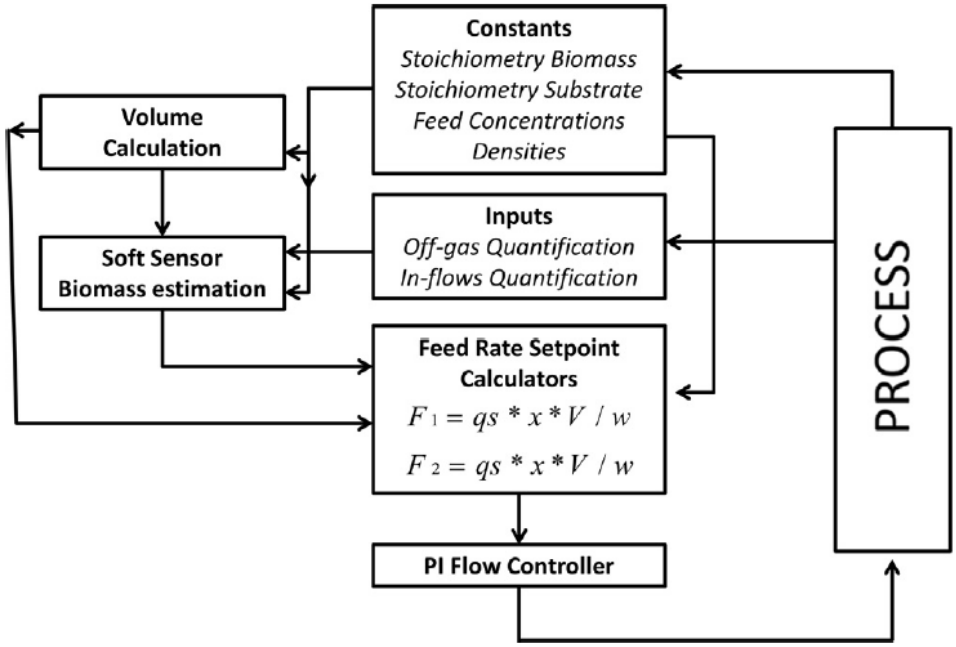

Figure 1: Flow diagram depicting the qs control based on first principle softsensor; Constants (Biomass elemental composition, Substrate elemental composition, Feed concentration, Densities), and online process signals (off-gas measurements and substrate inflow) are used as inputs for total biomass estimation; From this feed-rate set-points to maintain a certain qs are calculated (Sagmeister, Wechselberger et al. 2013).

The open loop control for q_S is shown in Figure 2. The feed rate is calculated accordingly to Equation 2 in ml/h. The setpoint of q_S in g/g/h was calculated using a visual basic script shown further below. Biomass in g was estimated in real-team as described above. Using a density correlation, the feed concentration c_S in g/ml was calculated beforehand.

$$\dot{F}_{0,i+1} = \frac{q_S SP_{i+1} \cdot X_i}{c_S}$$

Equation 2: Feed rate for desired q_S setpoint for next control interval

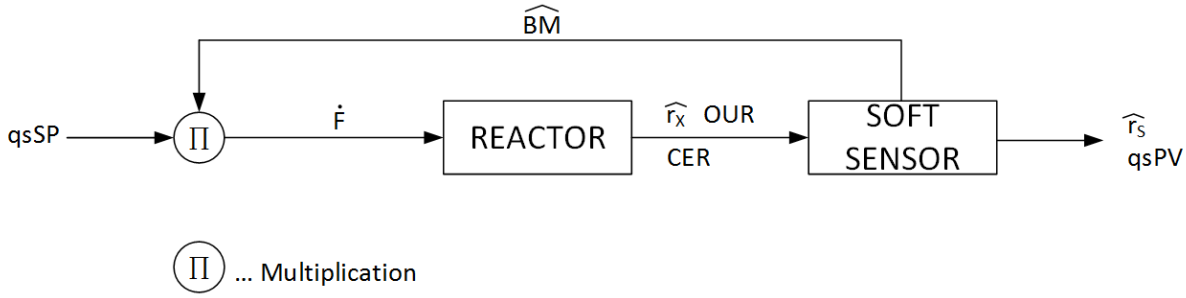

Figure 2: Open loop control for q_S including real-time biomass estimation based on a Softsensor with incremental calculations of the suitable substrate flow rate. The figure shows that the system does not return any information about the achieved process value of the controlled variable.

The code for data reconciliation of this softsensor K2S1 is shown in the appendix. The oscillations of q_S and temperature were controlled using the following Visual Basic code:

```
osc_time ...          time of constant qs/Temperature in hours
.InoculationTime...   timer started at induction

'oscillate qs between qs low and qs high
If (.InoculationTime_H/osc_time) mod 2D > 1
        qs = 0.2 'qs low
        else if (.InoculationTime_H/osc_time) mod 2D < 1
        qs = 0.6 'qs high
End if
'oscillate temp between temp low and temp high
If (.InoculationTime_H/osc_time) mod 2D > 1
        .TSP = 35 'Temp high
        else if (.InoculationTime_H/osc_time) mod 2D < 1
        .TSP = 20 'Temp low
End if
```

According to the Visual Basic code the maximum of q_s coincides with the minimum of the temperature and vice versa.

2.6.2 CLOSED LOOP CONTROL OF Q_S

2.6.3 CALCULATION AND CONTROL OF FEED RATE

A step controller was used for real-time q_S closed loop control during the post induction phase. Therefore, the feed rate $\dot{F}$ in ml/h was calculated with the base load $\dot{F}_0$ and the adjustment term $\Delta\dot{F}$. Feed rate was adjusted every 20 minutes. In Equation 3 the calculation of the feed rate $\dot{F}_{i+1}$ for the next control interval *i+1* is shown.

$$\dot{F}_{i+1} = \dot{F}_{0,i+1} + \Delta\dot{F}_{i+1}$$

Equation 3: Feed rate for next control interval *i+1*

The base load of the feed rate was calculated accordingly to Equation 4. The q_S setpoint for the next interval q_SSP_{i+1} in g/g/l was either a constant or a variable adjusted by the controller with setpoint adaption, used in some of the experiments. Using a density correlation, the

feed concentration c_S in g/ml was calculated beforehand. The current biomass X_i was estimated every 20 minutes as described in 2.6.4.

$$\dot{F}_{0,i+1} = \frac{q_S SP_{i+1} \cdot X_i}{c_S}$$

Equation 4: Feed rate base load for the next control interval *i+1*

Taking the difference of the q_S setpoint q_sSP_i and process value q_sPV_i during the last control interval *i* into account, the adjustment term $\Delta\dot{F}_i$ for the next control interval *i+1* was calculated accordingly to Equation 5. To prevent overshooting the adjustment term was limited to 30 % of the base load.

$$\Delta\dot{F}_{i+1} = \frac{(q_S SP_i - q_S PV_i) \cdot X_i}{c_S} \;\middle|\; \Delta\dot{F}_{i+1} \leq \dot{F}_{0,i+1} \cdot 0.3$$

Equation 5: Feed rate adjustment term for the next control interval *i+1*

2.6.4 REAL-TIME BIOMASS ESTIMATION

To decouple the estimation of biomass from the carbon balance, it was based on nitrogen balance, degree of reduction balance and a permittivity measurement. Results from prior experiments were taken to determine errors on the estimations based on the different approaches (N-, DoR-balance and permittivity measurement) and used to calculate an error-weighted average which was adapted from (Aehle, Simutis et al. 2010) according to Equation 6, to ensure more stable estimation during phase of induction. Real-time biomass estimation was started with time point of induction. The initial value for the estimation was calculated with a fixed-yield feed forward profile during fed-batch phase.

$$X_i = \frac{w_N \cdot X_N + w_{Dor} \cdot X_{Dor} + w_{Perm} \cdot X_{Perm}}{w_N + w_{DoR} + w_{Perm}}$$

Equation 6: Error-weighted biomass estimation in C-mol

Base concentration was determined from titration with standardized 1M HCl. The molecular composition of biomass was determined in previous work and provided the necessary information of mole nitrogen per mole carbon of biomass. In Figure 3 the linear correlation between biomass growth and consumption of base is shown.

$$X_{N,i} = X_{i-1} + \frac{\Delta V_{NH4} \cdot c_{NH4} \cdot M_N}{n_{N,XC} \cdot M_{NH4}}$$

Equation 7: Biomass estimation based on N-balance in C-mol

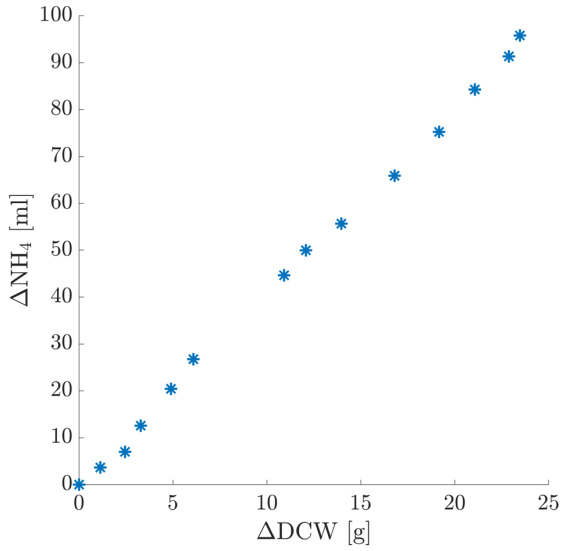

Figure 3: Base consumption – biomass growth correlation; The plot shows the linear correlation between biomass growth and consumption of base (NH4) during phase of induction of experiment WR45-A. The correlation was used to estimate the growth of biomass independently of the C-balance and therefor of substrate accumulation.

Beginning with the degree of reduction balance the biomass growth rate $r_{X,DOR}$ was calculated in C-mol/h/l according to Equation 8. r_{O2} was calculated from off-gas measurements and for r_S the estimation of the last interval was used.

$$r_{X,DOR} = \frac{4 \cdot (r_{O2} - r_S)}{4.13}$$

Equation 8: Biomass growth $r_{X,DOR}$ based in DOR-balance

With Equation 9 the biomass estimation in C-mol based on $r_{X,DOR}$ was calculated.

$$X_{DOR,i} = X_{i-1} + r_{X,DOR} \cdot V_{R,i} \cdot \Delta t$$

Equation 9: Biomass estimation in C-mol based on DOR-balance

During the fed-batch phase the permittivity measurement was calibrated with the feed forward profile with constant biomass yield, shown in Figure 4. According to Equation 10 the biomass estimation in C-mol based on the permittivity measurement was calculated.

$$X_{Perm,i} = \frac{V_{R,i} \cdot (\varepsilon_i \cdot a_\varepsilon + b_\varepsilon)}{M_{X,C}}$$

Equation 10: Biomass estimation in g based on permittivity measurement

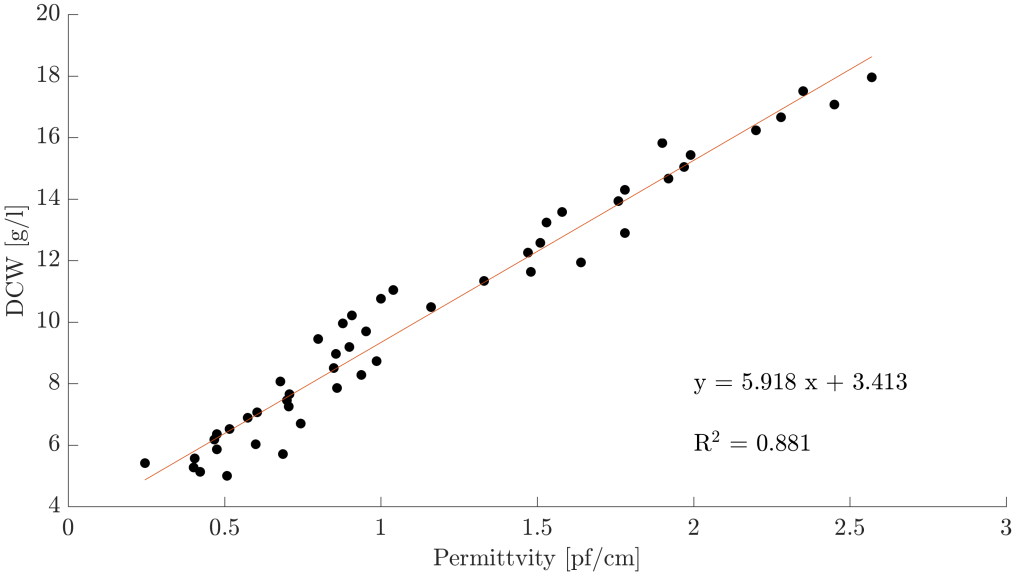

Figure 4: Biomass – permittivity correlation; The plot shows the linear correlation between DCW in g/l and permittivity measurement during the fed-batch phase of experiment WR45-A which was used to estimate DCW in g/l during phase of induction independently of the C-balance.

Based on former experiments the error e_i of each particular biomass estimation was estimated and used to calculate an error-weighted average of all estimations.

$$w_i = \frac{1}{X_i \cdot e_i}$$

Equation 11: Absolute error of particular biomass estimation

2.6.5 REAL-TIME ESTIMATION OF METABOLIC RATES

Biomass was estimated every 20 minutes and thus the amount of grown biomass within this time interval.

$$r_{X,i} = \frac{\Delta X_i}{\Delta t} = \frac{X_i - X_{i-1}}{t_i - t_{i-1}}$$

Equation 12: Estimation of biomass growth rate

To estimate the substrate conversion rate r_S the softsensor (described in 2.6.1) was modified to version K2S1_v2. Instead of the feed rate the biomass growth rate was an obligatory input enabling the real-time estimation and reconciliation of r_S. The code for data reconciliation for K2S1_v2 is shown in the appendix.

2.6.6 CLOSED LOOP CONTROL

Figure 5 shows the closed loop control for q_S. As described in 2.6.3 the process value $q_S PV$ is compared with the setpoint $q_S SP$ and provides Δq_S to calculate $\Delta \dot{F}$ (according to Equation 5) to improve the quality of the controller performance.

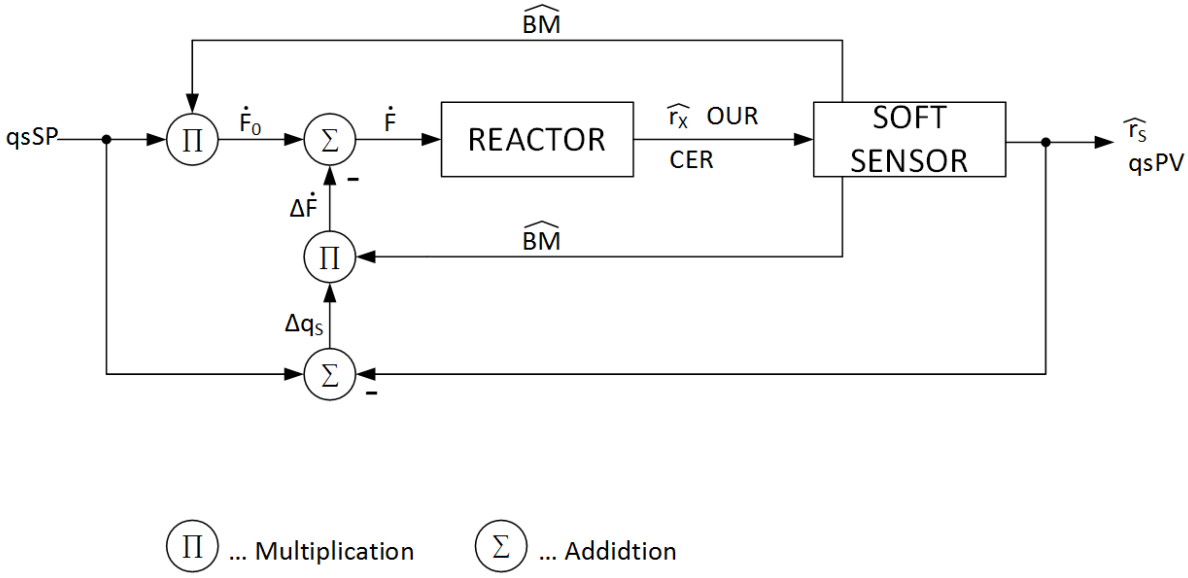

Figure 5: Closed loop control provides the possibility of comparing the wanted setpoint with the actually achieved process value. Therefore, an interference of the process can improved the quality of the control strategy

Additionally, to closing the control loop of q_S a logical query (shown in Figure 6) to adapt q_sSP if $q_s$crit is reached, was developed. In Figure 7 the closed loop control with adaption of q_sSP is shown. (Reichelt, Brillmann et al. 2016) showed a correlation between $q_s$crit decline and $q_s$mean. This coherence (see Equation 13) was used to assess the maximum possible decline of $q_s$crit within a control interval (20 minutes).

$$k = -0.26 \cdot q_{Smean} + 0.04$$

Equation 13: Calculation of $\boldsymbol{q_s crit}$ trajectory slope in dependency of $\boldsymbol{q_s mean}$

Based on the results for $q_s$mean of 0.4 g/g/h (-0.064 g/g/h^2) and 0.2 g/g/h (-0.012 g/g/h^2) it was decided to reduce the setpoint to 90 % in case of reaching the limit to ensure to fall below $q_s$crit again and maintain high growth rate. Three inputs need to be true to adapt the setpoint. The comparison of the q_sSP with the last q_sPV from point of time t_{i-1} 20 minutes ago, a comparison of q_sSP with the current q_sPV, point of time t_i, and the change of the q_sPV within the last 20 minutes (Δq_sPV). In case of all three conditions, q_sSP will be reduced to 90 %. Since q_sPV should always increase due to the adapted federate, if q_sPV is smaller than q_sSP and $q_s$crit is not reached.

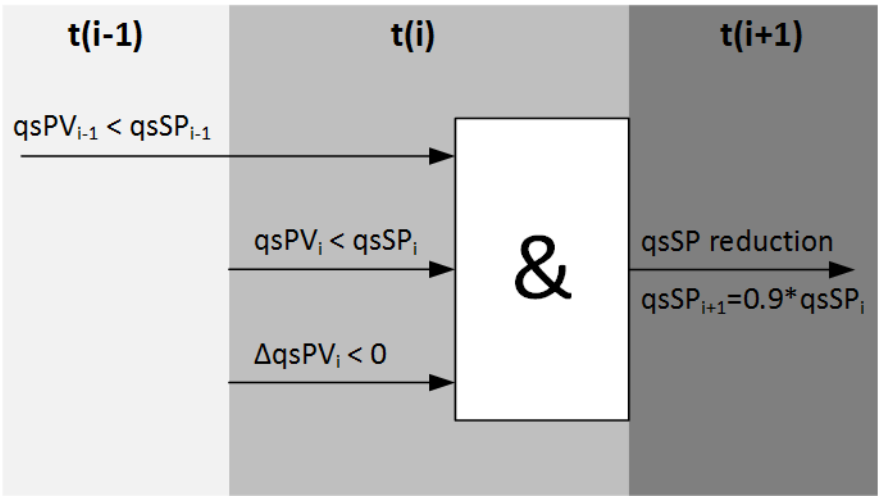

Figure 6: Logical query for closed loop control to adapt the setpoint of qs if qs crit is reached and the setpoint cannot be achieved.

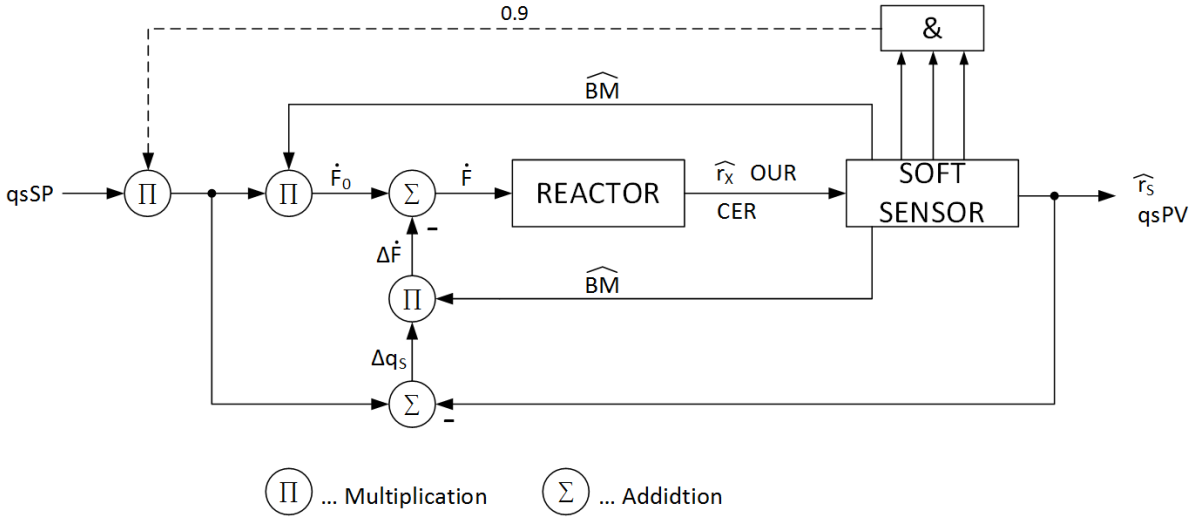

Figure 7: Closed loop control with logical query to adapt qsSP provides the possibility to stay within the oxidative maxima of the culture

2.7 ANALYTICS

2.7.1 BIOMASS DRY CELL WEIGHT

Throughout the induction phase BM samples were taken every 30 min using an automated sampling device consisting of two pump modules (two tubes per module), and an autosampler with a cooling block holding up to 50 vials. This autosampler consisted of four tubes each connected to one sampling port of one of the slave reactors, two tube pumps (two tubes per pump) and a robotic arm to maneuver the tubes to the corresponding sampling vials, which have been placed in the cooling block. The sampling procedure featured a tube flushing step of two minutes followed by the actual sampling into the sampling vials (30 seconds). Upon activation the robotic arm automatically counted the number of previous samples and therefore moved to the next sampling vial for the subsequent sampling cycle. The automated sampling device was controlled using Lucullus PIMS. Samples were stored at 4°C until the end of the fermentations. Vial volume was measured gravimetrically based on a density of 1 kg/m^3, afterwards the suspension was centrifuged (5000 rpm) and the pellet

was washed with deionized water, centrifuged again (5000 rpm) and finally dried at 110°C for at least 72 h before weighing on an analytical scale.

2.7.2 MANUAL FERMENTATION SAMPLES

6 ml aliquots of fermentation broth were centrifuged, two aliquots of 1 ml of supernatant stored in Eppendorf tubes (rest discarded) and both (Pellets and supernatant tubes) stored on -20°C until further analysis.

2.7.3 SUPERNATANT ANALYSIS

Glucose, acetate and ammonia contents in fermentation supernatant were measured using enzymatic test kits for Cedex Bio HT Analyzer (Roche, Switzerland) for the manually taken fermentation samples.

2.7.4 INCLUSION BODY SIZING

Methods for IB sizing was performed using nano-particle tracking analysis and is described in detail in the publication (Reichelt, Kaineder et al. 2017), which was a result of this work.

2.8 DATA ANALYSIS

Matlab R2015a (MathWorks, USA) was used for calculation of rates and yields from online data (feed rate, gassing rate, off-gas analysis) and offline data (BM dry cell weight, Glucose, acetate and ammonia measurements in SN). Reaction rates were all normalized by the reaction volume. For calculation of r_x a quadratic fit for the total biomass (shown in Figure 8) was used to minimize the effect of error propagation which could lead to artifacts caused by sampling interference. The novel approach used for calculation of q_s based on off-gas measurements and offline biomass is elsewhere (Reichelt, Brillmann et al. 2017).

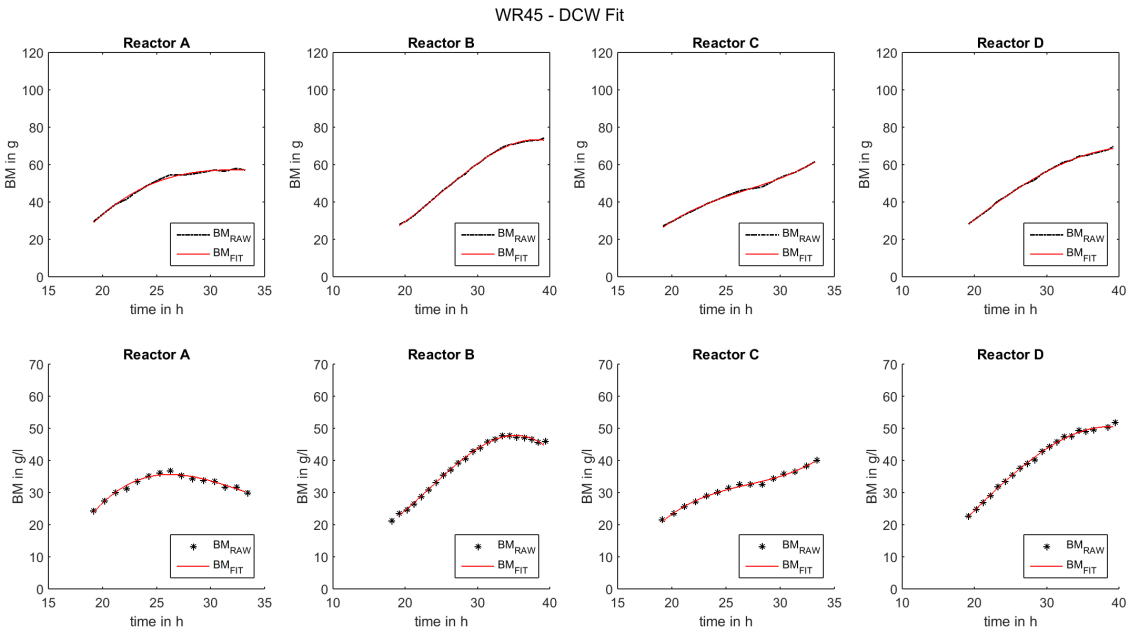

Figure 8: Biomass fitting for rate smoothing. The plots show the quadratic fits used for total biomass (top row) and biomass concentration (bottom row) used to give a smoothened r_X upon rate calculation. The red line shows the fit and the black dashed line shows the actual measured offline biomass.

3 RESULTS AND DISCUSSION

3.1 INCLUSION BODY QUALITY ATTRIBUTES

3.1.1 TEMPERATURE AND Q_S OSCILLATIONS

The results of experiment WR40-C led to ceasing further experiments dealing with temperature oscillations in combination with q_S oscillations because the reactor setup did not provide sufficient heating and cooling performance. As shown in Figure 9 the lower temperature setpoint could not be reached and the time to reach the higher setpoint was approximately 30 minutes. With the given setup it was not possible to run temperature oscillations within the desired conditions in terms of frequency and amplitude.

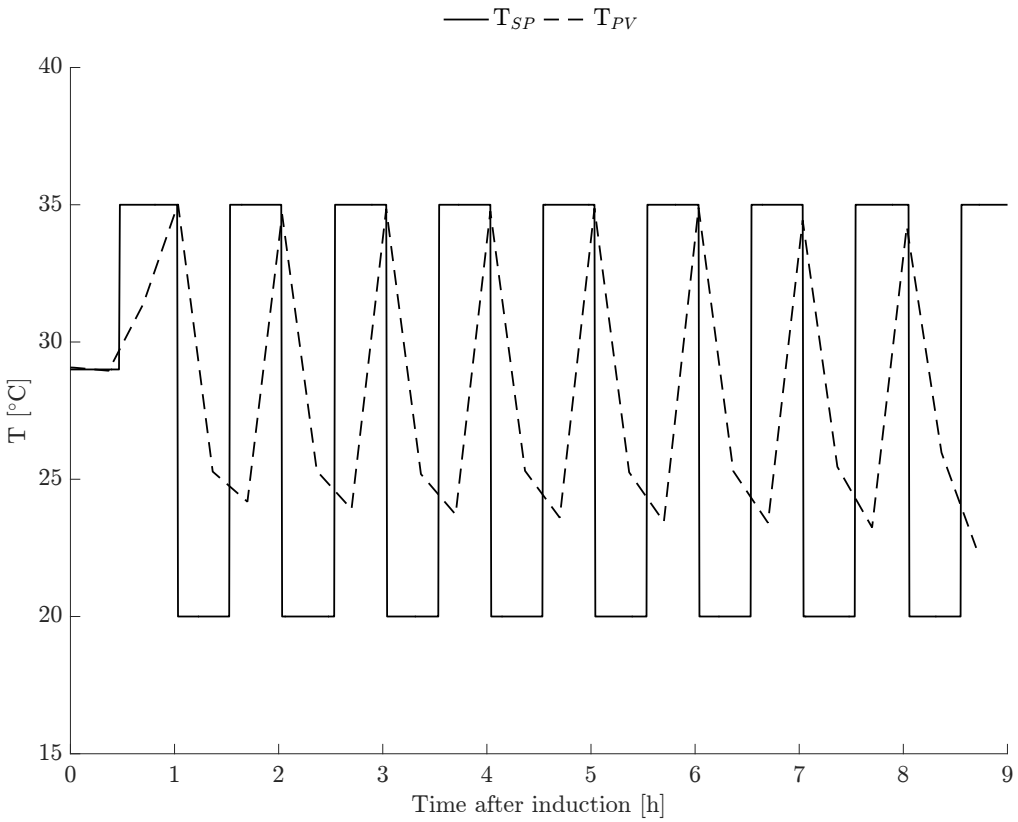

Figure 9: Temperature oscillations during phase of induction; The plot shows large deviations of the actual process value TPV (dashed line) from the desired setpoint TSP (solid line) due to insufficient heating and cooling performance

Furthermore the constantly changing temperature led to problems with pH and pO_2 control as shown in Figure 10. With the given setup it was not possible to provide stable pH and pO_2 control going hand in hand with temperature oscillations. It would have been necessary to

add acid to the pH control and determine the pH-temperature dependency of the system to compensate the temperature changes for the pH measurement.

The pO_2 control was done by a simple step controller as described in 2.3 and did not ensure a pO_2 higher than 30 % at all times. Brillmann (Brillmann 2015) showed in his master thesis a correlation of inclusion body properties (solubilisation kinetics) and the amplitude of q_S oscillations at constant temperature and pH as well as pO_2 higher than 30 % during induction phase. de Groot et al. (de Groot and Ventura 2006) showed that lower temperature also favors inclusion body solubility. A potential increase of these positive effects due to temperature oscillations could not be proved if pH and pO_2 cannot be controlled in a reproducible manner.

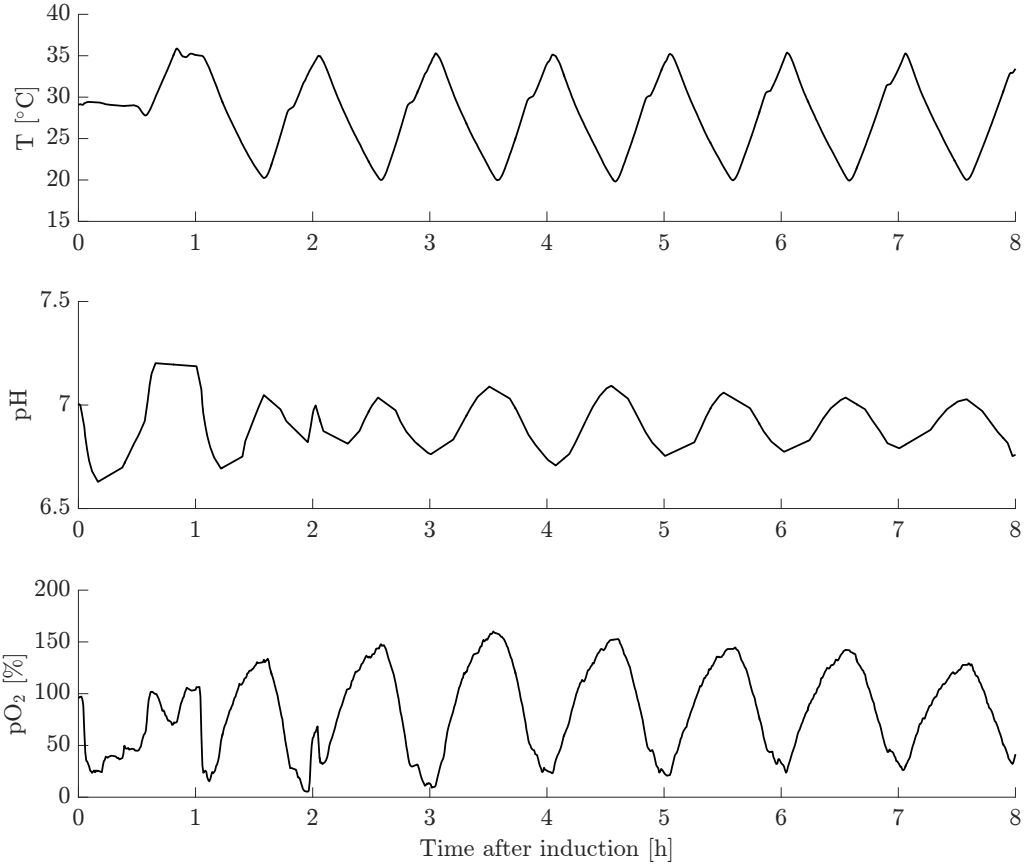

Figure 10: Quality control of WR40-C; Temperature oscillations (upper plot) impacts on pH (middle plot) and pO2 (lower plot) control, which could not be controlled in the desired way.

3.1.2 INCLUSION BODY SIZING – PUBLICATION PART

In the following section the accepted publication is covering the results of this thesis concerning inclusion body sizing. The contributions of the work to the following publication accompanied with the writing of this thesis lie in the execution of the fermentation experiments, the development and execution of the fixation and staining protocol as well as the execution of the NTA measurements.

Biotechnology Journal
www.biotechnology-journal.com

Biotechnol. J. 2017, 12, 1600471

DOI 10.1002/biot.201600471

Research Article

High throughput inclusion body sizing: Nano particle tracking analysis

Wieland N. Reichelt[1], Andreas Kaineder[1], Markus Brillmann[1], Lukas Neutsch[2], Alexander Taschauer[4], Hans Lohninger[3] and Christoph Herwig[1,2]

[1] Christian Doppler Laboratory for Mechanistic and Physiological Methods for Improved Bioprocesses, Institute of Chemical Engineering, Vienna University of Technology, Vienna, Austria
[2] Research Division Biochemical Engineering, Institute of Chemical Engineering, Vienna University of Technology, Vienna, Austria
[3] Research Group Electronic Media; Institute of Chemical Technologies and Analytics; Vienna University of Technology, Vienna, Austria
[4] Division of Clinical Pharmacy and Diagnostics, University of Vienna, Vienna, Austria

The expression of pharmaceutical relevant proteins in *Escherichia coli* frequently triggers inclusion body (IB) formation caused by protein aggregation. In the scientific literature, substantial effort has been devoted to the quantification of IB size. However, particle-based methods used up to this point to analyze the physical properties of representative numbers of IBs lack sensitivity and/or orthogonal verification. Using high pressure freezing and automated freeze substitution for transmission electron microscopy (TEM) the cytosolic inclusion body structure was preserved within the cells. TEM imaging in combination with manual grey scale image segmentation allowed the quantification of relative areas covered by the inclusion body within the cytosol. As a high throughput method nano particle tracking analysis (NTA) enables one to derive the diameter of inclusion bodies in cell homogenate based on a measurement of the Brownian motion. The NTA analysis of fixated (glutaraldehyde) and non-fixated IBs suggests that high pressure homogenization annihilates the native physiological shape of IBs. Nevertheless, the ratio of particle counts of non-fixated and fixated samples could potentially serve as factor for particle stickiness. In this contribution, we establish image segmentation of TEM pictures as an orthogonal method to size biologic particles in the cytosol of cells. More importantly, NTA has been established as a particle-based, fast and high throughput method (1000–3000 particles), thus constituting a much more accurate and representative analysis than currently available methods.

Received	06 AUG 2016
Revised	06 MAR 2017
Accepted	16 MAR 2017
Accepted article online	16 MAR 2017

Supporting information available online

Keywords: Electron microscopy · Grey scale image segmentation · Inclusion body · Nano particle tracking analysis · Stickiness

1 Background

The production of biosimilars is one of the main growth markets in pharmaceutical industry. Especially *Escherichia coli*, as well characterized expression host, has been established as easily accessible host for fast and efficient, high titer protein production. Hereby, high titer expression of heterologous protein frequently leads to inclusion body (IB) formation. This protein aggregation either coincides with high cytosolic concentrations of unfolded protein or can be induced, using a protein tag in order to reduce the toxicity of an otherwise toxic protein. However, while USP is hardly affected by IB formation, DSP

Correspondence: Prof. Christoph Herwig, Research Division Biochemical Engineering, Institute of Chemical Engineering, Vienna University of Technology, Gumpendorfer Strasse 1A/166-4, 1060 Vienna, Austria
E-mail: christoph.herwig@tuwien.ac.at

Abbreviations: BSA, bovine serum albumin; **DLS**, dynamic light scattering; **DSP**, down stream processing; **EDTA**, ethylenediaminetetraacetic acid; **FL**, fluorescence; **HPF**, high-pressure freezing; **IB**, inclusion body; **IPTG**, isopropyl-β-D-thiogalactopyranosid; **NTA**, nano particle tracking analysis; **PBS**, phosphate buffered saline; **TEM**, transmission electron microscopy; **USP**, up stream processing

constitutes the bottleneck in IB related production processes [1] and causes the bigger share of the total production costs.

As end product of USP, IBs are isolated by cell disruption prior to further processing during DSP. Industrial IB isolation is commonly conducted using high-pressure homogenization [2] and a continuous centrifugation to release and isolate IBs from the cells. Hereby, the continuous addition of washing buffer allows the combination of cell disruption and the removal of cellular debris in one unit operation. Post isolation, IBs are commonly solubilized in a chaotropic solubilization buffer, prior to refolding the protein into the native and therefore active protein conformation.

Integrated bioprocess development [7] strives to debottleneck the production process and to increase efficiency by addressing the impact of USP on DSP [8, 9]. This calls for sensitive response parameters describing the characteristics of IBs as intermediate product of USP and DSP. Solubilization in particular, as well as, refolding efficiency and purity appear to impair yields during DSP, which is why these steps and their efficiency have been investigated comprehensively [3–6].

Various methods have been investigated in order to characterize IBs, by paying attention to their different chemical properties. IB purity directly affects the necessary effort for further purification post-refolding and can be easily analyzed by SDS-PAGE [10]. Furthermore, IB solubility is critical for DSP performance, since highly soluble IBs would dissolve during the washing steps. In contrast, barely soluble IBs require high amounts of chaotrope reagents during solubilization, which increases buffer volume for refolding [11]. The increased volume in turn calls for bigger column diameters of economically expensive DSP purification columns [12]. Recent developments have enabled the concise measurement of solubility in respect to time [6, 8] as well as in respect to the concentration of chaotrope reagents [13].

Presumably mainly physical IB particle properties such as particle size and stickiness impact the yield of the isolation step prior IB solubilization during DSP. Nevertheless, the majority of available methods measure properties of the IB mass rather than properties of single particles. To date, analytical methods to quantify physical properties like the size and shape of IBs appear as being less developed.

Although a lot of effort has been invested especially into the quantification of IB size [14–18], the current, established methods have hardly been challenged/verified by an orthogonal verification method or are not single particle based (OD, DLS). Within this contribution we aim to establish an orthogonal verification method to analyze the size of large numbers of IBs and to assess sensitivity and information content of the nano particle tracking analysis (NTA) as a high throughput method to analyze IBs.

Besides being sensitive, a suitable method to effectively characterize the physical properties of IBs needs to be robust and reproducible. Mainstream adoption in academic as well as in industrial labs will only happen if the method is sufficiently simplistic. Highly sophisticated methods often lack technical transferability and comparability due to a certain degree of equipment and operator dependency.

1.1 Centrifugation based techniques

Early studies used centrifugation techniques, as centrifugal disc photo sedimentation [19] or cumulative sedimentation analysis [17], but require a particle density for the calculation of a size distribution of IBs. The more recently discussed approach of using an analytical centrifuge for IB sizing also relies on the density [20]. This dependency on IB density calls for an additional analytical method for IB density measurement, which makes methods relying on the density more laborious and less direct.

1.2 Dynamic light scattering

Dynamic light scattering (DLS) has extensively been utilized to size biological nano particles [11, 13–15, 21, 22]. Nevertheless, since this method only measures one variable its sensitivity is greatly impaired by multimodal distributions as well as by background particles [17]. As a counteraction, sample purification by serial washing steps [11, 13] or full-grown purification techniques as ultracentrifugation [15] have been investigated but these measures increase the risk of a measurement bias caused by sample preparation.

1.3 Field flow fractionation

Field flow fractionation (FFF) as separation or purification technique, as described elsewhere [23], has a wide dynamic range from 0.3–100 µm of particle separation capacity. The separation mechanism is a combination of Brownian motion, sedimentation and hydrodynamic lift forces [23] and facilitates bulk separation of nano particles according to their respective size and mass. Luo et al. used asymmetrical FFF in combination with multi-angle light scattering in order to analyze the size distribution of GFP inclusion bodies in response to induction time and temperature [14]. Using a sedimentation FFF in combination with a UV-Vis detector Margreiter et al. investigated the impact of inducer concentration and induction time on IB size [16]. Thereafter, an increase in the median spherical diameter of up to 140 nm over induction time was observed. Nevertheless, the effort for FFF method establishment is substantial and the potential interplay of different separation mechanisms implies that the interpretation of the results may not be so straightforward.

1.4 Imaging/TEM

The majority of the previously described methods and research contributions feature transmission electron microscopy, in an attempt to verify drawn conclusions. TEM facilitates conclusions based on single particle analysis by making single IBs visible. Given the overall goal of IB analytics of characterizing IBs as product of USP, the IBs should be analyzed in the most native conformation possible. Hereby, imaging of IBs in the cytosol excludes most of the otherwise necessary sample preparation and therefore a potential analytical bias. Nevertheless, it is indicated to analyze as many IBs per sample as possible in order to obtain a statistical representation of the IB population in the sample. Sizing IBs using TEM is commonly based on a laborious manual image analysis of the TEM image [16, 24]. In addition, the effort for sample preparation, the analytical technique, and image evaluation is substantial and basically excludes the possibility of using the usage of EM-based methods as routine analytical technique.

1.5 Nano particle tracking analysis

In this contribution we introduce nano particle tracking analysis (NTA) [25] as method to analyze and size a large number of biologic nano particles individually. For biologic particles the dynamic range of NTA spans 100–1500 nm, which fits the reported size range of IBs from 170 to 1300 nm [13, 16, 19, 24]. NTA uses a laser as light source, which passes through the sample particle suspension and illuminates the particles. In scatter mode, the scattered light, while in fluorescence mode the emitted light is recorded by a high-speed camera through a microscope. Due to the angle between the light beam and the camera axis, individual particles can be tracked and analyzed. At a constant temperature and a constant viscosity of the liquid, the size of each particle correlates to the Brownian particle movement. Using the Stokes–Einstein equation, the individual particle size can consequently be calculated resulting in a histogram of the particle size distribution of the particles in suspension.

The overall goal of this contribution is to provide a cost and time efficient method to quantify IB size. Firstly, for method verification, we aim to establish grey scale image segmentation of TEM pictures as an orthogonal method to assess IB size. Secondly, as cost and time efficient method, NTA is assessed as method to quantify IBs and their size in the background of cell debris. Finally, as an exemplary application the growth of IBs is investigated over process time.

2 Materials and methods

2.1 Bioreactor system

The fermentations were conducted in a DASGIP multi-bioreactor system (4Force; Eppendorf; Germany) with a working volume of 2 L each. The DASGIP control software v4.5 revision 230 was used for data logging and control: pH (Hamilton, Reno, USA), pO2 (Mettler Toledo; Switzerland;), temperature and stirrer speed (module DASGIP TC4SC4; Eppendorf; Germany), aeration (module DASGIP MX4/4; Eppendorf; Germany) and pH (module DASGIP MP8; Eppendorf; Germany). CO_2, O_2 concentrations in the off-gas were quantified by a gas analyzer (module DASGIP GA4; Eppendorf; Germany) using the non-dispersive infrared and zircon dioxide detection principle, respectively.

2.2 Cultivations

A recombinant BL21 DE3 *E. coli* strain was cultured, producing an intracellular protein (≈30 kDa) in the form of inclusion bodies, after a one-time induction with IPTG (1 mM). The synthetic media was based on the recipe from Korz et al. [26], where the limiting C-source was glucose.

Pre-cultures were grown to a OD600 of approx. 1.5 in 150 mL media. 2.5% of the batch volume was added as pre-culture for inoculation. The strain was cultivated at a controlled pH level, dissolved oxygen DO_2 (>30%) and temperature. The DO_2 was kept over 30% by supplementing oxygen to the air. After depletion of the C-source in an initial batch phase, the pre-induction fed-batch was started. The pre-induction feeding strategy was based on an exponential feed forward profile to maintain a predefined growth rate [27]. Upon induction, stirrer speed was set to 1400 rpm and aeration to 1.4 v/v/m for the whole process. The pH was maintained by adding 12.5% NH_4OH, which also served as nitrogen source.

2.3 Imaging

For high-pressure freezing (HPF) *E. coli* samples were pelleted and re-suspended in 5% BSA. After a second centrifugation step, the pellet was immediately frozen in a high-pressure freezer (HPF Compact 01; Wohlwend; Switzerland). The samples were then transferred into a freeze substitution unit (EM AFS2; Leica Microsystems; Germany) for water substitution with 2% uranyl acetate in anhydrous acetone over five days (–140°C to –90°C for 3 h; –90°C to –90°C for 25 h; –90°C to –54°C for 18 h; –54°C to –54°C for 8 h; –54°C to –24°C for 10 h; –24°C to –24°C for 15 h; –24°C to 0°C for 12 h;0 to 0 for 2 h)

For chemical fixation, the supernatant of the pelleted *E. coli* samples was carefully aspirated and cells were fixed using 2.5% glutaraldehyde in 100 mM cacodylate

buffer at pH 7.4 for 1 h at room temperature. After washing in the same buffer samples were post-fixed in 2% osmium tetroxide in cacodylate buffer, washed and dehydrated in a graded series of ethanol.

The dehydrated specimens were embedded in agar 100 resin (AGR 10131; Agar Scientific Ltd; UK) and after hardening, ultrathin sections (70 nm) were prepared (Leica ultramicrotome UCT; Leica Microsystems; Germany). The 70 nm sections were collected on 100 mesh Cu/Pd grids with a supporting formvar film. Sections were post-stained with 2% aqueous uranyl acetate, followed by incubation with Reynold's lead citrate. Images were collected using a transmission electron microscope (Morgagni 268D; FEI; Netherlands) operated at 80 kV and equipped with an 11 megapixel camera (Morada CCD; Olympus-SIS; Germany). Images were collected in both, regions of random sections and regions of one section.

2.4 Image segmentation

To quantify the IB size, the relative area of IB per cell was calculated based on grey scale image segmentation. The thresholds for background and IBs were selected manually by the operator, specific for each picture. The difference in area of the background and total image area corresponds to the area covered by cells. A pre-test with a larger number of operators (9) substantiated that the image segmentation is insignificantly impacted by the operator and can be regarded as transferable in-between operators (data not shown). Image segmentation of the 17 different samples and one negative sample with three to six images for each sample was conducted in Image Lab (v.1.02, Epina GmbH, Pressbaum, Austria; http://www.imagelab.at). 160–380 individual cells were repetitively analyzed per sample.

2.5 Cell disruption

2 mL of the fresh culture broth were centrifuged (4500 × *g*; 10 min; 4°C). The cell pellets were re-suspended in 20 mL 0.1 M Tris-buffer; 10 mM EDTA (pH 7.4) buffer and were disrupted in a high-pressure homogenizer (EmulsiFlex; Avestin; Canada) at 1400 ± 100 bar in six passages. For chemical fixation, 0.2% glutaraldehyde (G7776; Sigma Aldrich; Austria) was added dropwise to the re-suspended pellet and incubated 1 h at 4°C prior to homogenization.

2.6 Fluorescence stain

To discriminate cell debris from IBs, the homogenized cell pellet (5000 × *g*; 5 min) was re-suspended and incubated for 30 min in a 1x PBS solution containing 1% BSA and 2.2 mg/L of a product specific biotinylated primary antibody (courtesy of Sandoz GmbH; Austria). After washing with 1x PBS 1% BSA once, the pellet was re-suspended and incubated for 30 min in 1x PBS 1% BSA containing 10 µg/mL secondary IgG antibody labelled with Alexa 488 (AT11001; Invitrogen Life Technologies; Austria). Prior to measurement, the suspension was washed and re-suspended in 1x PBS.

2.7 Nano particle tracking (NTA)

A NS500 (Malvern, UK) software release (Nano Sight 3.0) equipped with a 488 nm laser and a CMOS camera (Hamamatsu Photonics, Japan) was used for the conducted NTA measurements. Most of the software parameters and algorithms are proprietary and are not known to the authors. The measurement chamber was primed prior to each measurement with 1x PBS to minimize particle drift. In-between measurements the chamber was flushed twice to avoid sample carryover. All samples were sonicated 1 min prior to measurement and diluted 1:10 in PBS. The focus level was set automatically, a standardized camera level of 16 was used in combination with a detection threshold of 20. Six replicates, 90 s each were measured with a 5 s time delay at a controlled temperature of 25°C.

2.8 Titer quantification

Product titer was measured using RP-HPLC after solubilizing the washed pellet of disrupted cells in guanidine hydrochloride. To calculate the specific titer, biomass concentrations were gravimetrically quantified after drying at 105°C for 72 h. Therefore 2 mL of culture broth were centrifuged (4500 × *g*, 10 min, 4°C) in a pre-weighted glass tube and the pellet was washed once with 5 mL RO water. The determination was done in duplicates. After drying in the drying oven, the biomass dry weight was measured on a scale.

2.9 Statistical data analysis

Data were subjected to statistical analysis using Datalab Version 3.5 (Epina GmbH, Pressbaum, Austria, http://datalab.epina.at/). Based on an $\alpha = 0.05$ the significance of the correlation was evaluated based on the p-value. Data were not transformed. Serial correlation was tested using the Durbin-Watson statistic.

3 Results

3.1 TEM+HPF as a gentle method for IB visualization

Currently, there is no verified method available to quantitatively size a representative number of IBs. This circumstance hinders the establishment of a high throughput method for quantitative IB sizing. Consequently, method

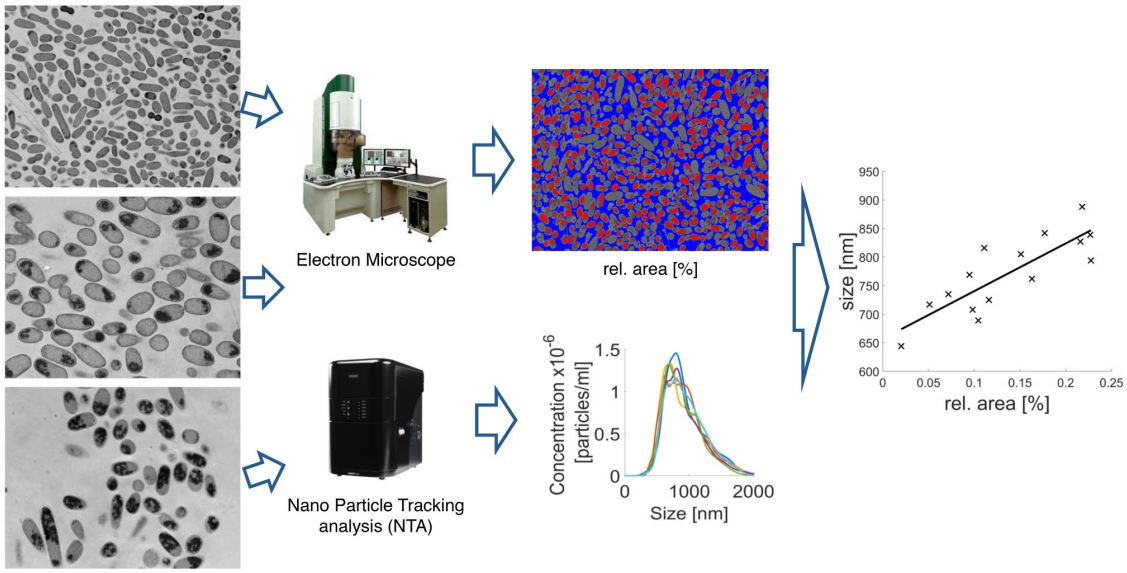

Figure 1. Method assessment by relative method verification. Flow chart of the targeted relative verification of NTA and TEM derived quantification of IB size. Independent IB process samples are analyzed by NTA as well as by TEM. TEM images on the left only represent the IBs as such, not the method specific sample preparation. While NTA measures IBs in the background of homogenized cells (Supporting information 2), the verification method TEM is based on ultrathin sections of whole cells in combination with grey scale image segmentation (Supporting information 1). NTA yields the hydrodynamic particle diameter distribution which corresponds to the IB size (nm) based on the utilized specific FL stain. TEM derived images of process samples are segmented according to the grey scale. Based on the image segmentation of IB area and cell area the relative area IB/cell [%] is calculated.

assessment can only be based on relative confirmation by comparing the results of two otherwise orthogonal methods (Fig. 1). Since an absolute measurement method for the particle size of a distribution of biologic nanoparticles is not available, a relative verification is targeted by comparing the relative area (%) derived from TEM and the hydrodynamic diameter (nm) derived from NTA. To avoid measuring artefacts it is of utmost importance to minimize the impact of sample preparation in order to preserve the most native IB form. While TEM is capable of visualizing IBs even in the cytosol, NTA can only measure particles in suspension. For this reason, HPF has been used as fixation approach for TEM, due to the gentle fixation properties. Supporting information 1 illustrates the conservation of cellular structures for different induction time points of two representative experiments. In contrast, sample preparation for NTA requires cell homogenization and a consequent FL stain in order to facilitate IB analysis in the background of cell debris.

3.2 Grey scale image segmentation for quantitative IB sizing is not significantly operator dependent

Based on deduced the images from Supporting information 1, qualitative IB growth over time can be observed. But for a quantitative assessment of IB size/growth over time, a standardized approach for IB sizing is necessary. Using grey scale image segmentation from TEM images. The relative IB size was quantified as IB area per cell (%). Basal grey values of TEM images, have been found to be highly variable owned to background particles. This impairs a uniform background correction and consequently fully automated image segmentation. Targeting a sound science method to reproducibly quantify cytosolic IB size, the operators analyzed 17 independent samples by image segmentation (Fig. 2) . Using a software aided approach, the time per image segmentation decreased below 10 s. For each sample three to six TEM images were recorded and segmented by the individual operator in random order at least three times. In Fig. 2A–C the variance induced by the different operators is indicated. The respective results are not statistically significant operator dependent, rendering the method transferable between operators for IB sizing and thereby suitable for method verification. Consequently, TEM imaging using HPF as gentle fixation method for IBs in combination with the software aided image segmentation approach across operators, was employed as orthogonal verification method.

In order to minimize artefacts, sample preparation was simplified as far as possible. While mere homogenization and direct NTA measurement did not lead to satisfying results, the implementation of a FL stain increased sensitivity of the method (data not shown). Also, standard FL beads were identified with high precision even in the

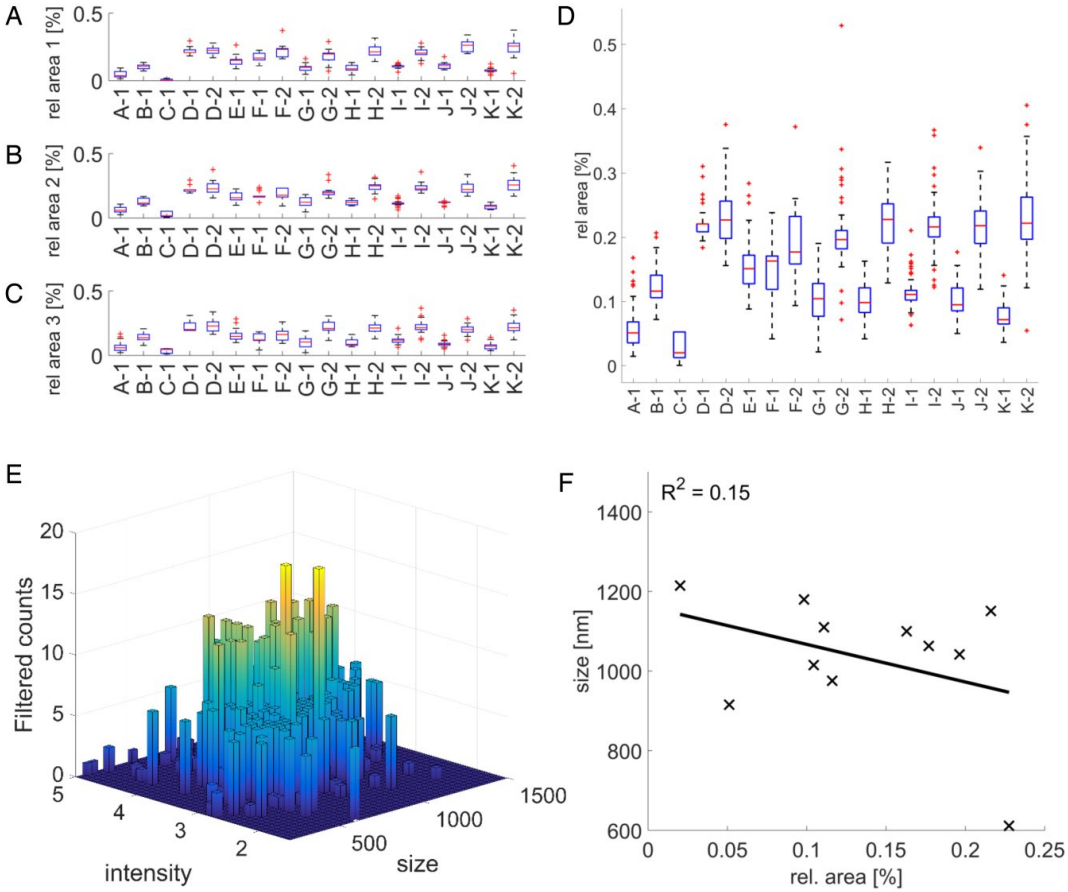

Figure 2. Grey scale image segmentation of TEM images for IB sizing is not significantly operator dependent. The relative area (rel. area [%]) corresponds to the area covered by IBs per cell background, 17 induced samples and one negative sample (C-1), of each sample three to six individual TEM images were segmented in random order ($n > 3$), each letter corresponds to an individual fermentation, samples with the same letter but different number correspond to different time points; (**A**–**C**) Grey scale image segmentation results operator specific, each subpanel corresponds to one individual operator; (**D**) all segmentations (> 550) results pooled, whiskers indicate 75% interval; (**E**) Filtered data of the size distribution of fixated IBs by NTA in the background of cell debris, all tracked particles of one sample measurement including the six replicate measurements, filtered by intensity and track length. (**F**) The correlation of relative IB area (%) to the hydrodynamic diameter derived from NTA (nm) was not found to be significant $p(t) = 0.18$, also the residuals were not found to be normally distributed.

background of stained homogenate (data not shown). Consequently, the samples were measured post homogenization and FL stain without any fixative (non-fixated). To illustrate the data basis for a size measurement by NTA, Fig. 2A displays a histogram of tracked and sized particles of an exemplary FL stained sample. Although a high number of particles was tracked in high number, NTA results and the relative IB areas from TEM-HPF were not significantly correlated (Fig. 2F). Judging from the TEM-HPF images (Supporting information 1) as well as from the image segmentation (Fig. 2D), a significant difference/growth over time in size of IBs can be observed. Nevertheless, this trend was not represented by NTA results.

3.3 Particle fixation for nano particle tracking analysis (NTA) increases method sensitivity

To investigate the impact of sample preparation on IB sizing by NTA, additional tests were conducted using a chemical fixative prior to cell homogenization. The IB sizing results of non-fixated IBs and chemical fixated IBs are compared in Fig. 3. Figure 3A and 3B illustrate the massive improvement of NTA raw data quality upon sample fixation prior to homogenization. The standard deviation is decreased and displays significantly less variance (Fig. 3C). Highly interesting is the observation that fixated samples display quantitatively more unspecific particles

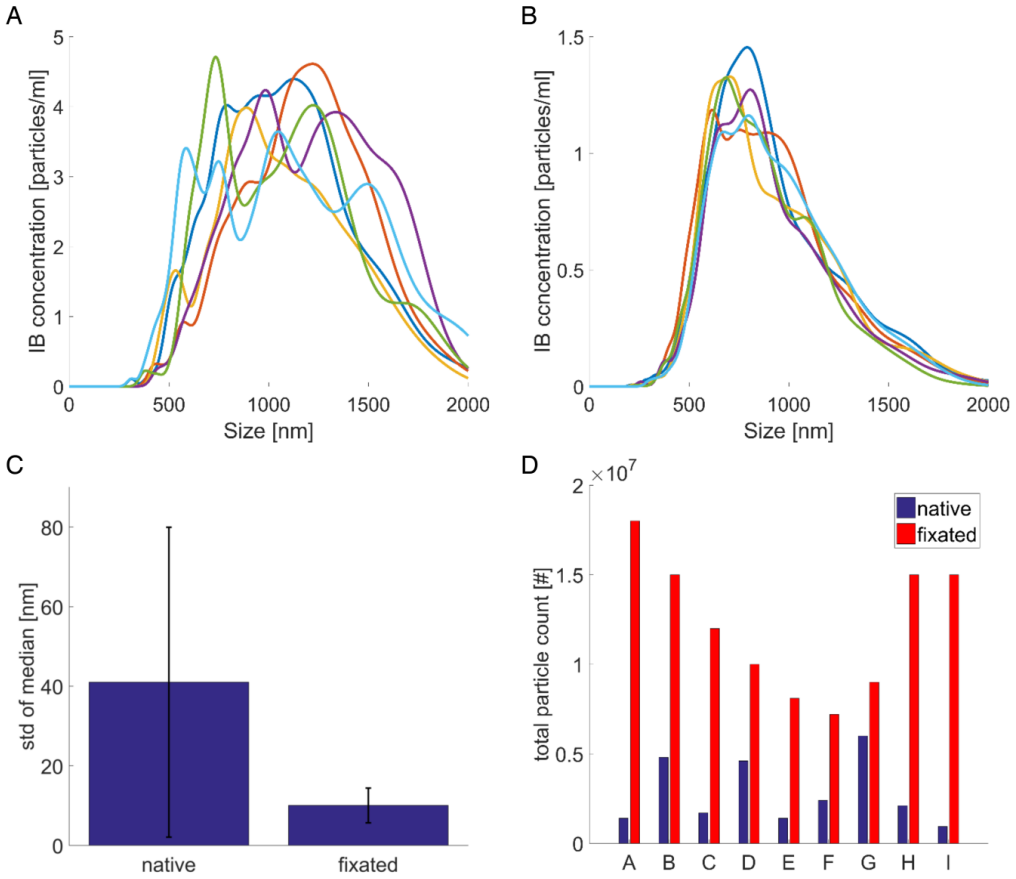

Figure 3. The positive impact of fixation prior to homogenization on NTA particle measurement. (**A**) raw data of particle size distribution of FL stained non-fixated IBs; (**B**) raw data of particle size distribution of FL stained fixated IB of the same sample as in (**A**); (**C**) Fixation prior to homogenization decreases standard deviation of the median size and boosts reproducibility, comparison of the standard deviation of the median of the particle size distribution of the six replicate measurements per sample, for not fixated (native) $n = 37$ and fixated (fixated) $n = 26$ samples, $p(t) < 0.001$; (**D**) total unspecific particle counts of NTA raw data, A-I correspond to sample names from various fermentations and time points, the observed trajectory in the particle count of fixated samples is presumably of coincidental nature.

than non-fixated samples (Fig. 3D). This is surprising in respect of the cross-linking properties of glutaraldehyde, which in theory should lead to generally bigger and fewer particles. Possibly, the addition of a fixative prior to homogenization prevents IBs from aggregating post homogenization during sample preparation.

Besides increasing measurement sensitivity, sample fixation appears to impact the intensity per particle (Fig. 4A). In order to put the IB size obtained from fixated samples into perspective, Fig. 4B illustrates the correlation of all measured TEM and NTA samples. Based on a p value of 0.002, it can be concluded that the rel. IB area (%) and the hydrodynamic diameter (nm) are correlated. To illustrate that the measured IB size is not a redundant measure of a simple titer quantification, the TEM areas as well as the IB diameters derived from NTA are compared to the respective specific titers (Fig. 4B and 4C). Based on the assumption of a uniform IB density within one sample the size of IBs should be tightly correlated to the amount of product contained in the particle. Nevertheless, the specific product titer does not display a highly significant correlation to the particle size, neither for particle sizes derived from TEM (Fig. 4C) nor from NTA (Fig. 4D).

The analysis of early and late time points of induction from different sets of experiments increases the observable differences in IB size. In comparison, the timely resolution of size over induction time is a greater challenge in regard to method sensitivity. In this respect, Fig. 5A illustrates the growth of IB size as well as the progression of specific product titer (g/g) over induction time. In accordance with Fig. 4D the IB size and product titer were not found to be

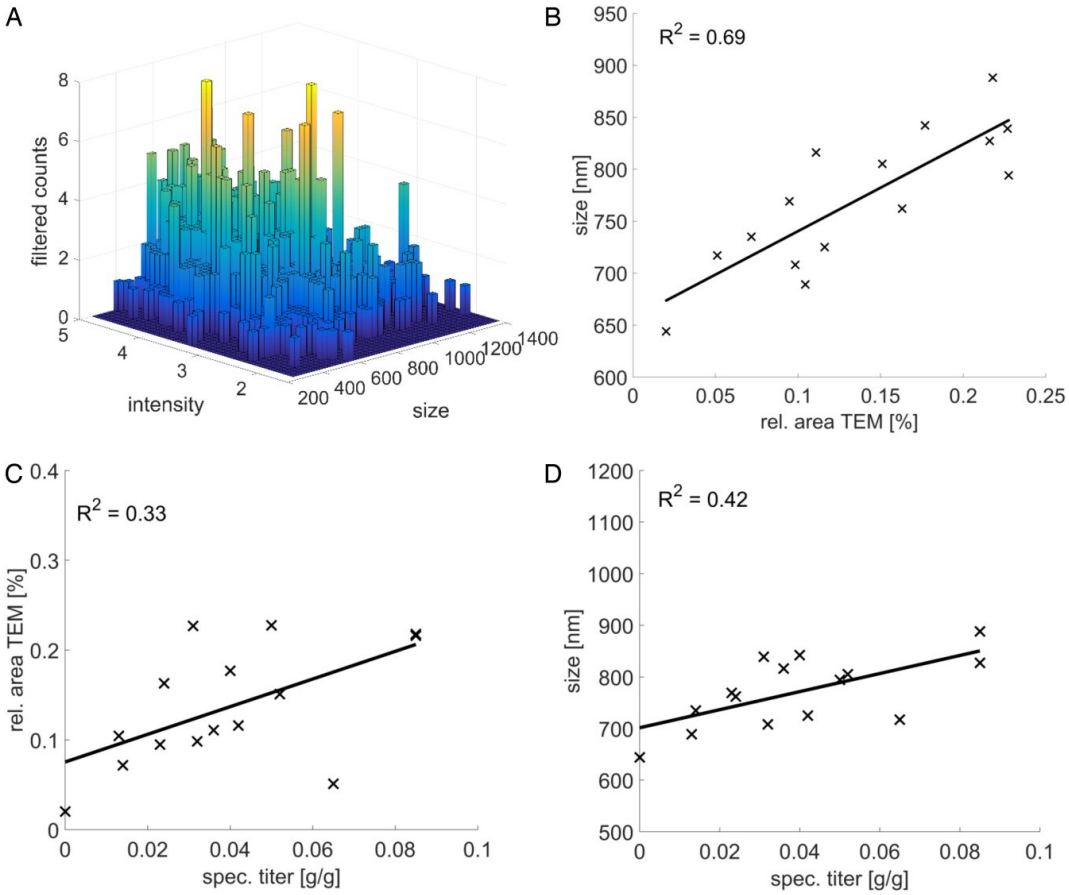

Figure 4. IB size of fixated samples derived from NTA is significantly correlated to the relative IB area derived from TEM; (**A**) Filtered data of the size distribution of fixated IBs by NTA in the background of cell debris, all particles of one sample measurement including the six replicate measurements, filtered by intensity and track length; (**B**) significant correlation of relative area (rel. area TEM [%]) and particle size derived from NTA (size), $n = 15$, $R^2 = 0.69$, $p(f) = 0.002$; (**C**) correlation of relative area (rel. area TEM [%]) and specific product titer (spec. titer [g/g]), $n = 15$, $R^2 = 0.33$, $p(f) = 0.026$, no serial correlation; (**D**) correlation of particle size derived from NTA (size) and specific product titer (spec. titer [g/g]), n = 15, $R^2 = 0.42$, $p(f) = 0.009$, no serial correlation.

closely correlated. The increase in size (+10 to 12%) was found to be comparably small given the substantial increase in specific titer (+300–400%) over induction time.

To investigate the impact of homogenization and correlated sample preparation on the sensitivity of the NTA measurement, additional samples after homogenization were analyzed (Fig. 5) by TEM. It can be observed that in case of HPF (Fig. 5C), the IBs are released into the supernatant and appear to maintain a more segregated state. The structure of these protein aggregates appeared porriferous and fragile. In contrast, the chemical fixation of the same sample prior to homogenization led to denser particles (Fig. 5B). Based on these images, it can be inferred that chemical fixation helps to maintain the IB conformation. A more pronounced impact of chemical fixation can be observed if the sample preparation for TEM is based on a thorough chemical fixation (Supporting information 3) instead of HPF/AFS (Supporting information 1). This conclusion is also in accordance with the previously discovered positive impact of fixation on NTA sensitivity.

4 Discussion

The investigation of the interface of USP and DSP requires sensitive analysis of the USP end product – the inclusion bodies. In this context, it are especially physical IB properties that presumably impact the isolation yield prior to solubilization and refolding during DSP. Accordingly, the goal of this contribution was to establish and verify a particle-based method to size a representative number of IBs with high sensitivity and high efficiency.

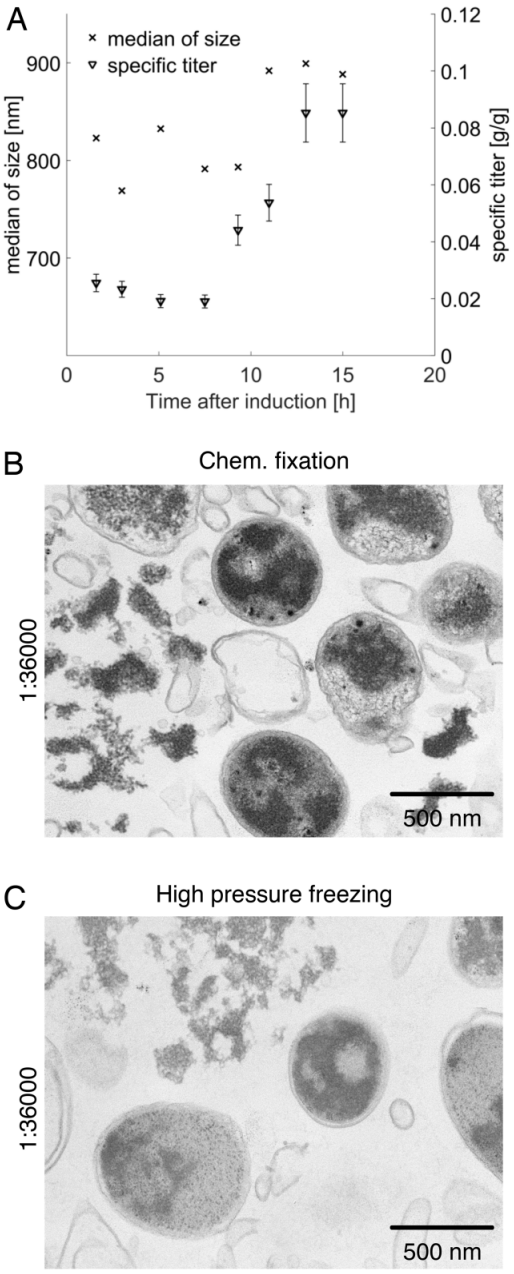

Figure 5. IB sizing by NTA features sufficient sensitivity to resolve IB growth over time; (**A**) The specific titer (spec. titer [g/g]) and IB size (median of size [nm]) over process time since induction (Time after induction [h]), as indicated before size and titer are correlated significantly, $n = 8$, $p(f) = 0.015$, including a serial correlation over time (**B**); Fixation leads to particle condensation in the homogenate as well as in the cells, TEM image of IBs after homogenization (700 bar, six passages) of chemically fixated cells, 1:36 000; (**C**) TEM image of IBs after homogenization (700 bar, six passages) of non-fixated cells, 1:36 000.

4.1 Grey scale segmentation of TEM images is a sensitive method for IB characterization

For method verification, a second, orthogonal method to assess IB size has been established. To minimize the effect of sample preparation and correlated artefacts, it was indicated to analyze the IBs in the most native conformation feasible. Using only centrifugation prior to HPF, sample preparation was reduced to a minimum. Besides sample preparation, the method of sample fixation has been a topic of vivid discussion. A common approach for sample fixation of IBs is chemical fixation [11, 13, 20, 24] or air drying of the specimen on a copper grid [15, 16]. Despite the wide usage of these methods, the specimen might be altered, subcellular structures can condense and shrinkage can occur. In this contribution, that used HPF-AFS, the structure of IBs was found to be far looser and more sensitive in contrast to the general opinion regarding the shape of IBs [13].

Regardless of the sample preparation, microscopy generally simplifies the shape of 3D specimens to 2D images. Although 3D-TEM offers an alternative, the effort per measurement renders 3D-TEM unfeasible for the analysis of over 100 particles per sample. To compensate for the drawback of a 2D image based method a representative number of particles needs to be analyzed. Especially since the IBs do not appear to have a fully symmetrical, spherical shape, different orientations need to be accounted for by sizing a larger number of IBs. Peternel et al [24] addressed the problem of the statistical significance and sized 250–350 IBs in order to obtain a histogram of IB size distribution of isolated and washed IBs. In accordance with this contribution, 160–380 individual cells (containing IBs) were analyzed repetitively by each of the three operators for every sample. But in contrast to the chemical fixation used by Peternel et al [24], HPF-AFS was used as a highly gentle method, renowned for its ability to preserve cellular substructures. Hereby, we established HPF-AFS TEM imaging of IBs in the cytosol in combination with grey scale image segmentation as valuable method to reproducibly quantify a representative number of IBs, independent of an operator. Nevertheless, the five days needed for sample preparation in combination with the undeniably time-consuming procedure of image segmentation, do not qualify the method as simplistic or easily transferable in a technological sense.

4.2 NTA is a sensitive method to size a representative number of IBs

IB sizing by NTA permitted the characterization of a representative number of particles (>1000) per sample. In contrast to HPF-AFS TEM, sample preparation for the NTA measurement as well as the actual measurement of several samples was done within one day. Data evaluation was fully automated and can be easily standardized,

which makes it less operator dependent. Consequently, NTA appears to offer a far more feasible approach to quantify IB properties on a single particle level, compared to currently available methods.

In the context of NTA measurements the chemical fixation of the cells prior to homogenization yielded an increase in sensitivity as well as a substantial increase in total, unspecific particle count. Although the total amount of recognized particles had increased (Fig. 3D), the amount of relevant particles (IBs) had decreased (Fig. 3A and 3B). This observation might be due to a certain aggregation tendency of native particles. This tendency could lead to an aggregation of the IBs with cell debris, especially in combination with a disintegration of the IBs triggered by shear stress during homogenization. Cellular debris after high-pressure homogenization is about 0.5 µm in size [17]. In combination with IB fragments, the resulting size would theoretically overlap with the expected size of native IBs. In case fixation impairs this aggregation and decreases the probability of IB disintegration during homogenization, fixation would lead to an increase in the total number of particles but a decreased number of product specific particles.

Despite using a product specific FL stain in combination with a chemical fixation of the cells, background particles were found to create a bias in the sizing of the standard beads present in the background of homogenate. The observed strong background signal might be attributed to a bleed through of scattered light through the long pass fluorescence filter, which decreases the method specificity. For future measurements, it would be advisable to circumvent such issues by increasing the distance between the excitation wavelength and fluorescence filter.

Besides the methodological advances for sizing IBs by using NTA and TEM, the results indicate that high-pressure homogenization greatly impacts IB properties. Consequently, the native size of the IBs is highly unlikely to be preserved throughout high pressure homogenization. Nevertheless, a tendency to adhere to surfaces (stickiness) during the isolation step, e.g. the wall of the continuous centrifuge, could cause significant product loss. This hypothesis substantiates the necessity of a sensitive characterization of IB particle properties, in order to enable the scientific community to investigate the correlated product loss.

The NTA method offers an approach to quantify the impact of USP and homogenization on the IB particle property size. Moreover, NTA allows to assess the ratio of particle count of non-fixated and fixated IB samples, which could potentially be used as a measure of stickiness.

In summary, the findings infer that even if different process parameters in USP elicit differences in IB particle properties, it is highly unlikely that these differences are preserved throughout high-pressure homogenization. However, IB sizing by NTA could help to better understand the molecular processes which lead to different aggregation tendencies and in turn, impact isolation efficiency. Consequently, NTA could be used to derive an additional response parameter on the basis of which integrated bioprocess development might succeed in investigating the interlink of USP and DSP.

4.3 Conclusions

The overall goal of this contribution was the establishment and assessment of a simplistic and sensitive method for high throughput IB sizing.

(i) TEM in combination with grey scale image segmentation is a sensitive and reproducible method to quantify the size of native, cytosolic IBs and can be used for method verification.
(ii) NTA is a particle-based method that allows to size a great number (>1000) of fluorescence labelled IBs in the background of cell debris.
(iii) Chemical fixation of IBs prior to homogenization leads to a decrease in standard deviation and particle count but increases the reproducibility of IB sizing with NTA.

Based on the observed effect of fixation, it can be hypothesized that high-pressure homogenization annihilates differences in IB size caused by USP. Nevertheless, the ratio in particle count of native homogenate and fixated homogenate offers a measure for IB stickiness.

We are grateful for the financial support of Sandoz GmbH. The TEM measurements were performed by the EM Facility of the Vienna Biocenter Core Facilities GmbH (VBCF), a member of Vienna Biocenter (VBC), Austria. We are grateful for the access to the NTA measuring device facilitated by Prof. Manfred Ogris and Haider Sami, from the Division of Clinical Pharmacy and Diagnostics, University of Vienna, Vienna, Austria.

The authors declare no conflict of interest.

5 References

[1] Hedhammar, M., Alm, T., Gräslund, T., Hober, S., Single-step recovery and solid-phase refolding of inclusion body proteins using a polycationic purification tag. *Biotechnol. J.* 2006, *1*, 187–196.
[2] Wong, H. H., O'Neill, B. K., Middelberg, A. P., Centrifugal processing of cell debris and inclusion bodies from recombinant *Escherichia coli*. *Bioseparation* 1996, *6*, 361–372.
[3] Singh, S. M., Panda, A. K., Solubilization and refolding of bacterial inclusion body proteins. *J. Biosci. Bioeng.* 2005, *99*, 303–310.
[4] Chen, Y., Wang, Q., Zhang, C., Li, X. et al., Improving the refolding efficiency for proinsulin aspart inclusion body with optimized buffer compositions. *Protein Expression Purif.* 2016, *122*, 1–7.
[5] Burgess, R. R., Refolding solubilized inclusion body proteins. *Methods Enzymol.* 2009, *463*, 259–282.

[6] Dürauer, A., Mayer, S., Sprinzl, W., Jungbauer, A., Hahn, R., High-throughput system for determining dissolution kinetics of inclusion bodies. *Biotechnol. J.* 2009, *4*, 722–729.
[7] Grote, F., Ditz, R., Strube, J., Downstream of downstream processing – Integrated bioprocess development from upstream to downstream. *Chem. Ing. Tech.* 2009, *81*, 1276–1277.
[8] Freydell, E. J., Ottens, M., Eppink, M., van Dedem, G., van der Wielen, L., Efficient solubilization of inclusion bodies. *Biotechnol. J.* 2007, *2*, 678–684.
[9] Wellhoefer, M., Sprinzl, W., Hahn, R., Jungbauer, A., Continuous processing of recombinant proteins: Integration of inclusion body solubilization and refolding using simulated moving bed size exclusion chromatography with buffer recycling. *J. Chromatogr. A* 2013, *1319*, 107–117.
[10] Strandberg, L., Enfors, S. O., Factors influencing inclusion body formation in the production of a fused protein in *Escherichia coli*. *Appl. Environ. Microbiol.* 1991, *57*, 1669–1674.
[11] Datta, I., Gautam, S., Gupta, M. N., Microwave assisted solubilization of inclusion bodies. *Sustainable Chem. Proc.* 2013, *1*, 1–7.
[12] Datar, R. V., Cartwright, T., Rosen, C.-G., Process economics of animal cell and bacterial fermentations: A case study analysis of tissue plasminogen activator. *Biotechnology* 1993, *11*, 349–357.
[13] Castellanos-Mendoza, A., Castro-Acosta, R. M., Olvera, A., Zavala, G. et al., Influence of pH control in the formation of inclusion bodies during production of recombinant sphingomyelinase-D in *Escherichia coli*. *Microb. Cell Fact.* 2014, *13*, 1–14.
[14] Luo, J., Leeman, M., Ballagi, A., Elfwing, A. et al., Size characterization of green fluorescent protein inclusion bodies in *E. coli* using asymmetrical flow field-flow fractionation–multi-angle light scattering. *J. Chromatogr. A* 2006, *1120*, 158–164.
[15] Upadhyay, A. K., Murmu, A., Singh, A., Panda, A. K., Kinetics of inclusion body formation and its correlation with the characteristics of protein aggregates in *Escherichia coli*. *PLoS One* 2012, *7*, e33951.
[16] Margreiter, G., Messner, P., Caldwell, K. D., Bayer, K., Size characterization of inclusion bodies by sedimentation field-flow fractionation. *J. Biotechnol.* 2008, *138*, 67–73.
[17] Wong, H. H., O'Neill, B. K., Middelberg, A. P. J., Cumulative sedimentation analysis of *Escherichia coli* debris size. *Biotechnol. Bioeng.* 1997, *55*, 556–564.
[18] Agerkvist, I., Enfors, S.-O., Characterization of *E. coli* cell disintegrates from a bead mill and high pressure homogenizers. *Biotechnol. Bioeng.* 1990, *36*, 1083–1089.
[19] Taylor, G., Hoare, M., Gray, D. R., Marston, F. A. O., Size and density of protein inclusion bodies. *Nat. Biotechnol.* 1986, *4*, 553–557.
[20] Walther, C., Mayer, S., Sekot, G., Antos, D. et al., Mechanism and model for solubilization of inclusion bodies. *Chem. Eng. Sci.* 2013, *101*, 631–641.
[21] Elena, G.-F., Joaquín, S.-F., Esther, V., Antonio, V., Tunable geometry of bacterial inclusion bodies as substrate materials for tissue engineering. *Nanotechnology* 2010, *21*, 205101.
[22] Cano-Garrido, O., Rodríguez-Carmona, E., Díez-Gil, C., Vázquez, E. et al., Supramolecular organization of protein-releasing functional amyloids solved in bacterial inclusion bodies. *Acta Biomater.* 2013, *9*, 6134–6142.
[23] Moon, M. H., Lee, S., Sedimentation/steric field-flow fractionation: A powerful technique for obtaining particle size distribution. *J. Microcolumn Sep.* 1997, *9*, 565–570
[24] Peternel, Š., Jevševar, S., Bele, M., Gaberc-Porekar, V., Menart, V., New properties of inclusion bodies with implications for biotechnology. *Biotechnol. Appl. Biochem.* 2008, *49*, 239–246.
[25] Filipe, V., Hawe, A., Jiskoot, W., Critical evaluation of Nanoparticle Tracking Analysis (NTA) by NanoSight for the measurement of nanoparticles and protein aggregates. *Pharm. Res.* 2010, *27*, 796–810.
[26] Korz, D. J., Rinas, U., Hellmuth, K., Sanders, E. A., Deckwer, W. D., Simple fed-batch technique for high cell density cultivation of *Escherichia coli*. *J. Biotechnol.* 1995, *39*, 59–65.
[27] Lee, S. Y., High cell-density culture of *Escherichia coli*. *Trends Biotechnol.* 1996, *14*, 98–105.

3.2 PHYSIOLOGICAL PROCESS CONTROL - q_S CONTROL

3.2.1 OPEN LOOP CONTROL OF q_S - OFFLINE VALIDATION OF SOFTSENSOR K2S1

The first principal softsensor described in 2.6.1 works precisely under the condition of no substrate accumulation, accurate feed pump calibration and correct initial biomass. Over- or underestimation of initial biomass, wrong calibration of the feed pump or substrate accumulation lead to inadequate feed rates and consequently to mistakenly process values of q_S.

Data analysis of previous work (experiments WR27 and WR35) showed some examples of failing of this setup as shown in Figure 11. In case of WR27-A an underestimation of the initial biomass or inaccurate feed pump calibration led to an in real-time unknown deviation of q_SPV from q_SSP. In WR27-D the assumption that r_S and therefore q_S can be calculated directly from the feed rate is not valid due to severe overestimation of biomass through substrate accumulation. Using the consistency check, by monitoring the h-value and making control actions on it, it could be possible to detect such situations. However, this was not implemented in the used script and was not used in real-time.

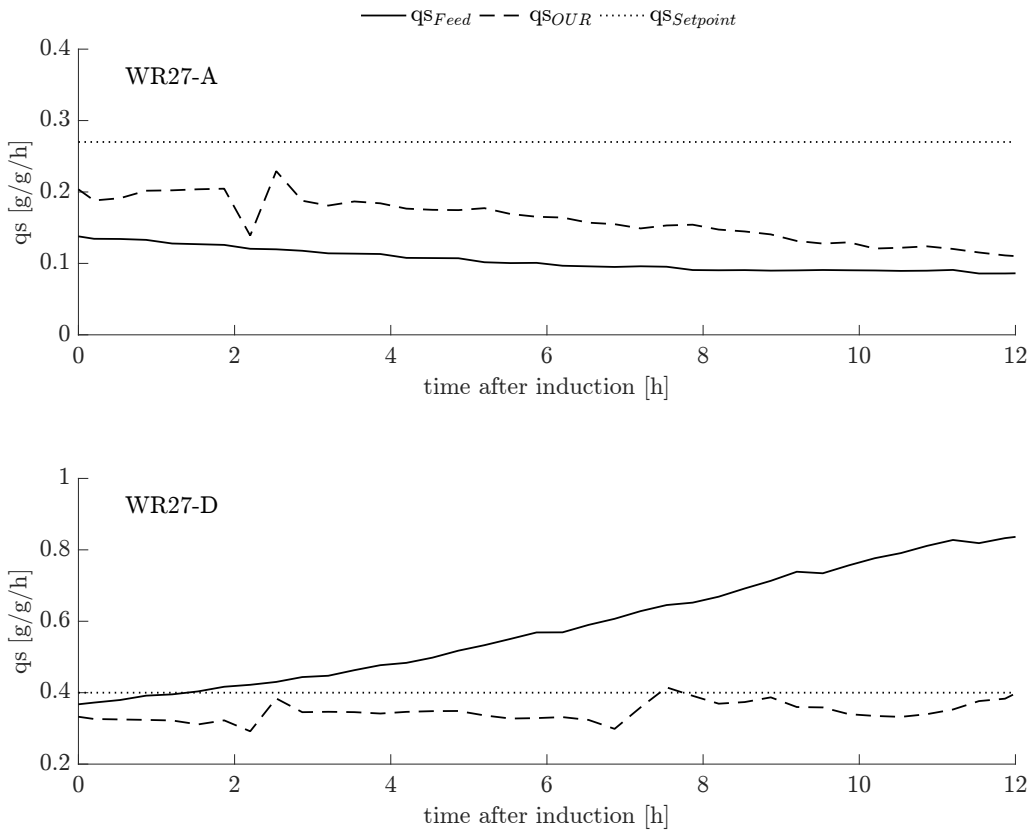

Figure 11: Specific substrate uptake rate calculated from feed rate and OUR of WR27; The upper plot (WR27-A) shows continuous underestimation of q_{SFeed} (solid line) compared to q_{SOUR} (dashed line) because of to low biomass estimation respectively inaccurate feed pump calibration. The lower plot (WR27-D) shows escalating of q_{SFeed} due to substrate accumulation. Both plots show deviations from the setpoint (dotted line)

Using the h-value as a consistency check certainly would have some difficulties too, if applied in real-time. For WR27-A the h-value spikes about two hours after induction because of mixing pure oxygen to the in-gas to control pO_2 above 30 % as shown further down for WR35 (shown in Figure 14). In this case, since no substrate accumulation occurred, the h-value could be misinterpreted and would cause wrongful feed rate adjustment. And even though the biomass underestimated throughout the experiment and q_SPV is far of its setpoint, the h-value stays low all along. In such a case the h-value would have not been of any use because the deviation of q_SPV couldn't be detected in real-time. The data of WR27-D show steady increase of the h-value in case of severe overestimation of biomass (shown in Figure 12) through substrate accumulation.

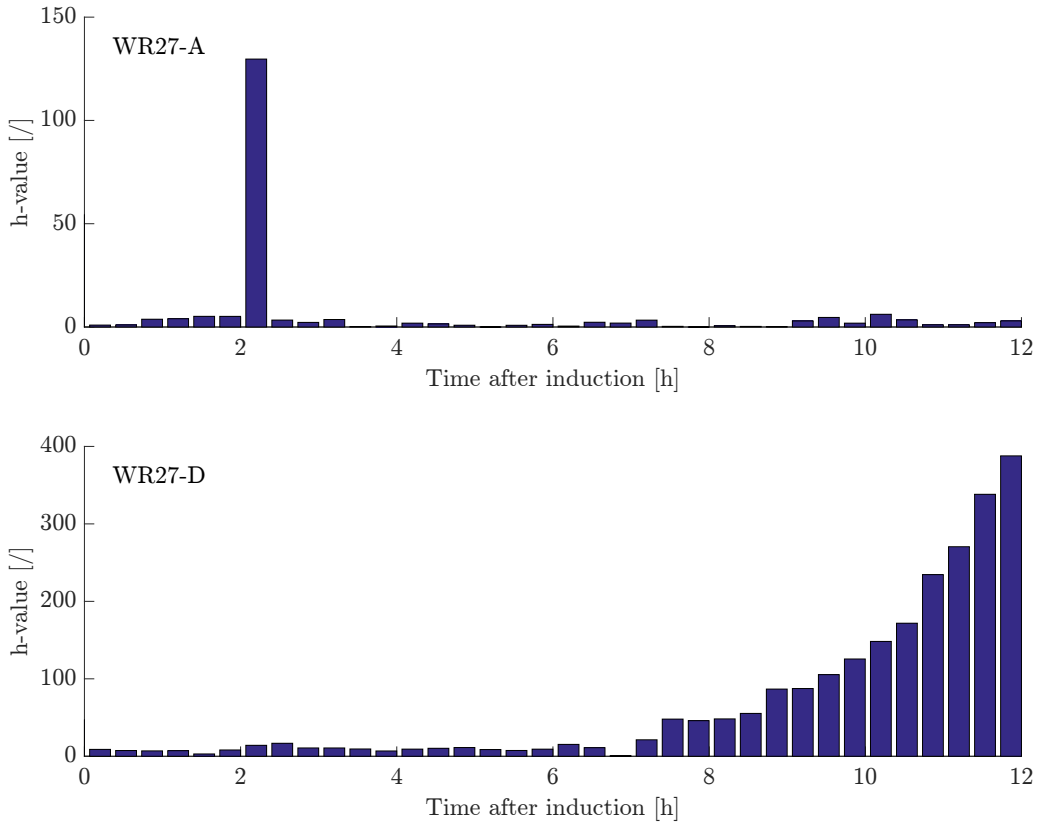

Figure 12: h-value of WR27-A (upper plot) and QR27-D (lower plot); The upper plot shows a spike of the h-value due to mistakenly calculated OUR and the constant increase of the h-value during growing substrate accumulation.

As WR35 was evaluated the RQ showed spikes, which could not be explained physiologically. Data analysis and investigation of the calculation method showed that the spikes originate in the changes of oxygen concentration of the gassing (shown in Figure 13). It is clearly seen that with every increase of the oxygen concentration, the RQ shows a steep negative slope and vice versa. In Equation 14 the calculation of RQ is stated.

$$RQ = \frac{CER}{OUR}$$

Equation 14: Calculation of the respiratory quotient RQ with carbon evolution rate CER and oxygen uptake rate OUR as inputs

Investigating the two inputs, CER and OUR, led to the finding that spikes in the OUR signal causes the spikes of the RQ. Equation 15 shows the calculation of OUR and confirms the assumption that changes of the oxygen concentration cause the spikes in the OUR respectively the RQ signal.

$$OUR = \frac{\dot{V}}{V_m} \cdot \left(y_{O2,in} - y_{O2,out}\right)$$

Equation 15: Calculation of oxygen uptake rate OUR in mol/h

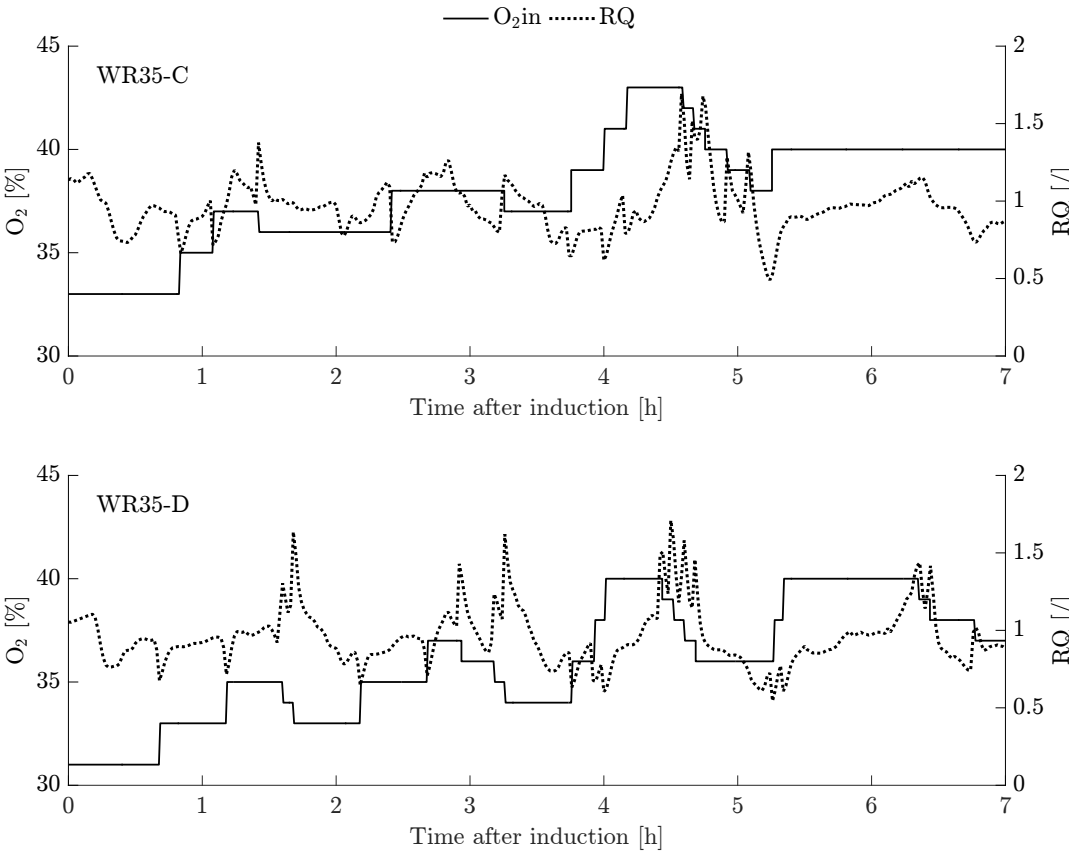

Figure 13: Spikes in RQ (dotted line) occur at changes of oxygen concentration of gassing (solid line) shown in WR35-C (upper plot) and WR35-D (lower plot)

Further data analysis of WR35C and WR35-D showed a direct correlation between the changes of oxygen concentration respectively the OUR spikes and the h-value (shown in Figure 14). The h-value increases just because of the mistakenly spikes of OUR and would have indicated substrate accumulation accidently.

The increase of the h-value around 0.4 h after induction results from the sudden increase of the feed rate (oscillations) and indicates a severe change of the physiological state of the culture.

A small experiment was conducted to determine the delay of off-gas measurement. Therefore, a reactor was filled with 1.5 l of water and aeration was done with 1.4 vvm in accordance to the setup of the fermentations. The goal was to measure the time delay between increasing the oxygen concentration at the inlet and detecting it at the off-gas analyzer. The results are shown in Figure 15.

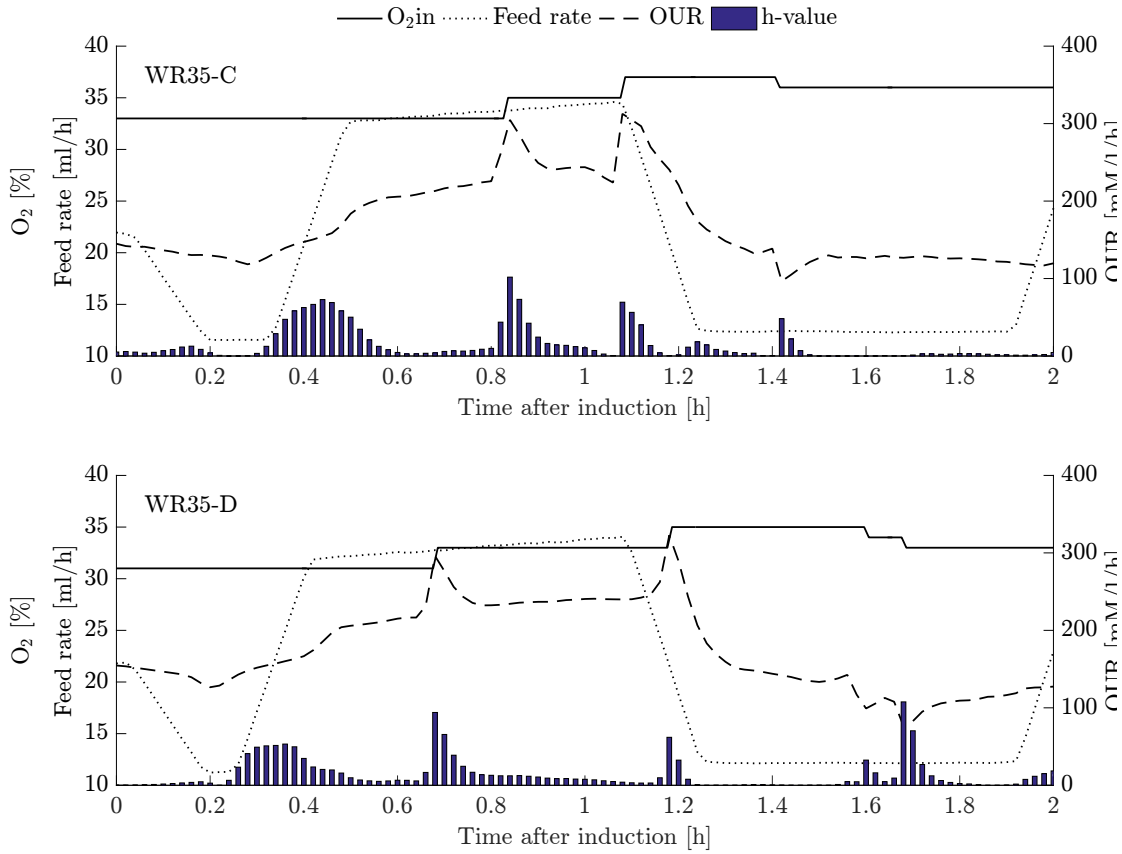

Figure 14: Correlation of OUR (dashed line) and h-value (bars) with changes in oxygen concentration of gassing (solid line) shown in WR35-C (upper plot) and WR35-D (lower plot). The feed rate (dotted line) is shown to understand the trend of OUR.

It takes approximately two minutes to detect a change in the off-gas signal and approximately six minutes to measure the same concentration at the outlet as at the inlet. The commonly used visual basic script was executed every five minutes. With each execution of the script it was possible to change the setpoints of the parameters and calculate biomass growth etc. using the softsensor. This means that the effects of any changes in the gassing or the feed rate five minutes ago could not be detected correctly.

This was the reason to increase the intervals for executing the script from five to twenty minutes. Especially when thinking about adjusting the feed rate to reach a setpoint, one has to make sure that the next measurement respectively estimation of the process value reflects the most recent changes to influence it.

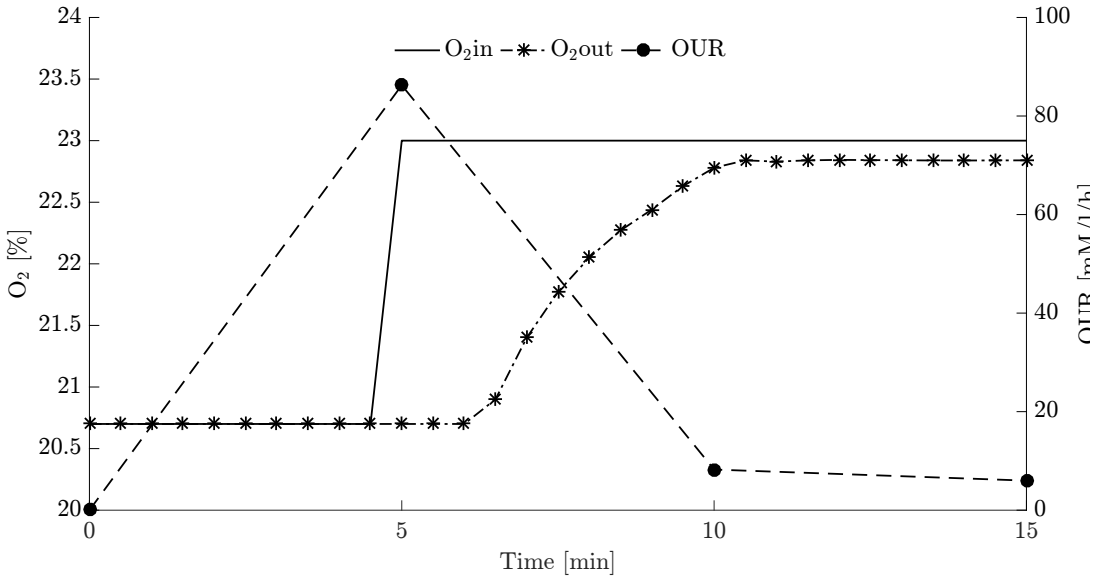

Figure 15: Oxygen concentration of inlet (solid line) and outlet (dot-dashed line) of reactor and calculated OUR (dashed line)

3.2.2 REAL-TIME ESTIMATION OF Q_S - VALIDATION OF SOFTSENSOR K2S1_V2 ON BASIS OF WR44

The experiment WR44 was conducted to test the new version of the softsensor K2S1_v2 as described in 2.6.2. Biomass estimation in real-time was performed by the new softsensor K2S1_v2 however no feed rate adjustment nor setpoint adaption was performed in real-time.

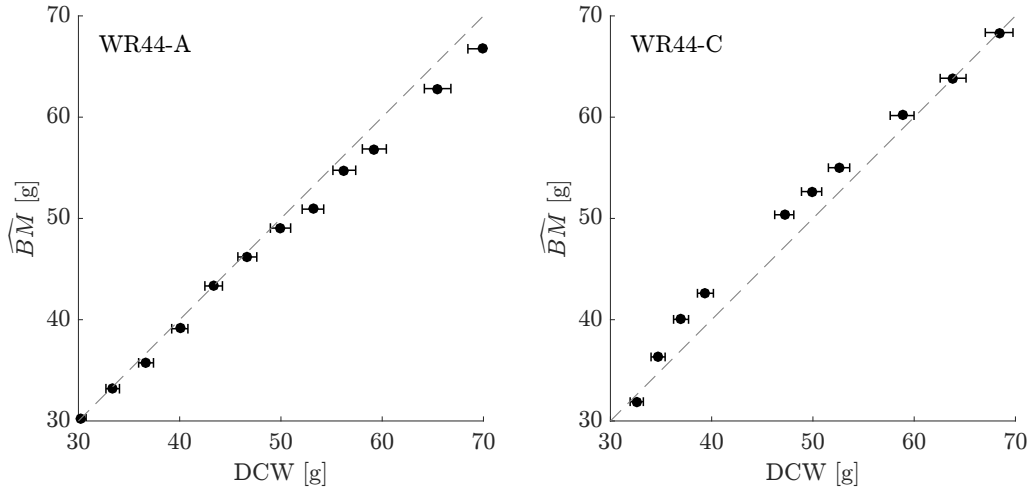

Figure 16: Real-time biomass estimation of WR44-A (upper plot) and WR44-C (lower plot) compared to DCW; horizontal error bars indicate accuracy of DCW measurement of approximately +/- 1.5 %

The biomass estimation is consistent with the DCW measurements from 30-70 g (shown in Figure 16). Therefor it was suitable to apply for real-time estimations and provides a solid basis for estimation of biomass growth rate to feed the softsensor.

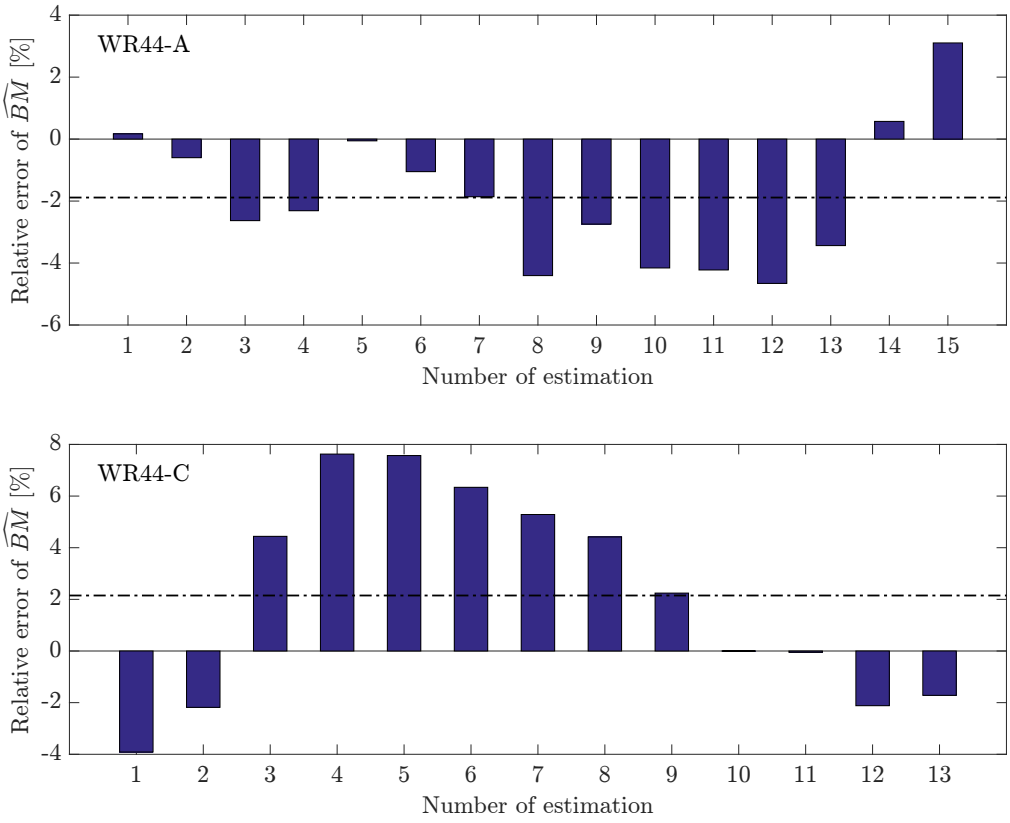

Figure 17: Relative error (bar) and average of relative error (dash-dotted line) on real-time biomass estimation from WR44-A (upper plot) and WR44-C (lower plot).

With -1.9 % and 2.1 % the average relative error of biomass estimation is close to the DCW measurement accuracy (shown in Figure 17), which is approximately 1.5 %. The results of this weighted average approach correspond with data shown by Reichelt et al. (Reichelt, Thurrold et al. 2016). The maximum relative error of estimation amounts to -4.7 % in case of WR44-A respectively 7.6 % in case of WR44-C.

When calculating the biomass growth rate r_x from DCW, error propagation has to be taken into account. The possible error of the growth rate was calculated accordingly to Equation 16, Equation 17 and Equation 18. The magnitude of the possible error of r_x from DCW is shown in Figure 18. In most cases the real-time estimation $\widehat{r_x}$ lies well within the range of r_x and more importantly follows the trend of r_x correctly.

$$e_{rx} = r_{x,max} - r_x = \frac{\Delta X_{max}}{\Delta t} - \frac{\Delta X}{\Delta t}$$

Equation 16: Error biomass growth rate due to limited accuracy of DCW measurements

$$\Delta X_{max} = X_i \cdot \left(1 + e_{x,i}\right) - X_{i-1} \cdot \left(1 - e_{x,i-1}\right)$$

Equation 17: Possible maximum change of DCW based on error-prone DCW measurement

$$\Delta X = X_i - X_{i-1}$$

Equation 18: Change of DCW based on correct measurements

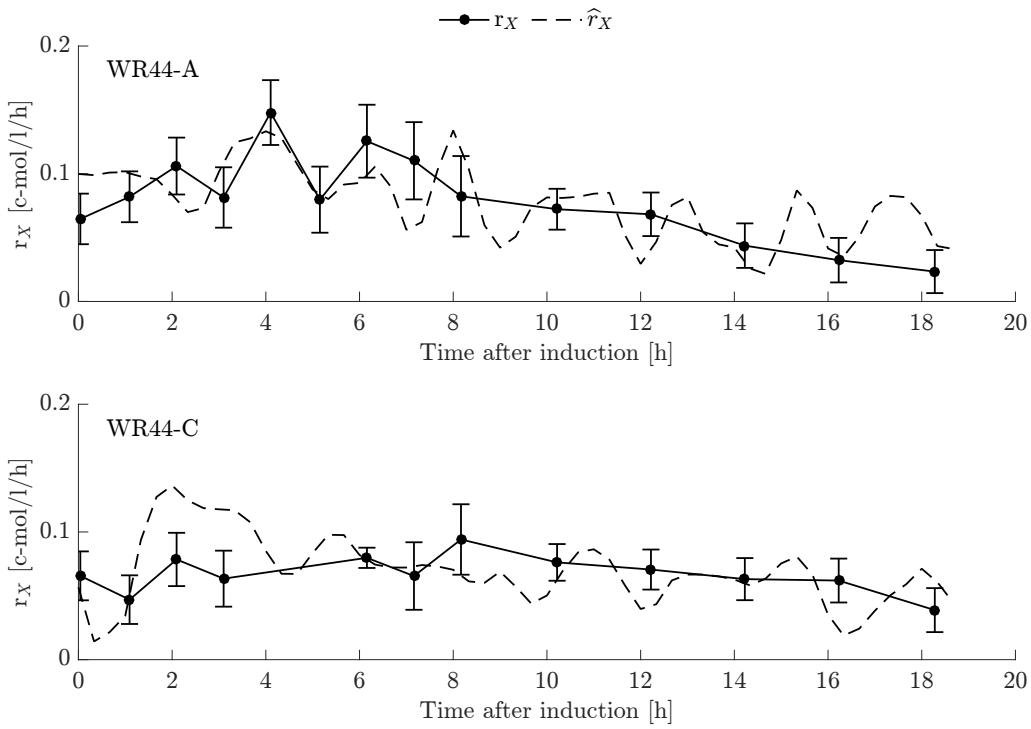

Figure 18: Real-time estimation of biomass growth rate (dashed line) from WR44-A (upper plot) and WR44-C (lower plot) compared to r_x from DCW (solid line)

Beginning with r_x, OUR and CER the softsensor K2S1_v2 calculated the reconciled substrate conversion rate r_s and furthermore the specific substrate uptake rate q_s in real-time every 20 minutes. The results of the real-time estimation are shown in Figure 19. It can be seen that the real-time estimation of q_s clearly follows the trend of actually achieved q_s correctly. The mean error of 17.7 % for WR44-A and 18.4 % for WR44-C with a standard deviation of 21.8 % and respectively 22.0 % originates in the previously mentioned errors of biomass estimation, biomass growth rate estimation and off-gas measurements. Considering the novel approach of real-time estimation these values were adequate to continue with an experiment for closed loop control of q_s.

The consistency check respectively the h-value is plotted in Figure 20. After a transient condition of approximately two hours the h-value is constantly single-digit, what points to correct input values of the softsensor.

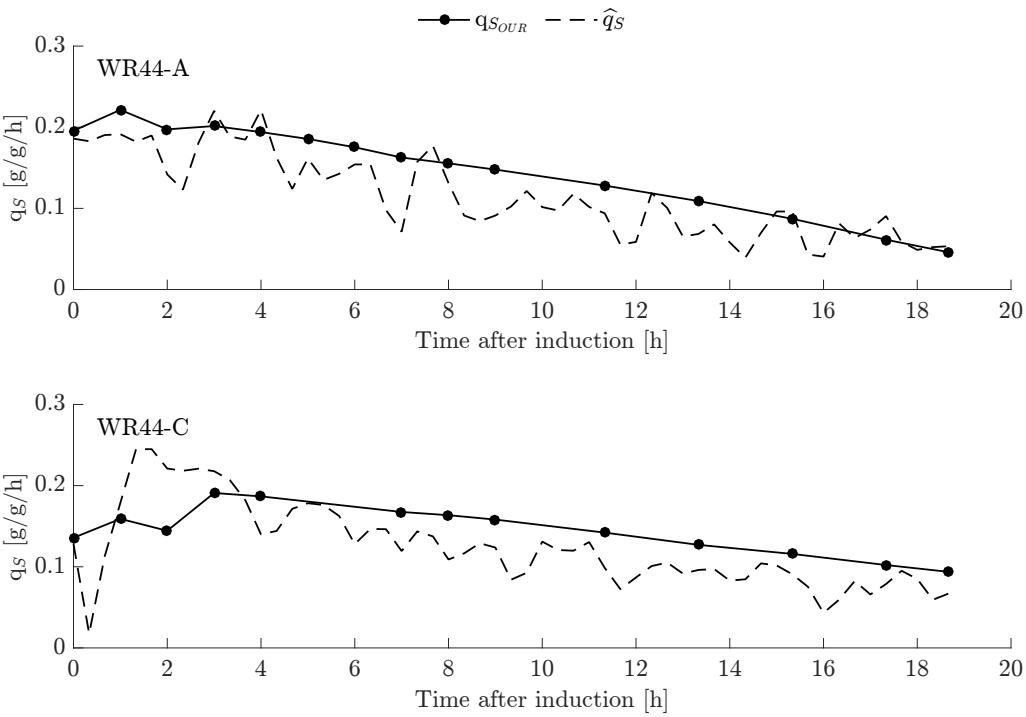

Figure 19: Real-time estimation of q_s (dashed line) compared to actually achieved specific substrate uptake rate q_{sour} (solid line)

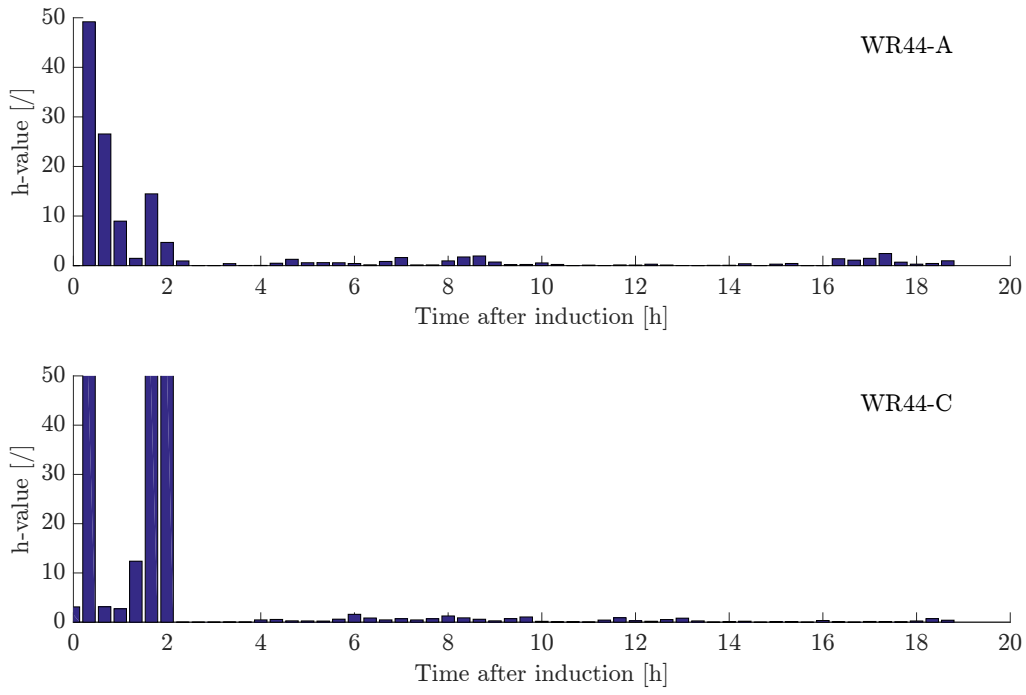

Figure 20: h-value of WR44-A (upper plot) and WR44-C (lower plot); The figure shows spikes of the h-value at the beginning of induction phase due to transient condition

3.2.3 CLOSED LOOP CONTROL OF q_S – PUBLICATION PART

Based on the results described in 3.2.2 experiments for closed loop control of q_s with and without setpoint adaption were performed. Results of these experiments are shown in the current draft of the publication, which covers an aspect of this thesis.

The contributions of the work to the following publication accompanied with the writing of this thesis lie in the development of methods, execution and data analysis of the fermentation experiments, writing of materials and methods, preparation of all figures and figure captions and the preparation and description of the equations as well as abbreviations.

Physiological feedback control to avoid substrate accumulation in *E. coli*

Wieland N. Reichelt[a], Andreas Kaineder[a], Jens Fricke[a], Christoph Herwig[a,b*]

[a] Christian Doppler Laboratory for Mechanistic and Physiological Methods for Improved Bioprocesses, Vienna University of Technology, Getreidemarkt 9/166, A-1060 Vienna, Austria

[b] Research Division Biochemical Engineering, Institute of Environmental and Bioscience Engineering, Vienna University of Technology, Gumpendorfer Strasse 1A/166-4, 1060 Vienna, Austria;

Abstract

Producing pharmaceutically relevant proteins in microbial bioprocesses, substrate accumulation has to be avoided for the sake of productivity and controllability. But during late induction phase unexpected substrate accumulation is a phenomena often observed in microbial bioprocesses, despite comprehensive strain characterization and quantification of the physiological capacity to metabolize substrate without accumulation of substrate or metabolites (q_{Scrit}).

Recent literature has illustrated a clear dependency of q_{Scrit} not only on induction time but also on the level of metabolic activity. In other words in addition to time after induction the rate of substrate metabolization severely impacts the decline in q_{Scrit}. To effectively avoid substrate accumulation, the dynamics and dependencies of the q_{Scrit} highlight the necessity to sense this physiological capacity real time. In this contribution a combination of mass balances was used to estimate the process value of q_S independently of substrate accumulation as well as biomass yield and accurately real time (q_SPV). In the context of physiological control of the specific substrate uptake rate (q_S) the latter approach allowed for the first time physiological feedback control by the comparison of q_SPV and the setpoint of q_S (q_SSP). Moreover, using a simple algorithm reaching of q_{Scrit} was detected real time in order to react upon such a breach by q_SSP adaptation. By successfully avoiding substrate and metabolite accumulation throughout induction phase of an industrial relevant production process, we were able to illustrate the feasibility of the physiological feedback control.

Keywords

Physiological feedback control; critical physiological capacity; substrate accumulation; real time biomass estimation; oxidative metabolism

Abbreviation

$\widehat{BM}_{mean}$	estimated biomass based on the N-balance, the DoR balance and a permittivity measurement (g)
$\widehat{BM}_{C-Bal}$	estimated biomass based on the C-balance (g)

r_X	biomass growth rate calculated offline (g/h)
$\widehat{r_X}$	real time estimated biomass growth rate (g/h)
r_S	substrate conversion rate calculated offline (g/h)
$\widehat{r_S}$	real time estimated substrate conversion rate (g/h)
q_S	the specific substrate uptake calculated offline (g/g/h)
q_SSP	setpoint of q_S (g/g/h)
q_SPV	real time process value of q_S (g/g/h)
Δq_SPV_i	change of q_SPV within the last control interval (20 min) (g/g/h)
c_S	substrate concentration in feed (g/L)
F_0	starting feed rate (L/h)
r_{acc}	rate of accumulating substrate and acetate (c-mol/h)
r_{CO2}	CER, carbon dioxide evolution rate (mol/h)
r_{O2}	OUR, oxygen uptake rate (mol/h)
r_X	biomass conversion rate (mol/h)
μ	specific biomass growth rate (1/h)
V_0	volume at t = 0 (L)
X_i	biomass dry cell weight at t = 0 batch end or at time point t = I (g)
$Y_{X/S}$	biomass yield on substrate (g/g or c-mol/c-mol)

Introduction

Bioprocesses are increasingly employed for the production of pharmaceuticals owned to correlated cost efficacy. Given the simple genetic accessibility, high growth rates and low demands concerning media composition, *Escherichia coli* is one of the most exploited hosts for industrial production of recombinant proteins [1,2]. To optimize productivity of a given bioprocess, process parameters and their interrelations are investigated during process development within a given range – the characterization space [3]. One of the most promising factors used frequently to increase productivity - the substrate feed rate [4-6,8-10] is also the most challenging one. Overfeeding the physiological capacity to metabolize substrate (q_{Scrit}) impairs productivity [11] and leads to unwanted overflow metabolism or even substrate accumulation. But rather than being a constant q_{Scrit}, is dynamically changing in response to process parameters e.g. pH, temperature and induction time q_{Scrit}=f(t,pH,T,…). This dependency of q_{Scrit} makes its quantification especially laborious but necessary to avoid overflow metabolism and substrate accumulation.

Conventional process development investigates the impact of process parameters on productivity within a design of experiment (DoE). To ensure feasibility of the experiments the DoE is commonly been designed within the technological and

physiological feasible space. While technologic constraints (e.g. k_La, heat transfer rate) are setup specific and therefore commonly known, physiological constraints have to be assessed strain and product specific. Especially in respect of the substrate feeding rate the definition of the physiological feasible space is of great concern. Exceeding the physiological feasible space leads to metabolite formation and substrate accumulation, which negatively affect physiology [12]. Moreover substrate accumulation has been shown to negatively affect productivity [11] as well as controllability of the respective bioprocess. Consequently, q_{Scrit} constitutes the main constraint to the physiologic feasible space. Given the importance of q_{Scrit} various approaches for the quantification of q_{Scrit} have been discussed in literature [13-16].

Other factors investigated within the respective DoE can potentially impact q_{Scrit} e.g. temperature and pH. Additionally, q_{Scrit} has been shown to be dependent on induction time [17,18] as well as on the level of metabolic activity [17]. Concluding, q_{Scrit} comprises a highly dynamic nature, which boosts the necessary effort for strain characterization regardless of the experimental approach utilized for quantification of q_{Scrit}. To reduce dependencies and to increase transferability bioprocess development has increasingly focused on specific physiological rates [19-21,14] rather than on volumetric feeding rates. To overcome the challenge of a dynamically changing q_{Scrit} and to avoid substrate accumulation a feedback control approach of a physiological variable is necessary. Focusing on specific physiological rates for process physiological development requires real time biomass estimation and physiological bioprocess control.

Various approaches have been outlined to overcome the challenge of biomass estimation and consequently facilitate the control specific physiological variables. For biomass estimation in general, literature favors data driven models or hybrid models [22,23]. But for bioprocess development available historic process data is commonly scarce, which restricts the use of data based algorithms. In this context hard type sensor and first principle mass balance based approaches for biomass estimation are regarded as more feasible than data driven approaches. A weighted average based combination of redundant biomass estimations has been shown as highly beneficial to increase accuracy and robustness of the biomass estimation[24].

Controllers

Besides biomass estimation approach physiological control approaches can be discriminated by the controller category.

Two main categories of controllers are employed for physiological process control. While open loop controllers (feed forward) do not measure or estimate the current state of the controlled variable closed loop respectively feedback controllers derive the current state of the controlled variable as process value from a direct or indirect measurement. Direct measurements refer to online measurements of e.g. pH, temperature, DO2 and off gas. Indirect measurements or estimations refer to computational values derived from direct measurements. For physiological process control the biggest challenge is the estimation of the variable of interest, since it commonly cannot be measured directly. Consequently the quality of the primary data as basis for subsequent computations is of often underestimated importance. In respect of transferability to industrial scale the number of necessary direct measurements is crucial, since manufactures tend to minimize of measurement device ports to avoid contamination sites.

Open loop controllers are most commonly employed for physiological process control . The high level of simplicity and its robustness concerning measurement errors are the main reasons for the common use.

Feedback control

While technological feedback controllers are widely used e.g. for temperature and pH, examples of closed loop control of physiological variables are extremely scarce. The reason for the limited examples for feedback control is the necessity to

determine the variable of interest with sufficient accuracy. Only if the variable of interest can be determined fast enough and with a sufficient signal to noise ratio closed loop control is feasible. Nevertheless, closed loop control is of great interest as real time quality insurance [20], owned to its capability to react on process perturbations constitutes the basis of the attributed robustness of closed loop control approaches.

In the stage of bioprocess development strain specific historic process data is commonly scarce the accuracy of multivariate approaches (e.g. artificial neuronal networks), which commonly feature a high degree of accuracy does not appear feasible. An light data driven example, merely using a set of three experiments, Jenzsch et al. used an extended Kalman filter for biomass estimation in combination with generic model control [22] for closed loop control of the specific growth rate. Examples of algorithms independent of apriori information and complex mathematical models have been introduced [21,25] although provided experimental data is scarce.

q_s maximation for productivity

Given the frequent correlation of substrate uptake rate and productivity a control objective can be to run a bioprocess at the highest possible substrate level below q_{Scrit}. These control approaches do not necessarily qualify as closed loop control approaches since the system response used for controller actions is of qualitative nature and not directly correlated to the controlled variable. Hereby usually a basic open loop control approach is combined with a probing technique e.g. in order to assess the reaction of the culture to sudden substrate starvation or excess. Monitoring the DO2 as response to short substrate up-pulsing it can be determined whether the culture is being overfed or if the open loop feed profile can continue increasing the volumetric feed flow rate [26,16]. Based on the same principle also down pulsing by temporary intermittence of the substrate feed rate can be utilized to assess the substrate supply situation in order to maximize the substrate uptake during induction phase [14].

More advanced techniques target a quantitative assessment of the metabolic state and substrate accumulation. Since the direct sensing of metabolites and substrate is only possible with substantial analytical effort using an online HPLC or FTIR this approach is commonly not regarded as feasible to pilot or even production scale. Merely based on off-gas analysis the respiratory quotient (RQ) can provide valuable insight into the metabolism of the cell [27]. But sensing overflow metabolism based on the RQ is only feasible if the metabolite has a different degree of reduction than the substrate. In case of glucose as substrate and acetate as metabolite this approach is consequently not feasible. Nevertheless, using first principle mass balances differentiating between oxidative and oxido-reductive metabolic states is possible as shown by the approach of Jobe et al. [27].

We target an independent estimation of biomass and the oxidative metabolism of the culture using first principle mass balances. Since this approach requires off-gas analysis besides a volume balance and merely the biomass composition as strain specific information this approach shall feature a high degree of transferability. Using a simple algorithm the controller shall facilitate avoiding substrate accumulation rather than mere sensing of the latter.

Goals

The goal of this contribution is the introduction of a transferable control concept independent of historic process data. Thereby the approach shall be employable even in early bioprocess development. Given the relevance of physiological bioprocess development this approach shall facilitate a reduction in necessary strain characterization experiments, by sensing

q_{Scrit} in real time. The feasibility of the introduced approach to effectively avoid substrate accumulation, despite a rapidly declining q_{Scrit} , shall assessed within a fermentation.

Materials and Methods

Cultivations

Bioreactor system

Fermentations were conducted in a DASGIP multi-bioreactor system with 4 parallel reactors with 2L of working volume each (Eppendorf; Hamburg, Germany). The reactors were equipped with baffles and three disk impeller stirrers. The DASGIP control software v4.5 revision 230 was used for control: pH (Hamilton, Reno, USA), pO2 (Mettler Toledo; Greifensee, Switzerland; module DASGIP PH4PO4), temperature and stirrer speed (module DASGIP TC4SC4), aeration (module DASGIP MX4/4) and pH (module DASGIP MP8). CO2, O2 concentrations in the off-gas were quantified by a gas analyzer (module DASGIP GA4) using the non-dispersive infrared and zircon dioxide detection principle, respectively.

Strain and media

A recombinant BL21 DE3 *E.coli* strain was cultivated, producing an intracellular protein (~30 kDa) in form of inclusion bodies, after a one-time induction with IPTG (1 mM). The synthetic media was based on the recipe of Korz, Rinas et al. [28], where the limiting C-source was glucose.

Process parameters

Pre-cultures were incubated at 30°C and 170 rpm to an OD_{600} of approx. 1.5 in 150 mL batch media and 2.5% batch volume aliquots were used for inoculation. After depletion of the C-source in an initial batch phase, the pre-induction fed-batch was started. The pre-induction feeding strategy was based on an exponential feed forward profile to maintain a predefined growth rate. On attainment of the predefined biomass the cultures were induced after 30 min adaption time. Stirrer speed was set to 1400 rpm and aeration to 1.4 v/v/m for the whole process. The pH was maintained at 6.9 by adding 12.5% NH_4OH, which also served as nitrogen source. The dissolved oxygen (DO2) was kept over 30% by supplementing oxygen to the air.

Pre induction: exponential feed forward profile

The starting feed rate in L/h (F_0) was calculated using a gravimetrical biomass yield in g/g ($Y_{X/S,g}$), the starting biomass in g (X_0), the concentration of the feed solution in g/L ($c_{S,g}$) as well as the specific biomass growth rate as described elsewhere [29]

Post induction feeding strategy

After the depletion of C-source in the batch phase the culture was induced with IPTG (1 mM). A step controller was used for *real time q_S feedback control* during the post induction phase. Therefore the feed rate is dynamically adjusted every 20 minutes, which is calculated with the base load (calculated with the setpoint) and the adjustment term (calculated with the difference between the setpoint and the process value) to actually reach the given set-point of q_S.

$$\dot{F}_{i+1} = \frac{q_S SP_{i+1} \cdot X_i}{c_S} + \frac{(q_S SP_i - q_S PV_i) \cdot X_i}{c_S}$$

Equation 1: Feed rate for next control interval

Two fermentations with $q_S SP = 0.02\ g/g/h$ were performed parallel. For on fermentation the algorithm to detect reaching $q_S SP$ was deactivated to keep $q_S SP$ unmodified, while the second fermentation was performed with activated algorithm to adapt $q_S SP$ automatically, if necessary.

Process evaluation and data analysis

Metabolic rates and yield coefficients were calculated with Matlab r2013 b (Mathworks; Natick, Massachusetts, USA). The calculation of specific rates and yield coefficients was conducted as described elsewhere [30].

Offline analytics

Biomass dry weight (CDW)

Biomass concentrations were gravimetrically quantified after drying at 105°C for min. 72 h. Therefore 2 mL of culture broth were centrifuged (4500 x g, 10 min, 4°C) in a pre-weighted glass tube and the pellet was washed once with 5 mL RO water. The determination was done in duplicates. After drying in the drying oven, the biomass dry weight was measured on a scale.

Substrate concentration and small metabolites

The C-source concentration in the feed media was calculated using the gravimetrically determined density. NH_4OH concentration was determined by titration with 1 M HCl. Acetate concentrations were quantified from the supernatant by enzymatic photometric principle in a robotic system (Cedex BioHT, Roche, Switzerland). The analysis was used as a quality control to exclude possible acetate production due to oxygen limitation or overflow metabolism.

Results

Physiological maxima for physiologic DoE design

Rather than controlling a technological variable at a constant level, physiological process control targets maintaining a physiological variable. To illustrate the goal of physiological bioprocess control **Fig. 1** illustrates biomass growth and the accordingly increased substrate flow rate. The most common approach for physiological control is based on a fixed biomass yield, neglecting the dynamics of the biomass yield for the sake of simplicity. **Fig. 1**a shows a more advanced control approach, including a real time biomass estimation from a softsensor. Following the control scheme of **Fig. 1**b despite real time biomass estimation no estimate of the controlled variable is made. Instead of following a predefined substrate flow profile, the substrate flow rate is calculated incrementally via softsensor biomass estimation.

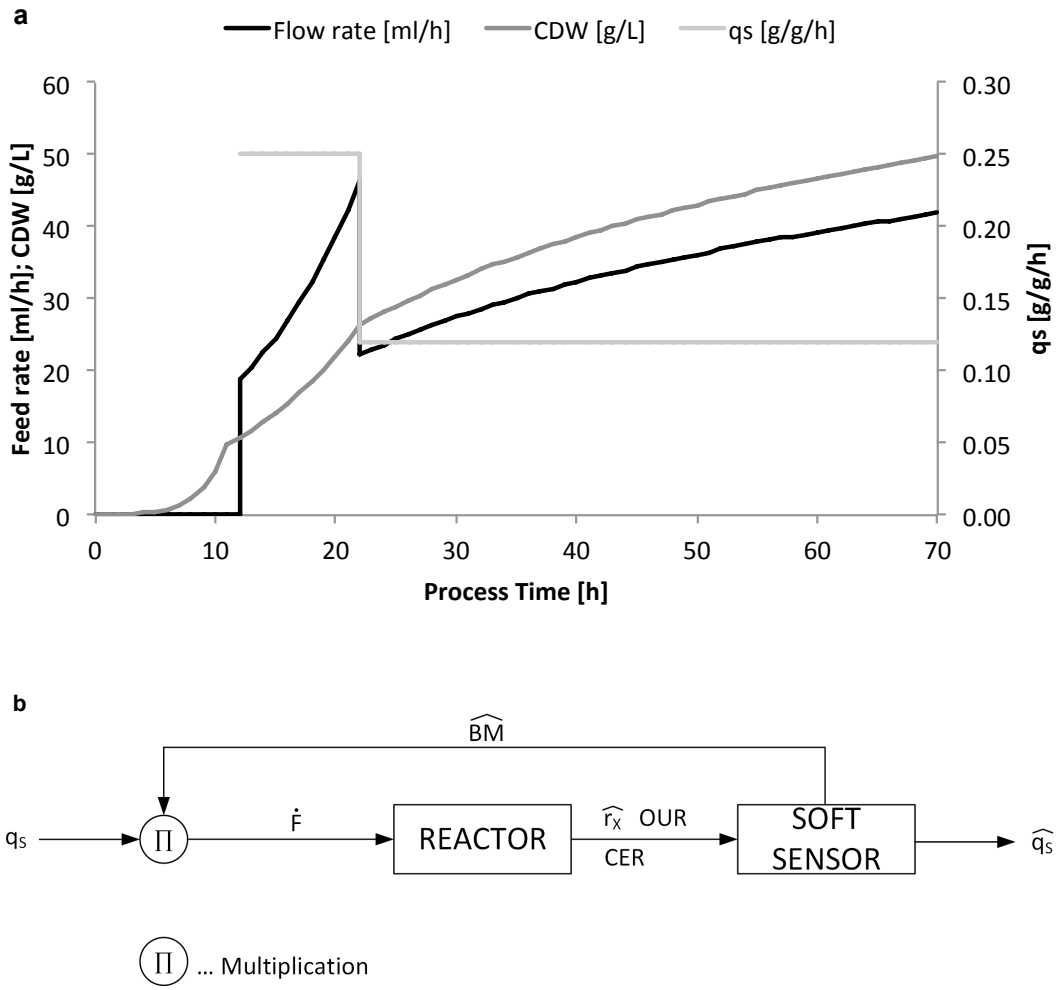

Fig. 1 State of the art physiological open loop process control does not provide quantitative real time information about the controlled variable **a** Open loop control including real time biomass estimation based on a softsensor with incremental calculations of the suitable substrate flow rate. **b** Open loop controller design for incremental feed forward calculations

In order to bring the limitations of the previously described controller setup to the reader's attention **Fig. 2** outlines a fed batch process including substrate accumulation. Upon glucose accumulation due to overfeeding, the biomass estimation of the softsensor clearly deviates from the offline verification data. This is owned to the fact that the used softsensor was based on the assumption that the accumulation term in the C-balance equals zero. Accordingly biomass estimation works well until substrate accumulation.

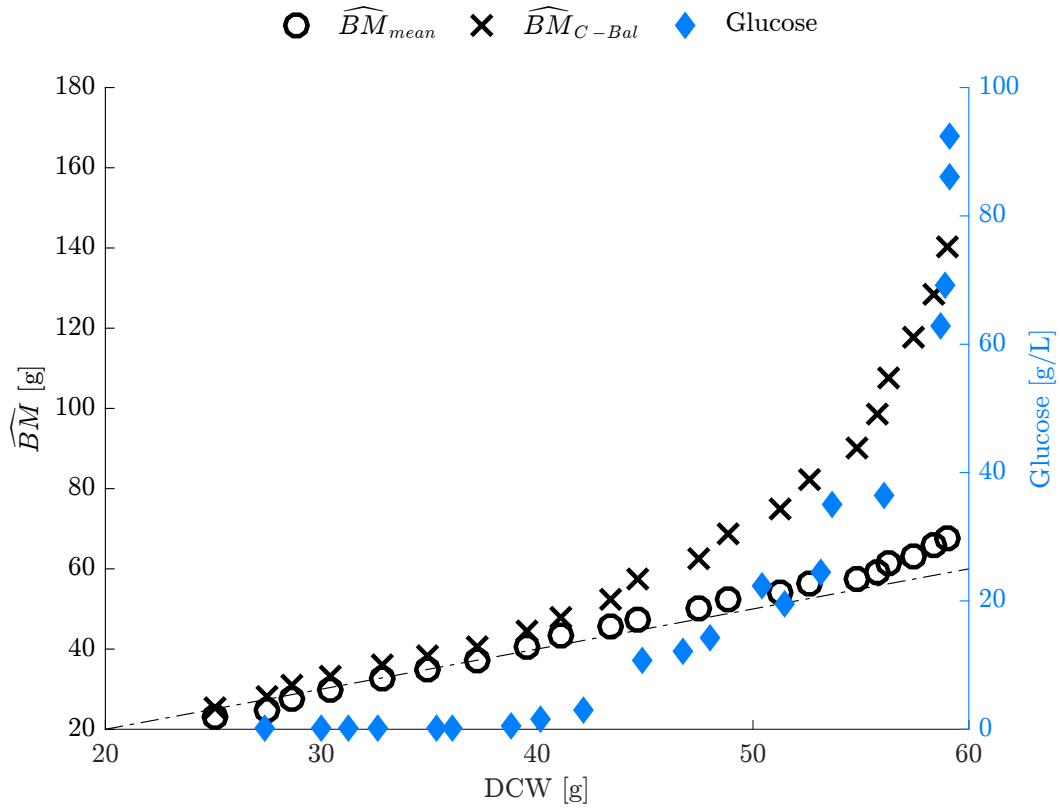

Fig. 2 BM estimation based on C-balance is impaired by substrate accumulation;BM real-time estimations (○,×) and accumulation of glucose (♦). The BM estimation via the C balance (×) shows an increasing error upon glucose accumulation and is not suitable for high q_s fermentations due to the possibility of reaching $q_{s,crit}$ and accumulating substrate. The BM estimation via weighted average of permittivity, DOR- & N balance (○) provides a more reliable estimation during glucose accumulation and was therefore chosen for the experiments

Given the sensitivity of the C-balance based softsensor an alternative approach for biomass estimation is necessary. Altering the softsensor setup can greatly benefit the quality of estimation as well as the robustness of estimation. **Fig. 3** shows the result of using balancing approaches in combination with a permittivity measurement in order to obtain reliable estimates of the biomass rate (r_X) and the oxidative substrate metabolization rate (r_{Sox}). Using a weighted average approach, the biomass estimations based on the N-balance, the DoR balance and a permittivity measurement are combined to make the approach robust against accumulation. Including the DoR balance is in this case straight forward since the primary metabolite (acetate) shares the same degree of reduction as the substrate. **Fig. 3** illustrates the correlated noise on the estimation of the most crucial rates: the biomass rate (**Fig. 3**a) and the oxidative substrate uptake rate (**Fig. 3**b). In case of substrate accumulation (data not shown) the C-balance based estimation quickly deviates from the rate derived from offline references biomass analytics. In comparison to the estimations merely based on the C-balance the noise on the rates of the weighted estimations lies within the order of magnitude of noise observed for the verification data that were calculated offline.

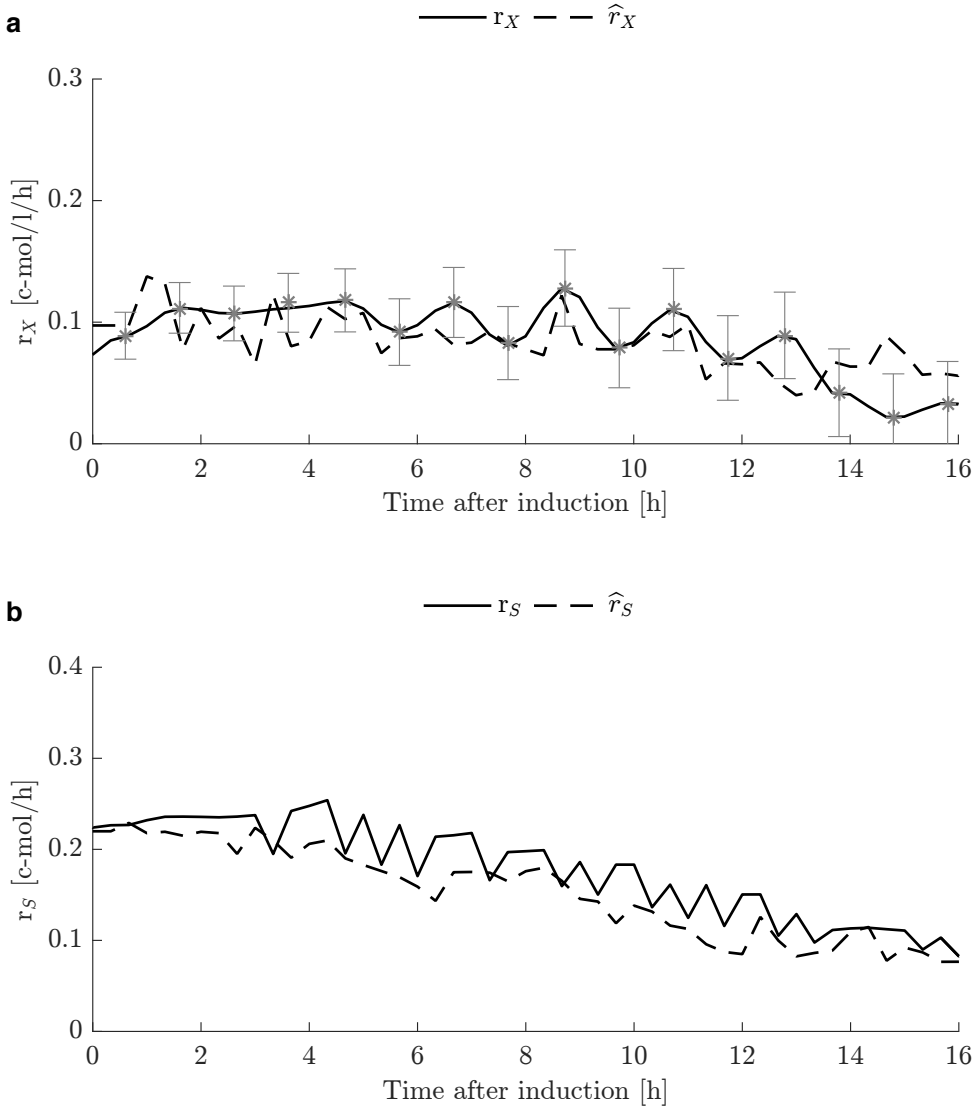

Fig. 3 Biomass and substrate uptake rate estimation vs. verification; **a** Biomass growth rate estimation (dashed line) based on N and DoR balance and permittivity is congruent to offline data (solid line). The BM accumulation rate r_x is one of three needed values to estimate the actual substrate conversion rate and furthermore the specific substrate uptake rate q_s. The biomass growth rate based on C balance (dotted line) shows large deviations due to glucose accumulation. **b** Estimation of substrate conversion rate (dashed line) reconciled from r_x, CER and OUR shows great congruence to offline data (solid line). The actual substrate conversion rate r_s in combination with the biomass is needed to estimate the actual specific substrate uptake rate q_s. The substrate conversion rate based on C balance (dotted line) shows large deviations due to glucose accumulation.

Upon the proof of principle for the real time estimation of r_X and r_S the calculation of a process value of q_S, as controlled variable, becomes possible. **Fig. 4**a explains the feedback controller design to close the loop for the controlled variable q_S by incorporating the obtained values into the previously introduced controller scheme. In contrast now a process value for q_S can be calculated and used for controller action. To increase robustness only the feedback of the process value of q_S impacts the controller action. The basal substrate flow rate is being calculated in time increments based on the biomass estimation, as

explained previously (**Fig. 1**b). Nevertheless, despite closing the loop the q_S setpoint is not maintained and glucose accumulation occurs quickly during induction phase.

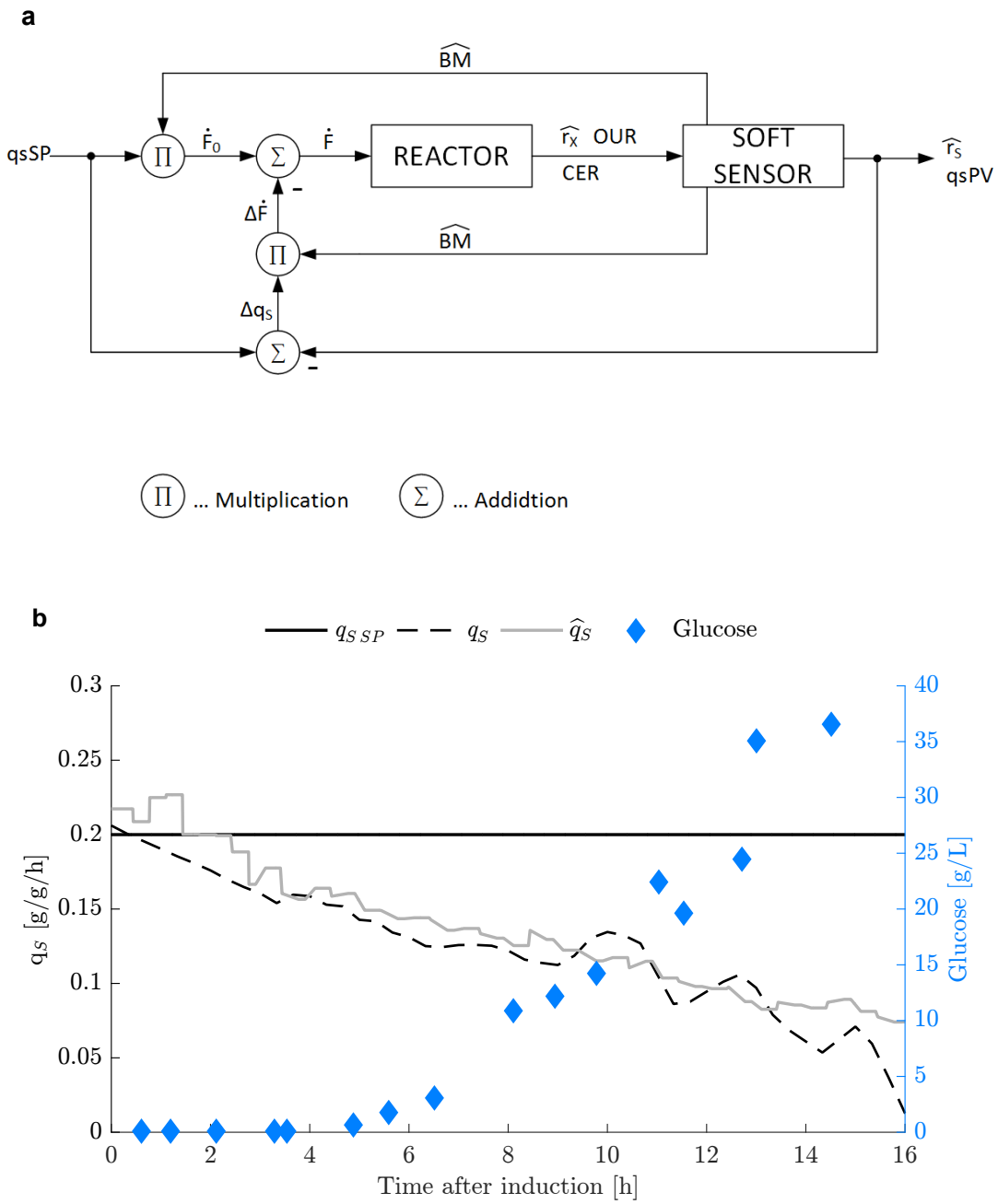

Fig. 4 Closed loop control approach and its limitations; **a** Closed loop control provides the possibility of comparing the wanted setpoint with the actually achieved process value. Therefore an interference of the process can improved the quality of the control strategy. **b** Closed loop control of q_s is insufficient without adapting the setpoint, if $q_{s,crit}$ is reached. Significant deviation from q_sSP (solid line) due to

q_{Scrit} decline. Real time estimation of q_SPV (dotted line) compared to actually achieved q_S (dashed line) is within 25% error. High glucose accumulation (blue dots) due to reaching q_{Scrit}

Owned to the physiological decline of q_{Scrit} the setpoint of q_S cannot be maintained without substrate accumulation. The cells are simply not capable of metabolizing the amount of substrate defined by the setpoint. As a consequence the setpoint has to be adapted according to the decline in q_{Scrit}. Given the target of reducing the effort for strain characterization, using a predefined q_SSP limitation is not viable. Instead a simple algorithm is required to assess whether q_{Scrit} is reached or not.

Fig. 5 illustrates the underlying principle of the q_{Scrit} control. By comparing the behavior of the culture upon controller action to previous behavior q_{Scrit} becomes obvious. If the process value of q_S does not increase after consecutive increase in substrate flow the setpoint of q_S needs to be decreased since q_{Scrit} has been reached. **Fig. 5**a displays the in cooperation of the step controller displayed in **Fig. 5**b.

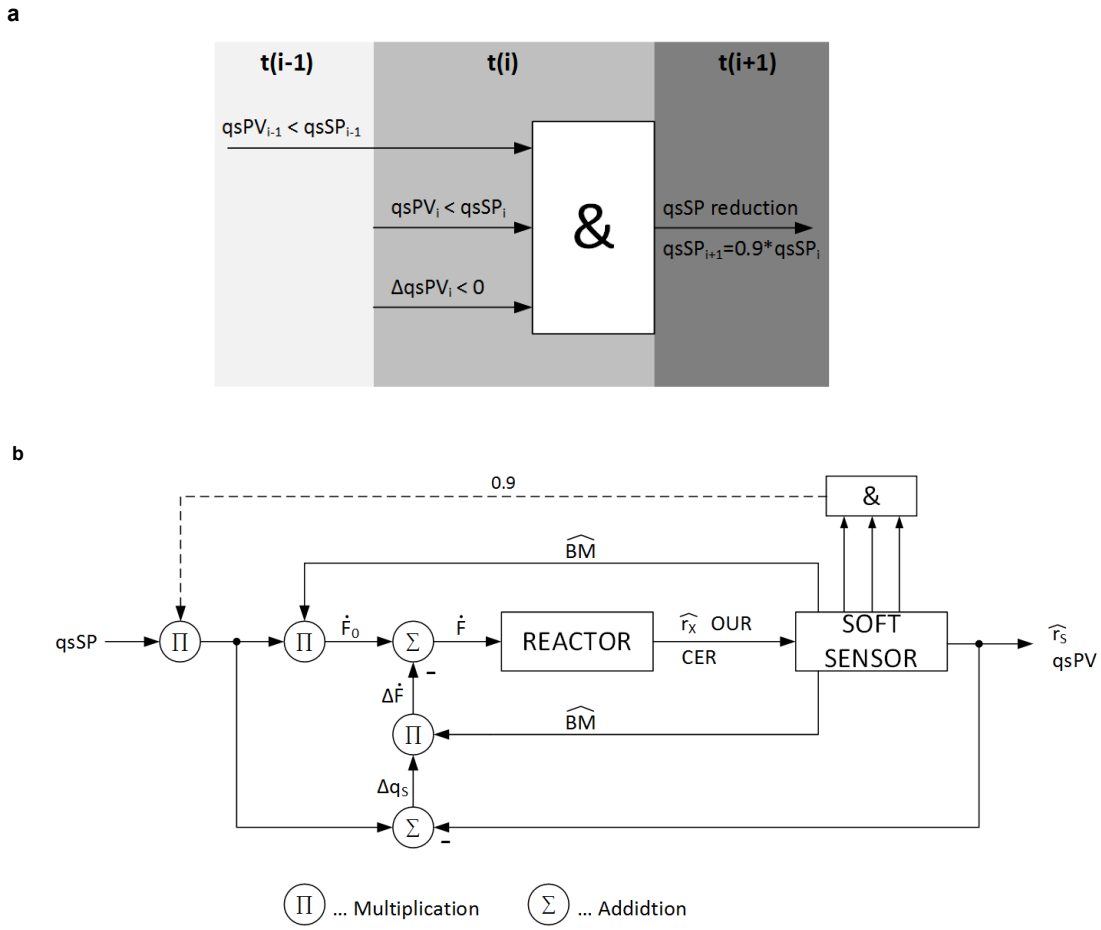

Fig. 5 Closed loop control with setpoint adaptation; **a** A logical query to adapt q_SSP, if q_{Scrit} is reached was implemented. Three conditions need to be true to adapt q_SSP: 1) The q_SSP from time point t(i-1) 20 minutes ago was larger than last q_SPV, 1) The q_SSP from time point t(i) is larger than the current q_SPV, 3) the change of the q_SPV within the last 20 minutes (Δq_SPV) was negative. In case all three conditions are true, q_SSP is be reduced by 10%. Since in case of substrate limitation q_SPV should always increase upon a q_SSP (federate) increase, if q_SPV is smaller than q_SSP and q_{Scrit} is not reached, this logical query provides a simple method to detect q_{Scrit}. **b** If the system hits its natural limit, a setpoint higher than this limit cannot be achieved and therefore has to be adapted

The benefit of the introduced control approach is displayed in **Fig. 6**. Although the same q_SSP as in **Fig. 4** has been used, no accumulation occurs upon decline of q_{Scrit}. The controller effectively avoids substrate accumulation throughout the whole induction phase by reducing q_SSP in case of reaching q_{Scrit}.

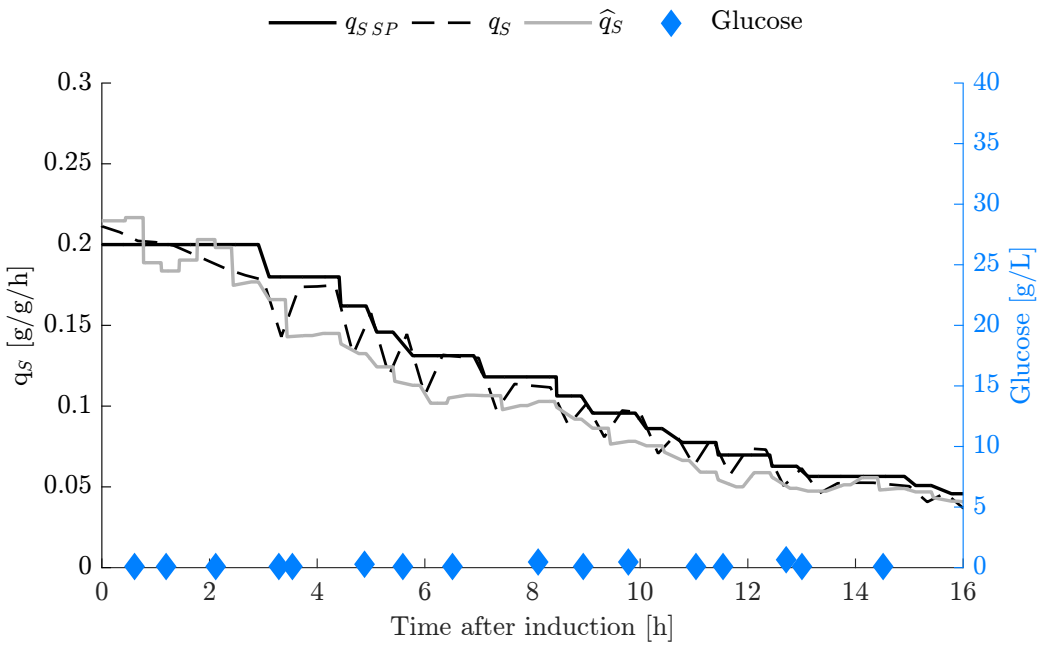

Fig. 6 Closed loop control of q_s with adaption of q_sSP (solid line) helps to avoid a breach of q_{Scrit} and facilitates an accurate control of q_s; Real-time estimation of q_s process value (dotted line) compared to actually achieved q_s process value (dashed line). No significant glucose accumulation (blue dots) due to adaption of q_sSP

Owned to the fact that this controller is merely based on first principle balances and a permittivity probe, the analytical as well as the computational effort is relatively lean. Consequently the introduced concept appears highly transferable even to an industrial environment.

Discussion

Within this contribution we introduced a process control approach capable of physiological feedback control and real time sensing of the physiological capacity to metabolize substrate (q_{Scrit}).

Feedback control approaches are highly challenging and require a high accuracy of estimation. Only if a viable signal to noise ratio can be obtained a physiological closed loop approach becomes feasible. Nevertheless, various contributions have claimed closed loop control up to this date.

In the contribution of Sagmeister et al., using a first principle softsensor, closed loop control of q_S was claimed. A real time estimation of biomass was used to control the substrate feed rate in order maintain q_S throughout induction [31]. Although real time process data was utilized for the estimation of biomass, no process value of q_S or other physiological variables was obtained. Hence, the feed rate is being controlled but lacks the calculation of the process value of any physiological variable not as physiological closed loop control approach.

A very similar concept has been introduced by Jobe et al. [27] to sense the metabolic status of the cell in real time. Nevertheless, this approach did not fulfill the requirements of a closed loop control, since no process value of controlled physiological variable was calculated. Calculating an oxidative and an oxireductive metabolic model every 4 min a statistical test was used for the evaluation of the current metabolic state. The substrate feed rate was controlled by an exponential feeding profile, of which the exponent was subjected to controller actions based on the decision concerning the statistical test. Hereby, no process value of the controlled variable μ was calculated and consequently the prerequisites for a closed loop control approach not met. Besides the nomenclature, although the acetic acid was being accumulated the approach of Jobe et al. lacked the sensitivity to take action. The accumulation of acetate was not pronounced enough to trigger controller action; instead the μ controller remained idle. The growth in biomass subsequently decreased the specific growth rate to a level which allowed the uptake of acetate. In contrast the approach presented within this contribution did effectively circumvent the accumulation of acetate and substrate.

Dabros et al. [32] introduced an algorithm for physiological closed loop control and illustrated its feasibility in non-induced *E.coli* cultures. Using an exponential feed profile a basic substrate feed rate was calculated. The deviation of the process value of μ from μ setpoints triggered an additional PI controller action. Although accumulation was measured using FTIR, the data was only used for reconciliation to improve biomass estimation. If the decline in q_{Scrit} leads to an inevitable decrease in specific growth rate, despite substrate accumulation the controller action would lead to a continuous feed rate increase.

As physiological closed loop controller Jenzsch et al. [22] used an extended Kalman filter for biomass estimation in combination with generic model control. Merely using a set of three experiments the model was trained and verified with an additional experiment prior the utilization for process control. This approach was shown to accurately achieve different distinct μsp in induction phase of a microbial bioprocess producing GFP. Despite the good performance the generic model control was based on a constant value for the q_{Scrit} as well as a static value of the biomass yield. Owned to the underlying hypothesis of a constant biomass $Y_{X/S}$ and q_{Scrit} this approach appears sensitive to reaching the physiological capacity q_{Scrit}.

Schaepe et al. [18] used a feedback control of total OUR to avoid overfeeding in high cell density cultures of *E. coli*. Three experiments were needed to first determine the q_{Scrit} profile and subsequently the OUR setpoint profile for feedback control.

q_{Scrit} was quantified using a sigmoid function to detect saturation of total OUR as a result of increasing feed rate. Taking a safety distance from q_{Scrit} the derived q_S profile was used to obtain the OUR setpoint profile. The approach showed high reproducibility in the cultivations although it neglected the dynamics of the biomass yield to calculate the feed rate profile. During OUR feedback control neither real time detection of reaching q_{Scrit} nor biomass estimation was realized. Accumulation of substrate could occur in case of deviations in the initial conditions, e.g. the initial biomass or substrate concentrations or disturbances appearing during the process, e.g. problems with the oxygen mass transfer or temperature control.

Conclusions

The goal of this paper was the introduction of a transferable control concept capable of effectively avoiding substrate accumulation as well as the illustration of the feasibility of the introduced approach. In respect of the state of the art we were able to establish the following points:

- ➔ Using a combination of first principle mass balances and permittivity measurement we were able illustrate a transferable concept to independently estimate biomass concentration as well as the rate of oxidative substrate metabolization.
- ➔ Physiological feedback control makes a process value of the physiologic variable accessible but is not sufficient in order to avoid substrate accumulation. This circumstance is owned to the fact that physiologic capacity to metabolize substrate declines over time, making it impossible to maintain the setpoint if q_{Scrit} is reached.
- ➔ Using a simple step controller substrate accumulation can be effectively avoided by setpoint adaptation in response to the violation of physiologic capacities.

The illustrated approach facilitates robust process development without relying on comprehensive strain characterization. This concept is theoretically not limited to the specific substrate uptake rate. Owned to the underlying first principle mass balances this concept appears highly transferable in comparison to data driven alternatives. Furthermore, physiologic feedback control including the introduced controller could be used to complete replace conventional strain characterization in terms of physiological capacities. This will ultimately decrease the effort for process development significantly since not only the effort for strain characterization can be minimized but also number of fail batches owned to substrate accumulation can be greatly reduced.

Declarations

Acknowledgements

We are grateful for the financial support of the Sandoz GmbH and the Christian Doppler Society Austria.

Authors' contributions

CH supervised the study. WR, AK and FJ conceived the study and the experimental design. AK performed the cultivations as well as analytics and created all figures. WR and AK processed and analyzed the data. WR wrote the first manuscript version. All authors participated in the final manuscript version, read and approved the final manuscript.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Terpe K (2006) Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Applied Microbiology and Biotechnology 72 (2):211-222. doi:10.1007/s00253-006-0465-8

2. Walsh G (2010) Biopharmaceutical benchmarks 2010. Nat Biotech 28 (9):917-924

3. Rathore AS, Winkle H (2009) Quality by design for biopharmaceuticals. Nat Biotech 27 (1):26-34

4. Levisauskas D, Galvanauskas V, Henrich S, Wilhelm K, Volk N, Lübbert A (2003) Model-based optimization of viral capsid protein production in fed-batch culture of recombinant Escherichia coli. Bioprocess and Biosystems Engineering 25 (4):255-262. doi:10.1007/s00449-002-0305-x

5. Ramalingam S, Gautam P, Mukherjee KJ, Jayaraman G (2007) Effects of post-induction feed strategies on secretory production of recombinant streptokinase in Escherichia coli. Biochemical Engineering Journal 33 (1):34-41. doi:http://dx.doi.org/10.1016/j.bej.2006.09.019

6. Sagmeister P, Schimek C, Meitz A, Herwig C, Spadiut O (2014) Tunable recombinant protein expression with E. coli in a mixed-feed environment. Appl Microbiol Biotechnol 98 (7):2937-2945. doi:10.1007/s00253-013-5445-1

7. Chen X, Liu L, Li J, Liu J, Du G, Chen J (2011) Optimization of glucose feeding approaches for enhanced glucosamine and N-acetylglucosamine production by an engineered Escherichia coli. J Ind Microbiol Biotechnol 39 (2):359-365. doi:10.1007/s10295-011-1046-0

8. Babaeipour V, Shojaosadati Seyed A, Khalilzadeh R, Maghsoudi N, Tabandeh F (2008) A proposed feeding strategy for the overproduction of recombinant proteins in Escherichia coli. Biotechnology and Applied Biochemistry 49 (2):141. doi:10.1042/ba20070089

9. Kavanagh JM, Barton GW (2008) Productivity improvement of recombinant Escherichia coli fermentation via robust optimization. Bioprocess Biosyst Eng 31 (2):137-143. doi:10.1007/s00449-007-0156-6

10. Sanden AM, Prytz I, Tubulekas I, Forberg C, Le H, Hektor A, Neubauer P, Pragai Z, Harwood C, Ward A, Picon A, De Mattos JT, Postma P, Farewell A, Nystrom T, Reeh S, Pedersen S, Larsson G (2003) Limiting factors in Escherichia coli fed-batch production of recombinant proteins. Biotechnol Bioeng 81 (2):158-166. doi:10.1002/bit.10457

11. Jensen EB, Carlsen S (1990) Production of recombinant human growth hormone in Escherichia coli: expression of different precursors and physiological effects of glucose, acetate, and salts. Biotechnol Bioeng 36 (1):1-11. doi:10.1002/bit.260360102

12. Luli GW, Strohl WR (1990) Comparison of growth, acetate production, and acetate inhibition of Escherichia coli strains in batch and fed-batch fermentations. Applied and environmental microbiology 56 (4):1004-1011

13. Akesson M, Karlsson EN, Hagander P, Axelsson JP, Tocaj A (1999) On-line detection of acetate formation in Escherichia coli cultures using dissolved oxygen responses to feed transients. Biotechnol Bioeng 64 (5):590-598

14. Henes B, Sonnleitner B (2007) Controlled fed-batch by tracking the maximal culture capacity. Journal of Biotechnology 132 (2):118-126. doi:http://dx.doi.org/10.1016/j.jbiotec.2007.04.021

15. Lin HY, Mathiszik B, Xu B, Enfors SO, Neubauer P (2001) Determination of the maximum specific uptake capacities for glucose and oxygen in glucose-limited fed-batch cultivations of Escherichia coli. Biotechnol Bioeng 73 (5):347-357

16. Åkesson M, Hagander P, Axelsson J (1999) A probing feeding strategy for Escherichia coli cultures. Biotechnology Techniques 13 (8):523-528. doi:10.1023/A:1008906517033

17. Reichelt WN, Brillmann M, Thurrold P, Keil P, Fricke J, Herwig C (2017) Physiological capacities decline during induced bioprocesses leading to substrate accumulation. Biotechnology journal 12 (7)

18. Schaepe S, Kuprijanov A, Simutis R, Lubbert A (2014) Avoiding overfeeding in high cell density fed-batch cultures of E. coli during the production of heterologous proteins. Journal of Biotechnology 192:146-153. doi:10.1016/j.jbiotec.2014.09.002

19. Levisauskas D, Simutis R, Borvitz D, Lübbert A (1996) Automatic control of the specific growth rate in fed-batch cultivation processes based on an exhaust gas analysis. Bioprocess Engineering 15 (3):145-150. doi:10.1007/BF00369618

20. Gnoth S, Jenzsch M, Simutis R, Lubbert A (2008) Control of cultivation processes for recombinant protein production: a review. Bioprocess Biosyst Eng 31 (1):21-39. doi:10.1007/s00449-007-0163-7
21. Levisauskas D (2001) Inferential control of the specific growth rate in fed-batch cultivation processes. Biotechnology Letters 23 (15):1189-1195. doi:10.1023/A:1010528915228
22. Jenzsch M, Simutis R, Luebbert A (2006) Generic model control of the specific growth rate in recombinant Escherichia coli cultivations. J Biotechnol 122 (4):483-493. doi:10.1016/j.jbiotec.2005.09.013
23. de Assis AJ, Filho RM (2000) Soft sensors development for on-line bioreactor state estimation. Computers & Chemical Engineering 24 (2–7):1099-1103. doi:http://dx.doi.org/10.1016/S0098-1354(00)00489-0
24. Reichelt WN, Thurrold P, Brillmann M, Kager J, Fricke J, Herwig C (2016) Generic biomass estimation methods targeting physiologic process control in induced bacterial cultures. Engineering in Life Sciences 16 (8):720-730
25. Rocha I, Veloso ACA, Carneiro S, Costa R, Ferreira EC (2008) Implementation of a Specific Rate Controller in a Fed-Batch E. Coli Fermentation. IFAC Proceedings Volumes 41 (2):15565-15570. doi:10.3182/20080706-5-kr-1001.02632
26. Velut S, de Maré L, Hagander P (2007) Bioreactor control using a probing feeding strategy and mid-ranging control. Control Engineering Practice 15 (2):135-147. doi:http://dx.doi.org/10.1016/j.conengprac.2006.05.009
27. Jobe AM, Herwig C, Surzyn M, Walker B, Marison I, von Stockar U (2003) Generally applicable fed-batch culture concept based on the detection of metabolic state by on-line balancing. Biotechnol Bioeng 82 (6):627-639. doi:10.1002/bit.10610
28. Korz DJ, Rinas U, Hellmuth K, Sanders EA, Deckwer WD (1995) Simple fed-batch technique for high cell density cultivation of Escherichia coli. Journal of Biotechnology 39 (1):59-65. doi:http://dx.doi.org/10.1016/0168-1656(94)00143-Z
29. Wechselberger P, Sagmeister P, Engelking H, Schmidt T, Wenger J, Herwig C (2012) Efficient feeding profile optimization for recombinant protein production using physiological information. Bioprocess Biosyst Eng 35 (9):1637-1649. doi:10.1007/s00449-012-0754-9
30. Sagmeister P, Wechselberger P, Herwig C (2012) Information Processing: Rate-Based Investigation of Cell Physiological Changes along Design Space Development. PDA J Pharm Sci Technol 66 (6):526-541. doi:10.5731/pdajpst.2012.00889
31. Sagmeister P, Wechselberger P, Jazini M, Meitz A, Langemann T, Herwig C (2013) Soft sensor assisted dynamic bioprocess control: Efficient tools for bioprocess development. Chem Eng Sci 96:190-198. doi:10.1016/j.ces.2013.02.069
32. Dabros M, Schuler M, Marison I (2010) Simple control of specific growth rate in biotechnological fed-batch processes based on enhanced online measurements of biomass. Bioprocess and Biosystems Engineering 33 (9):1109-1118. doi:10.1007/s00449-010-0438-2

4 SUMMARIZING DISCUSSION AND OUTLOOK

The goal to investigate effects of combined q_S and temperature oscillations during induction phase on inclusion body solubility in *E. coli* could not be achieved due to technical limitations of the reactor setup. Literature and previous work of the working group describe effects of temperature and q_S in USP on IBs, which could be used to customize IBs (Luo, Leeman et al. 2006, Margreiter, Schwanninger et al. 2008, Brillmann 2015). Brillmann showed a negative correlation between the q_S amplitude during q_S oscillations and the IB solubility (Brillmann 2015), which could possibly be compensated for by diametrically opposed temperature oscillations as solubility increases with lower temperatures. Nevertheless, one must look at the bigger picture because temperature also affects the maximum physiological capacity. This again could have negative effects on specific product titer, which contradict the original idea of process optimization. The direct investigation of the effects of q_S and temperature oscillations on IB properties remains to be done.

Using NTA, a particle-based method, allowed high throughput IB sizing. IBs were labeled with Alexa 488 and measured with and without chemical fixation. Results indicate that chemical fixation on the one hand leads to decreased standard deviation as well as particle count but on the other hand increases reproducibility of NTA measurements. Furthermore, it can be concluded that high-pressure homogenization eliminates differences in IB size caused by USP. Additionally, comparing particle count of native and fixated samples could allow a measure for IB stickiness.

The decline of q_{Scrit} over induction time has been shown previously (Schaepe, Kuprijanov et al. 2014, Reichelt, Brillmann et al. 2017) and was also demonstrated in the experiments of this thesis. This decline represents a challenge for physiological process control since substrate accumulation needs to be avoided to maximize productivity. A first-time real-time estimation of q_S allowed a closed loop control of q_S. Furthermore, a logical query to detect reaching q_{Scrit} enabled automatic adaption of the setpoint to avoid substrate accumulation. Schaepe et al. (Schaepe, Kuprijanov et al. 2014) established a similar physiological closed loop control, using the CER signal of earlier batches to keep the culture just below q_{Scrit} during induction time. Although this approach lacks a real-time estimation of q_S, it offers a simpler setup since off-gas analysis is the only measurement necessary. The method developed in this work offers the advantage that no previous batches are necessary and different temperatures or biomass concentrations in USO can easily be used without new experiments to obtain needed data. Using this approach strain characterization experiments can be greatly reduced or even eliminated.

5 APPENDIX

5.1 LIST OF ABBREVIATIONS

CPP	Critical process parameters
DCW	Dry cell weight
DSP	Downstream processing
IB	Inclusion body
IB-QA	Inclusion body quality attributes
IPTG	Isopropyl-β-D-thiogalactopyranosid
NTA	Nano-particle tracking analysis
QbD	Quality by Design
USP	Upstream processing

5.2 LIST OF SYMBOLS

$\widehat{BM}_{C-Bal}$	Real-time estimated biomass based on the C-balance (g)
$\widehat{BM}_{mean}$	Real-time estimated biomass based on the N-balance, the DoR balance and a permittivity measurement (g)
c_S	Substrate concentration in feed (g/L)
$\dot{F}_0$	Starting feed rate (mL/h)
μ	Specific biomass growth rate (1/h)
q_S	Specific substrate uptake calculated offline (g/g/h)
q_{Scrit}	Critical specific substrate uptake rate as defined by Åkesson, Hagander, & Axelsson, 1999 (g/g/h)
q_SPV	Real-time process value of q_S (g/g/h)
q_SSP	Setpoint of q_S (g/g/h)
Δq_SPV_i	Change of q_SPV within the last control interval (20 min) (g/g/h)
r_{acc}	Rate of accumulating substrate and acetate (c-mol/h)
r_{CO2}	CER, carbon dioxide evolution rate (mol/h)
r_{O2}	OUR, oxygen uptake rate (mol/h)
r_S	Substrate conversion rate calculated offline (g/h or c-mol/h)
$\widehat{r_S}$	Real-time estimated substrate conversion rate (g/h or c-mol/h)
r_X	Biomass growth rate calculated offline from DCW (g/h or c-mol/h)
$\widehat{r_X}$	Real-time estimated biomass growth rate (g/h or c-mol/h)

RQ	Respiratory quotient (mol/mol)
V_0	Volume at t = 0 (L)
X_i	Biomass dry cell weight at t = 0 batch end or at time point t = I (g)
$Y_{X/S}$	Biomass yield on substrate (g/g or c-mol/c-mol)

5.3 LIST OF FIGURES

5.4 DATA RECONCILIATION FOR K2S1 SOFTSENSOR

```
% reconcilation

function [ output ] = K2S1_v2(F_s, c_s, M_s, F_a_in, F_o2_in, y_o2_out,
y_co2_out, y_wet, y_o2_in_air, y_co2_in_air, gamma_s, gamma_o2,
gamma_x, Vol)

% inputs:-----------------------------------------------------------
---

% calculated variables (internally)---------------------------------
---

y_o2_in=(F_a_in.*y_o2_in_air+F_o2_in)./(F_a_in+F_o2_in);
y_co2_in=(F_a_in.*y_co2_in_air)./(F_a_in+F_o2_in);

ex_h2o=1-(y_wet./y_o2_in_air);
Ra_inert = (1-y_o2_in_air-y_co2_in_air)./(1-y_o2_out-y_co2_out-ex_h2o);
V_m=22.4;

% calculated rates--------------------------------------------------
---

rco2 = (((F_a_in+F_o2_in).*60)./V_m).*(y_co2_out.*Ra_inert-y_co2_in);
ro2  = (((F_a_in+F_o2_in).*60)./V_m).*(y_o2_out.*Ra_inert-y_o2_in);
rs_ox = (rco2+rx_ext)./Vol; % rs oxidative in c-mol/h/l, for comparison
rx   = rx_ext % estimation comes from external measurement

xm=[r_s; r_o2; r_co2];
E=[+1 0 +1 +1;gamma_s gamma_o2 gamma_x 0];
Em=[+1,0,+1;gamma_s,gamma_o2,0]; % First row C balance 2nd row DR
balance
Ec =[1; gamma_x]; % Biomass Estimation

e_s=0.03;
e_o2=0.03;
e_co2=0.03;

Xi=[e_s 0 0;0 e_o2 0 ;0 0 e_co2];

Ec_star=(inv(Ec'*Ec))*Ec';

R=Em-Ec*Ec_star*Em;
[U,S,V]=svd(R);
Sconv=[1 0];
C=Sconv*S;
K=C*S'*U';
Rred=K*R;
eps = Rred * xm;
sai = diag(diag(xm * xm' * Xi * Xi'));
Phi = Rred * sai *Rred';
delta = (sai*Rred'*inv(Phi) * Rred)* xm;
xmbest=xm-delta;calc
xcbest = -Ec_star*Em*xmbest;
h = eps' * inv(Phi) * eps;

calc.(Fexn).rS_rec(j,2)=xmbest(1);
calc.(Fexn).OUR_rec(j,2)=xmbest(2);
```

```
calc.(Fexn).CER_rec(j,2)=xmbest(3);
calc.(Fexn).rX_rec(j,2)=xcbest;
calc.(Fexn).h_value(j,2)=h;

end
```

5.5 DATA RECONCILIATION FOR K2S1_V2 SOFTSENSOR

```
function [ output ] = K2S1_v2(F_s, c_s, M_s, F_a_in, F_o2_in, y_o2_out,
y_co2_out, y_wet, y_o2_in_air, y_co2_in_air, gamma_s, gamma_o2,
gamma_x, Vol, rx_ext, e_x, e_o2, e_co2, e_s)

% inputs:-------------------------------------------------------------
----

% calculated variables (internally)-----------------------------------
----

y_o2_in=(F_a_in.*y_o2_in_air+F_o2_in)./(F_a_in+F_o2_in);
y_co2_in=(F_a_in.*y_co2_in_air)./(F_a_in+F_o2_in);

ex_h2o=1-(y_wet./y_o2_in_air);
Ra_inert = (1-y_o2_in_air-y_co2_in_air)./(1-y_o2_out-y_co2_out-ex_h2o);
V_m=22.4;

% calculated rates----------------------------------------------------
----

rco2 = (((F_a_in+F_o2_in).*60)./V_m).*(y_co2_out.*Ra_inert-y_co2_in);
ro2  = (((F_a_in+F_o2_in).*60)./V_m).*(y_o2_out.*Ra_inert-y_o2_in);
rs_ox = (rco2+rx_ext)./Vol; % rs oxidative in c-mol/h/l, for comparison
rx   = rx_ext % estimation comes from external measurement

xm=[rx; ro2; rco2];
Em=[+1,0,+1;gamma_x,gamma_o2,0]; % input rates: First row C balance 2nd
row DoR balance
Ec =[1; gamma_s]; % estimated rate: oxidative substrate conversion
Estimation

Xi=[e_x 0 0;0 e_o2 0 ;0 0 e_co2];

Ec_star=(inv(Ec'*Ec))*Ec';

R=Em-Ec*Ec_star*Em;
[U,S,V]=svd(R);
Sconv=[1 0];
C=Sconv*S;
K=C*S'*U';
Rred=K*R;
eps = Rred * xm;
sai = diag(diag(xm * xm' * Xi * Xi'));
Phi = Rred * sai *Rred';
delta = (sai*Rred'*inv(Phi) * Rred)* xm;
xmbest=xm-delta;
xcbest = -Ec_star*Em*xmbest;
h = eps' * inv(Phi) * eps;
```

```
%Outputs-----------------------------------------------------------------
---

%h_value-----------------------------------------------------------------
---
h_value=h;

% valumetric reconciled rates---------------------------------------------
---
rx_rec=xmbest(1,1)./Vol;     %rs from C-balance in c-mol/h/l
ro2_rec=xmbest(2,1)./Vol;    %OUR in mol/h/l
rco2_rec=xmbest(3,1)./Vol;   %CER in c-mol/h/l
rs_rec=xcbest./Vol;          %rx from C-balance in c-mol/h/l

% reconciled yields-------------------------------------------------------
---
Yx_s=-rx_rec./rs_rec;
Y_o2_s=ro2_rec./rs_rec;
Y_co2_s=-rco2_rec./rs_rec;
Y_o2_x=-ro2_rec./rx_rec;
Y_co2_x=rco2_rec./rx_rec;

output=[h_value,rs_rec,ro2_rec,rco2_rec,rx_rec,Yx_s,Y_o2_s,Y_co2_s,Y_o2
_x,Y_co2_x,rs_ox];

end
```

6 REFERENCES

Aehle, M., R. Simutis and A. Lubbert (2010). "Comparison of viable cell concentration estimation methods for a mammalian cell cultivation process." Cytotechnology **62**(5): 413-422.

Ami, D., A. Natalello, G. Taylor, G. Tonon and S. M. Doglia (2006). "Structural analysis of protein inclusion bodies by Fourier transform infrared microspectroscopy." Biochimica et Biophysica Acta (BBA)-Proteins and Proteomics **1764**(4): 793-799.

Babaeipour, V., S. A. Shojaosadati, R. Khalilzadeh, N. Maghsoudi and F. Tabandeh (2008). "A proposed feeding strategy for the overproduction of recombinant proteins in Escherichia coli." Biotechnology and applied biochemistry **49**(2): 141-147.

Brillmann, M. (2015). "Controlled physiological dynamics for quantification of physiological capacities and influencing inclusion body properties & the specific product titer."

Calo-Fernández, B. and J. Martínez-Hurtado (2012). "Biosimilars: Company Strategies to Capture Value from the Biologics Market." Pharmaceuticals **5**(12): 1393.

de Groot, N. S. and S. Ventura (2006). "Effect of temperature on protein quality in bacterial inclusion bodies." FEBS letters **580**(27): 6471-6476.

Dürauer, A., K. Ahrer, W. Kaar, C. Achmüller, W. Sprinzl, S. Mayer, B. Auer, A. Jungbauer and R. Hahn (2010). "Npro autoprotease fusion technology: development, characteristics, and influential factors." Separation Science and Technology **45**(15): 2194-2209.

Gottschalk, U. (2008). "Bioseparation in antibody manufacturing: the good, the bad and the ugly." Biotechnol Prog **24**(3): 496-503.

ICH (2009). Pharmaceutical Development Q8(R2). International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use, November.

Jensen, E. B. and S. Carlsen (1990). "Production of Recombinant Human Growth-Hormone in Escherichia-Coli - Expression of Different Precursors and Physiological-Effects of Glucose, Acetate, and Salts." Biotechnology and Bioengineering **36**(1): 1-11.

Jenzsch, M., S. Gnoth, M. Beck, M. Kleinschmidt, R. Simutis and A. Lubbert (2006). "Open-loop control of the biomass concentration within the growth phase of recombinant protein production processes." J Biotechnol **127**(1): 84-94.

Jobe, A. M., C. Herwig, M. Surzyn, B. Walker, I. Marison and U. von Stockar (2003). "Generally applicable fed-batch culture concept based on the detection of metabolic state by on-line balancing." Biotechnology and Bioengineering **82**(6): 627-639.

Kavanagh, J. and G. Barton (2008). "Productivity improvement of recombinant Escherichia coli fermentation via robust optimization." Bioprocess and biosystems engineering **31**(2): 137-143.

Levisauskas, D., V. Galvanauskas, S. Henrich, K. Wilhelm, N. Volk and A. Lübbert (2003). "Model-based optimization of viral capsid protein production in fed-batch culture of recombinant Escherichia coli." Bioprocess and Biosystems Engineering **25**(4): 255-262.

Lionberger, R. A., S. L. Lee, L. Lee, A. Raw and X. Y. Lawrence (2008). "Quality by design: concepts for ANDAs." The AAPS journal **10**(2): 268-276.

Luo, J., M. Leeman, A. Ballagi, A. Elfwing, Z. Su, J.-C. Janson and K.-G. Wahlund (2006). "Size characterization of green fluorescent protein inclusion bodies in E. coli using asymmetrical flow field-flow fractionation–multi-angle light scattering." Journal of Chromatography A **1120**(1-2): 158-164.

Margreiter, G., P. Messner, K. D. Caldwell and K. Bayer (2008). "Size characterization of inclusion bodies by sedimentation field-flow fractionation." Journal of biotechnology **138**(3-4): 67-73.

Margreiter, G., M. Schwanninger, K. Bayer and C. Obinger (2008). "Impact of different cultivation and induction regimes on the structure of cytosolic inclusion bodies of TEM1-β-lactamase." Biotechnology journal **3**(9-10): 1245-1255.

Pan, S., M. Zelger, A. Jungbauer and R. Hahn (2014). "Integrated continuous dissolution, refolding and tag removal of fusion proteins from inclusion bodies in a tubular reactor." Journal of biotechnology **185**: 39-50.

Ramalingam, S., P. Gautam, K. J. Mukherjee and G. Jayaraman (2007). "Effects of post-induction feed strategies on secretory production of recombinant streptokinase in Escherichia coli." Biochemical Engineering Journal **33**(1): 34-41.

Reichelt, W., M. Brillmann, P. Thurold and C. Herwig (2016). "Substrate accumulation in induced bioprocesses caused by declining physiological capacities."

Reichelt, W. N., M. Brillmann, P. Thurrold, P. Keil, J. Fricke and C. Herwig (2017). "Physiological capacities decline during induced bioprocesses leading to substrate accumulation." Biotechnology journal **12**(7).

Reichelt, W. N., A. Kaineder, M. Brillmann, L. Neutsch, A. Taschauer, H. Lohninger and C. Herwig (2017). "High throughput inclusion body sizing: Nano particle tracking analysis." Biotechnology journal **12**(6).

Reichelt, W. N., P. Thurrold, M. Brillmann, J. Kager, J. Fricke and C. Herwig (2016). "Generic biomass estimation methods targeting physiologic process control in induced bacterial cultures." Engineering in Life Sciences **16**(8): 720-730.

Sagmeister, P., C. Schimek, A. Meitz, C. Herwig and O. Spadiut (2014). "Tunable recombinant protein expression with E. coli in a mixed-feed environment." Applied microbiology and biotechnology **98**(7): 2937-2945.

Sagmeister, P., P. Wechselberger, M. Jazini, A. Meitz, T. Langemann and C. Herwig (2013). "Soft sensor assisted dynamic bioprocess control: Efficient tools for bioprocess development." Chemical Engineering Science **96**: 190-198.

Sandén, A. M., I. Prytz, I. Tubulekas, C. Förberg, H. Le, A. Hektor, P. Neubauer, Z. Pragai, C. Harwood and A. Ward (2003). "Limiting factors in Escherichia coli fed-batch production of recombinant proteins." Biotechnology and Bioengineering **81**(2): 158-166.

Schaepe, S., A. Kuprijanov, R. Simutis and A. Lubbert (2014). "Avoiding overfeeding in high cell density fed-batch cultures of E. coli during the production of heterologous proteins." Journal of Biotechnology **192**: 146-153.

Swartz, J. R. (2001). "Advances in Escherichia coli production of therapeutic proteins." Curr Opin Biotechnol **12**(2): 195-201.

Upadhyay, A. K., A. Murmu, A. Singh and A. K. Panda (2012). "Kinetics of inclusion body formation and its correlation with the characteristics of protein aggregates in Escherichia coli." PLoS One **7**(3): e33951.

Walsh, G. (2014). "Biopharmaceutical benchmarks 2014." Nat Biotechnol **32**(10): 992-1000.

Wechselberger, P., P. Sagmeister, H. Engelking, T. Schmidt, J. Wenger and C. Herwig (2012). "Efficient feeding profile optimization for recombinant protein production using physiological information." Bioprocess and biosystems engineering **35**(9): 1637-1649.

Wechselberger, P., P. Sagmeister and C. Herwig (2013). "Real-time estimation of biomass and specific growth rate in physiologically variable recombinant fed-batch processes." Bioprocess and biosystems engineering **36**(9): 1205-1218.

Williams, D. C., R. M. Van Frank, W. L. Muth and J. P. Burnett (1982). "Cytoplasmic inclusion bodies in Escherichia coli producing biosynthetic human insulin proteins." Science **215**(4533): 687-689.