Universität für Bodenkultur Wien
University of Natural Resources
and Life Sciences, Vienna

Master Thesis

Intensification of experimental design for *Escherichia coli* fed-batch fermentations

Submitted by

Lina VRANITZKY, BSc

in the framework of the Master programme

Biotechnology

in partial fulfilment of the requirements for the academic degree

Diplom-Ingenieur

Vienna, October 2020

Supervisor:

Univ.Prof. DI Dr. Gerald Striedner
Institute of Bioprocess Science and Engineering
Department of Biotechnology

EIDESSTATTLICHE ERKLÄRUNG

Ich erkläre eidesstattlich, dass ich die vorliegende Arbeit selbstständig verfasst, andere als die angegebenen Quellen/Hilfsmittel nicht benutzt, und die den benutzten Quellen wörtlich und inhaltlich entnommenen Stellen als solche kenntlich gemacht habe. Diese schriftliche Arbeit wurde noch an keiner Stelle vorgelegt.

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources, and that I have explicitly marked all material which has been quoted either literally or by content from the used sources. This written work has not yet been submitted anywhere else.

Wien/Vienna, 09.11.2020 Lina Vranitzky

DANKSAGUNG

Diese Arbeit entstand in der Arbeitsgruppe Mikrobielle Fermentation am Department für Biotechnologie an der Universität für Bodenkultur Wien.

In erster Linie gilt mein Dank meinem direkten Betreuer Benjamin Bayer, der mir nicht nur Fermentieren gelernt hat, sondern auch gezeigt hat wie man eine stressige Zeit mit Humor und Sarkasmus bewältigen kann. Dank seiner strukturierten Planung, seiner Hilfsbereitschaft bei kleinen Hürden und seiner fachlichen Kompetenz wurde mir eine lehrreiche, spannende Zeit während dem praktischen Teil meiner Arbeit ermöglicht.

Weiters möchte ich mich bei dem gesamten Team von Novasign bedanken, Mark Dürkop, Roger Dalmau Diaz und Maximilian Krippl, die mir einen Arbeitsplatz in ihrem Office zur Verfügung stellten und für Fragen und Hilfestellungen immer ein offenes Ohr hatten.

Herzlich bedanken möchte ich mich an dieser Stelle auch bei meinem Betreuer Gerald Striedner, der es mir schon ermöglichte meine Bachelorarbeit in seiner Arbeitsgruppe Mikrobielle Fermentation zu absolvieren.

Abschließend, aber keinesfalls letztrangig, möchte ich meiner Familie, meinen engsten Freunden und meinem Partner für ihre emotionale Unterstützung während meiner gesamten Studienzeit danken. Besonders möchte ich hier meinen Vater, Robert Vranitzky, erwähnen, der in Studienfragen immer meine erste Anlaufstelle war und mich bei jeder finanziellen als auch prüfungsbedingten Hürde unterstützt hat.

ACKNOWLEDGEMENT

This work was developed in the working group Microbial Fermentation at the Department of Biotechnology at the University of Natural Resources and Applied Life Sciences, Vienna.

First and foremost, my thanks go to my direct supervisor Benjamin Bayer, who not only taught me how to ferment but also showed me how to overcome a stressful time with humor and sarcasm. Thanks to his structured planning, his helpfulness with small hurdles and his professional competence, I gained an instructive, exciting time during the practical part of my work.

Furthermore, I want to thank the whole team of Novasign, Mark Dürkop, Roger Dalmau Diaz and Maximilian Krippl, who provided me a workplace in their office and have always been open to questions and assistance.

I would also like to take this opportunity to thank my supervisor Gerald Striedner, who has already made it possible for me to successfully complete my bachelor thesis in his working group Microbial Fermentation.

To conclude, last but not least, I want to thank my family, my closest friends and my partner for their emotional support throughout my studies. I would like to make particular mention of my father Robert Vranitzky, who has always been my first and main contact in study matters and who has supported me in every financial and exam-related hurdle.

ABSTRACT

To establish a new bioprocess in the biopharmaceutical industry is a time-consuming and expensive procedure, especially in the field of upstream processing. In order to determine the influences of various critical process parameters on the critical quality attributes in proximity to the already established optimum of a biopharmaceutical process, the statistical design of experiments is commonly used. To generate this demanded process knowledge, guaranteeing the best possible and stable quality output of the product, a so-called design space is set up by these critical process parameters. In this design space, different critical process parameter combinations are characterized and their influence on the critical quality attributes is evaluated. With respect to the chosen number of critical process parameters and levels for a full factorial process characterization and until the critical process parameters enabling the optimal process output are found, the number of required experiments can rapidly exceed a manageable amount. To drastically reduce the needed number of experiments, an alternative approach to classical design of experiments was developed, i.e., intensified design of experiments, thereby accelerating process characterization. During these experiments, intra-experimental process parameter shifts are carried out to characterize more than one critical process parameter combination setpoint per experiment.

This work aimed to investigate the usability of an intensified design of experiments approach for *Escherichia coli* fed-batch fermentations. By means of a static full factorial design of experiments with the same design space it was examined whether process comparability is attained. Based on the data received, a hybrid model to predict the cell dry mass concentration of all static fed-batch fermentations in the chosen design space was developed. Therefore, the following points were investigated:

- A three-dimensional design space was completely characterized by performing intensified *Escherichia coli* fed-batch fermentations in a 20-L stainless steel reactor.
- The possibility of an occurring irreversible memory effect of the cells due to the critical process parameter shifts was examined, which would highly complicate the applicability of intensified design of experiments.
- Consecutively, the intensified design of experiments data were used to train a hybrid model. To compare the performance of this hybrid model, a previously developed hybrid model based on a historical data set of the same design space, characterized by static fed-batch fermentations, was used.

It has been shown that, by using our set up, no memory effect on the cells emerged. Further, by using intensified design of experiments in combination with hybrid modeling, it is possible to accurately predict the cell dry mass concentration of the complete static design space with an error of 5.31% (±

RMSE 1.27 g/L), while the static hybrid model performed with an error of 4.24% (± RMSE 1.10 g/L). Hereby, a comparable model performance was maintained while the required number of experiments was reduced by 66%, highlighting the advantage of using intensified design of experiments for process characterization.

ZUSAMMENFASSUNG

Die Etablierung eines neuen Bioprozesses in der pharmazeutischen Industrie ist ein zeitaufwendiges und teures Verfahren, insbesondere im Bereich des Up-stream. Um die Einflüsse verschiedener kritischer Prozessparameter auf die kritischen Qualitätsmerkmale in der Nähe des bereits etablierten Optimums in einem biopharmazeutischen Prozess zu bestimmen, wird üblicherweise die statistische Versuchsplanung eines design of experiments verwendet. Um das geforderte Prozesswissen zu generieren und die bestmögliche und stabile Produktqualität zu garantieren wird durch die kritischen Prozessparameter ein sogenannter design space errichtet. In diesem design space werden verschiedene kritische Prozessparamter-Kombinationen charakterisiert und ihr Einfluss auf die kritischen Qualitätsmerkmale evaluiert. Für eine vollfaktorielle Prozesscharakterisierung, unter Berücksichtigung der Anzahl der kritischen Prozessparameter und Ebenen, und die Bestimmung der kritischen Prozessparameter, welche einen optimalen Prozessoutput ermöglichen, wird oft eine hohe Anzahl an Experimenten benötigt. Um die benötigte Anzahl an Experimenten drastisch zu senken wurde ein alternativer Ansatz zu einem klassischen design of experiments entwickelt, ein intensiviertes design of experiments, wodurch die Prozesscharakterisierung beschleunigt wird. Während dieser Experimente werden Prozessparameteränderungen durchgeführt um mehr als eine Prozessparameterkombination pro Experiment zu charakterisieren.

Ziel dieser Arbeit war es, die Anwendbarkeit eines intensivierten design of experiments Ansatzes für fed-batch Fermentationen mit *Escherichia coli* zu untersuchen. Mittels eines statischen vollfaktoriellen design of experiments desselben design spaces wurde untersucht, ob eine Prozessvergleichbarkeit vorliegt. Basierend auf den erhaltenen Daten wurde versuchsweise ein Hybridmodell entwickelt, um die Biomassekonzentration aller statischen fed-batch Fermentationen im gewählten design space hervorzusagen. Folgende Punkte wurden untersucht:

- Ein dreidimensionaler design space wurde vollständig durch die Durchführung intensivierter *Escherichia coli* fed-batch Fermentationen in einem 20 L-Edelstahlreaktor charakterisiert.
- Die Möglichkeit eines auftretenden irreversiblen Memory-Effekt der Zellen aufgrund der kritischen Prozessparameteränderungen wurde untersucht, da es die Anwendbarkeit eines intensivierten design of experiments stark erschweren würde.
- Anschließend wurden die intensivierten design of experiments Daten verwendet, um ein Hybridmodel zu trainieren. Um die Leistung dieses Hybridmodels zu vergleichen, wurde ein zuvor entwickeltes Hybridmodel verwendet, das auf einem historischen Datensatz desselben design spaces mit statischen fed-batch Fermentationen basiert

Es wurde gezeigt, dass bei Verwendung des intensivierten design of experiments kein Memory-Effekt der Zellen auftrat. Durch Anwendung eines intensivierten design of experiments in Kombination mit einem Hybridmodel ist es möglich, die Biomassekonzentration des gesamten statischen design space mit einem Fehler von 5.31% (± RMSE 1.27 g/L) genau vorherzusagen, während das statische Hybridmodel mit einem Fehler von 4.24% (± RMSE 1.10 g/L) aufzeigt. Hierbei wurde eine vergleichbare Modelleistung beibehalten und die erforderliche Anzahl von Experimenten um 66% reduziert, was den Vorteil der Verwendung von intensivierten design of experiments zu Prozesscharakterisierung hervorhebt.

CONTENT

1 INTRODUCTION

This chapter provides an introduction into bioprocessing and recombinant protein production and elaborates on process analytical technology (PAT) and quality by design (QbD). Subsequently, the design of experiments (DoE) approach is described in the context of fermentation operation modes in the pharmaceutical industry. Further aspects that are treated in this chapter are the different host systems that are applied and *Escherichia coli* (*E. coli*) as an expression system that is most widely used.

Finally, the characteristics of process modelling are provided, and some light was shed on hybrid modelling in particular, since it was used as part of this thesis for the validation of the generated intensified DoE (iDoE) process data

1.1 Bioprocessing and recombinant protein production

Biotechnology has gained great relevance when it comes to pharmaceutical and medical applications. Biotechnological processes are extremely complex, e.g., cells and microorganisms which form the biological system need to be cultivated under accordingly adjusted and defined conditions. To enable high productivity of the often sensitive biological systems, a specific physical (e.g., temperature, DO) and chemical (e.g., substrates, products) environment have to be provided (Meyer & Beyeler, 1984). The main focus lies on the viability, productivity of the cells and the reproductivity of the process. The interaction between these three compartments needs to be understood and subsequently controlled, i.e., detailed knowledge of the overall-system to drive the biochemical reaction network in the correct direction has to be provided (Scheper et al., 1999).

Especially in the upstream of a biotechnological process, many considerations must be taken into account to provide different perspectives and a comprehensive understanding, enabling the transition from small to larger scales (Marques, Cabral, & Fernandes, 2010). Usually, the development time for biopharmaceutical products is about 5 to 10 years from the first set up to a large-scale production, concluding, it is a long-lasting and expensive procedure.

There is a distinction between product and process development. The product development refers to the discovery of the product and the initial steps, which are obtained through research and are focussed on the mode of action. On the other hand, with goals such as persistent product quality and economic targets, e.g., cost-effective process flowsheets (Hassan et al., 2015), the manufacturability of the product is processed in the process development. This concurrent engineering of product and process development phases can be carried out with a certain degree of overlapping, e.g., through frequent information exchange of upstream and downstream activity (Krishnan, Eppinger, & Whitney, 1997).

The overlap of both phases continues form screening of product and process parameters to the scale-up phase until the full scale manufacturing is reached. Typically, the strategy used to set up a new bioprocess is trial and error based (Neubauer et al., 2013).

To enhance the process and early product development phase, structuring these approaches, engineering strategies are applied which lead to drastic paradigm shifts. Strategies typically used in process development are so-called statistical DoE, process modelling and control tools. Using these strategies, a systematically approach is provided to facilitate process characterization, offering a higher success rate and the opportunity to faster find the optimal process parameters (Kumar, Bhalla, & Rathore, 2014). In future, (semi-)automated platforms and models can support bioprocess development by increasing the product quality and productivity of the overall process (Alzari et al., 2006).

For the recombinant protein itself, the early development phase involves essential decisions for the final protein production process (Gräslund et al., 2008). A high number of variables need to be characterized and evaluated during this phase, e.g., host system and critical process parameters (CPPs) for the growth and production phase (Rohe, Venkanna, Kleine, Freudl, & Oldiges, 2012). In order to decrease the number of experiments that have to be carried out for process characterization, statistical approaches and strategies, such as DoE, are applied. Thereby, the maximum information for the accurate range and the interaction between the investigated CPPs are discovered (Neubauer et al., 2013).

When it comes to therapeutic protein production, the highest priority is to properly control the process to minimize process variability. Therefore, strategies for direct and targeted product quality control and assurance are developed. A detailed process characterization results in an all-embracing process understanding, identifying the influencing factors on the product yield and quality. As a result, the batch variability is lowered, leading to fewer rejected batches (Sommeregger et al., 2017).

1.2 Fermentation operation modes

The fermentation process can be performed in different modes of operation: batch, continuous and fed-batch fermentation. As the simplest and most common way, a batch mode is applied. For the manufacturing process, the production host (e.g., bacteria), the substrates and supplementary nutrients are inoculated in the reactor with controlled set points, e.g., temperature, pH and DO (Meyer & Beyeler, 1984). The process is finished when a certain CDM concentration is reached or the substrate is consumed, which results in stopped cell growth. An advantage of performing in batch mode is the ease of operation and low risk of contamination (Yang & Sha, 2020) . A disadvantage of batch processes

is that comparatively low cell densities are reached, which is not preferred in microbial fermentations. They are time-consuming due to the sterilization period of the reactor between the batches and the requirement to further maintain complete sterility during processing (Amelio et al., 2016; Yang & Sha, 2020). To avoid the expensive rejection of time-consuming batches it is of great importance to execute reproducible processes and in a predefined manner to enable the guarantee of safety (Gnoth, Jenzsch, Simutis, & Lübbert, 2008).

To overcome the disadvantages of the batch mode production, the superior production approach of fed-batch fermentation, e.g., for microbial processes and mammalian cell culture in industrial applications was developed. It is defined as a technique where one or more nutrient supplements are added to the bioreactor during the process. The cultured product remains in the vessel until the end of the run, consequently, no CDM is discharged (Yamanè & Shimizu, 2005). During a fed-batch fermentation the changing nutrient concentrations affect the productivity and thereby the yield of the desired product. By controlling the nutrient supply cell growth is increased or inhibited. The respectively needed amount of nutrients are added to the reactor during the process in a timely manner. Until finishing the production of e.g., the particular recombinant product no cells are removed. Various by-products like primary or secondary metabolites, proteins and biopolymers are produced using fed-batch fermentations (Lee, Lee, Park, & Middelberg, 1999). Often theoretical mathematical models are compared to experimental results to gain a better understanding of the process and to facilitate developing of a cultivation method that allows a cost-effective production for high product yield and high productivity of the production host (Yamanè & Shimizu, 2005).

Another approach is a continuous fermentation system, which offers important economic advantages and significantly improved rates as opposed to traditional systems. Steady-state operation, high volumetric productivity, streamlined process flow and low capital cost of continuous processing enable the slow introduction of this new process methodology into biotech industries. During the process, the reactor is fed with a continuous substrate flow and a continuous flow of product is transported out of the system. However, the higher productivity does not provide the same high product concentration compared to a batch process. Thereby the genetic stability of the host systems are insufficient the approach for recombinant protein production is still under development (Verbelen, De Schutter, Delvaux, Verstrepen, & Delvaux, 2006; Warikoo et al., 2012).

In the field of biotechnological industry, shortest development times and cost control on basis of strict quality and regulatory requirements are the driving force for development (Varsakelis, Dessoy, von Stosch, & Pysik, 2019). These requirements resulted in implementing new technologies like process modelling and process simulation, where mathematical equations describe physical systems (Brass,

Hoeks, & Rohner, 1997). Mechanistic, empirical, or hybrid models have been developed over the years and are rising more and more in the importance of process characterization (Varsakelis et al., 2019).

Besides the optimal manufacturing process, the choice of the ideal host system must be made to guarantee high yields and high quality of the recombinant products, as well as to guarantee the correct expression of the protein of interest (POI), e.g., the glycosylation pattern (Waegeman & Soetaert, 2011).

1.3 Host systems

There is no universal expression system for the production of recombinant proteins. Every antibody, enzyme, or other protein has its expression difficulties and limitations, e.g., different codon usage, the need for chaperones or post-translational modifications such as glycosylation or disulphide bridges (Pourmir & Johannes, 2012). Solubility, correct folding, stability and the size of the protein are examples of decisive factors that affect yield and biological activity of recombinant proteins. The unique and specific amino acid sequences are reasons for the difficulties, i.e., the optimal expression system for one antibody might not be suitable for another antibody. Which system is selected depends on many factors, such as the molecule being expressed (e.g., IgG or Fab), the antibody itself, the required quantity and the quality of the final product (Verma, Boleti, & George, 1998). The choice of the ideally suited system is therefore challenging due to the purpose and physicochemical characteristics of the POI. Existing expression systems are constantly improved and optimised while new approaches are developed to satisfy the demands of producing complex proteins (Gomes & Byregowda, 2016).

Typical host systems for the expression of recombinant products are bacteria, yeast, plants, insects and mammalian expression systems. Amongst others, they have been developed with the purpose of producing a functional molecule at adequate cost and minimal effort. To guarantee appropriate protein expression, the efficient and correct transcription and translation of the cloned gene must be provided (Waegeman & Soetaert, 2011).

Mammalian cells have become the dominant system for the production of recombinant proteins. Approximately 60-70% of all recombinant pharmaceuticals are produced in mammalian cells (Omasa, Onitsuka, & Kim, 2010). The most frequently used production host are immortalized Chinese hamster ovary (CHO) cells. Their abilities to proper protein folding and assembling are essential for the expression of some proteins and lead to higher quality and efficiency for recombinant proteins that require post-translational modifications such as antibodies, hormones and cytokines (Pourmir & Johannes, 2012). They can be considered as the host of choice in terms of correct synthesis and human-like modifications such as glycosylation patterns, which do not occur in lower organisms, e.g., bacteria and yeast. The achievement of high product titers are due to the long-term experience and knowledge

about CHO cells (F. M. Wurm, 2004). Disadvantages of the mammalian production system are high costs and a low specific growth rate, which causes long cultivation times and the time-consuming construction and selection of a cell line (Omasa et al., 2010).

Unlike mammalian expression systems, yeast has a high growth rate and is cultured on simple media (F. M. Wurm, 2004). As microorganism, a eukaryote yeast, e.g., *Saccharomyces cerevisiae*, combines the advantages of cost-efficiency, fast and technical feasibility similar to bacteria as well as high-density cell cultures reached in bioreactors. Furthermore, it possesses pathways for advanced heterologous protein folding. For this reason, the yeast system is industrially important for the expression of therapeutic proteins in industrial-scale fermentations. Although this host organism is related to N-linked glycosylation patterns which differ from higher eukaryotes (Ferrer-Miralles, Domingo-Espín, Corchero, Vázquez, & Villaverde, 2009). Yeasts are not an optimal host for large-scale production due to their high requirements especially in technical equipment (Mattanovich et al., 2012).

Using insect cells as an expression system has the benefit that most of the post-translational modifications, which are present in higher eukaryotes can be conducted. Over the last few years, insect cells are emphasized as an attractive alternative choice for the expression of recombinant molecules. The most popular expression system in insect cells is the baculovirus system, with the advantage that high amounts of functionally active proteins of interest can be produced. The genetic modification and the screening for cell lines is simple compared to mammalian cell lines. Moreover, it only infects invertebrates and therefore has a highly restricted host range (Verma et al., 1998).

The bacterial expression system is most widely used for primary cloning, genetic modification and small-scale production for research purposes and the production of recombinant proteins. Due to the long historical development of microbial physiology and molecular genetics, which is mainly based on this species, the first-choice bacterial microorganism is *E. coli* (Ferrer-Miralles et al., 2009).

1.4 *E. coli* as an expression system

E. coli has been established as the most frequently used bacterial host for producing recombinant proteins. With nearly 30% of all marketed recombinant therapeutics, a significant amount of commercial therapeutic proteins are produced by *E. coli* (Huang, Lin, & Yang, 2012). With well-characterized genetics and the public available genome sequences, *E. coli* offers a wide range of cloning vectors and mutant strains. The organism quickly reaches high cell density cultivated in inexpensive media and a short cultivation time compared to a mammalian system is required (Verma et al., 1998). With the possibility of growing facultatively anaerobic and the easy adaptability to metabolic stress, an extensive variety of compounds have been produced (Theisen & Liao, 2016).

A number of practical problems need to be considered when working with *E. coli* as expression host, e.g., large and complex proteins forming disulfide bridges or post-translational modifications are not suitable as a potential product. Furthermore, overexpressed proteins often lead to the formation of inclusion bodies (IBs) which require an additional and complicated denaturation and refolding process to achieve functionality (Ni & Chen, 2009). As well as the accumulation of lipopolysaccharides (LPS), which are pyrogenic for humans, which provides a laborious but solvable challenge for the downstream process when using therapeutic proteins expressed in *E. coli* (Terpe, 2006).

Nevertheless, the detailed and comprehensive understanding of that organism and its regulations make *E. coli* an appropriate model system for various mathematical prediction approaches, e.g., process modelling (Opalka et al., 2010).

1.5 Process analytical technology (PAT) and Quality by Design (QbD)

To reach higher health and economic benefits in pharmaceutical manufacturing, the U.S. federal development agency (FDA) first presented the PAT initiative for modern and advanced process control. Herein, a risk-based approach is suggested, providing a key element of the pharmaceutical current good manufacturing practise (cGMP) (FDA, 2006). With the objective to build a dynamic, adaptable QbD approach, intending to provide global manufacturing processes. As opposed to statically testing the quality of the final product, the QbD method aims to an approach where the overall process is considered, guaranteeing a more stable and robust product quality output (FDA, 2018).

By identifying the critical quality attributes (CQA), a robust control strategy for the process itself can be designed. To ensure a consistent process performance during the whole time, continuous on-line monitoring and dynamic process control plays an important role. Here, PAT comes into play. One main goal of the PAT framework is to design well-understood processes by utilizing different tools to enable a comprehensive measurement base, i.e., in a well-understood process, all causes of variability can be identified and explained (FDA, 2004).

In order to ensure a robust and uniform output of the process, the development of PAT led to a further improvement in the acquisition of process data and associated measurement tools. The FDA defined three categories of measurement approaches:

- Off-line: After a sample was taken from the bioprocess, the measurement is carried out independently from time and place, e.g., determination of the main carbon source glucose and by-products like acetate (Peuker et al., 2004).

- At-line: A sample is taken from the process stream, isolated and analysed immediately, in close proximity to the bioprocess, e.g., determination of cell dry mass (CDM) and product concentration (Rao, Moreira, & Brorson, 2009; Sommeregger et al., 2017).
- On-line and In-line: Many parameters can be monitored online, such as pH and dissolved oxygen (DO), without diverting a sample from the manufacturing process. This measurement can be invasive (In-line) or non-invasive though connected to the process (On-line) (Abu-Absi et al., 2011). A continuous process control is therefore permitted.

In order to achieve such a PAT and QbD implementation, process development, characterization and optimization as well as automation are carried out in different stages. The development of a new process starts with the design phase, in which the CQAs and the CPPs that affect the CQAs, e.g., product quality, are determined. Based on the chosen CPPs, a design space is created which is characterized to gain process knowledge and later on to find the process optimum (Finkler & Krummen, 2016; Rathore, Bhambure, & Ghare, 2010).

The utilization of a design space to investigate the impact of materials used, process and ecological conditions and manufacturing on the CQAs enables accurate and reliable development of tools to guarantee product quality requirements. The final goal of utilizing this framework is that a certain product quality can be predicted and assured at the end of a production process (Garcia, Cook, & Nosal, 2008). Figure 1 (p. 8) shows a schematic display of a bioprocess with a state-of-the-art quality by testing approach and a process where QbD tools are considered.

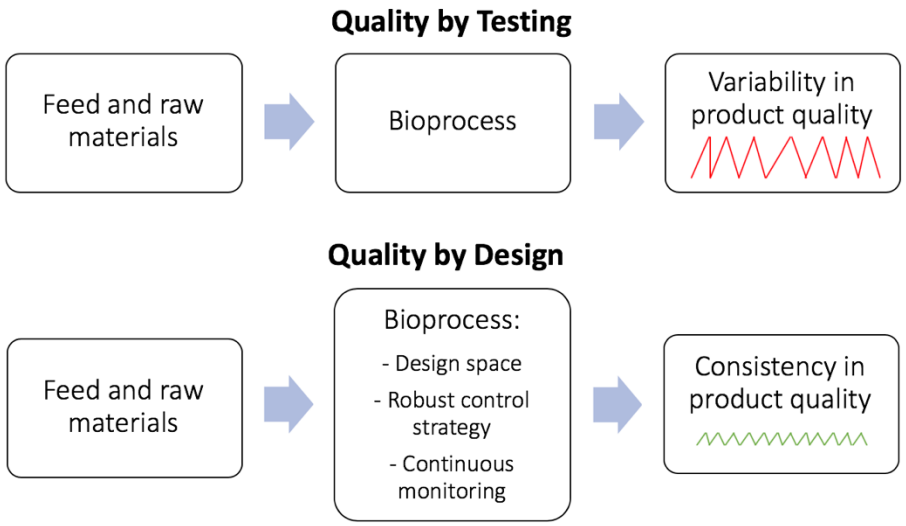

Figure 1: Difference between the product output of Quality by Testing and a QbD approach. In a static process where the quality is only tested in the final product a higher variability in product quality, which is shown as the red high-jagged line, occurs. By applying a QbD approach to a bioprocess, the whole process is considered and results in consistency in product quality (green less-jagged line).

Figure 1 illustrates the high variability in product quality by only considering the final product. To overcome these fluctuations in quality, a QbD approach is used, e.g., continuous monitoring during the bioprocess, and consistency in product quality is obtained.

The application of PAT in upstream operations such as microbial fermentation processes is gaining more and more attention. With the benefits of better process understanding using on-line and at-line measurements, improved yields and a decrease in energy consumption, the production cycle time can be reduced (Read et al., 2010). To avoid the rejection of batches due to the occurrence of major problems during the process or deviations of the batches from the quality specifications, the implementation of QbD and PAT already needs to take place at the process development level and further continue in the scale-up procedure. As consequence costs can be saved due to the prevention of reprocessing and the lowering of energy consumption and waste rates (Rathore et al., 2010).

In batch and fed-batch processes, e.g., CDM production and product formation have already been monitored on-line (Nickel, Cruz-Bournazou, Wilms, Neubauer, & Knepper, 2017; Wechselberger, Sagmeister, & Herwig, 2013). Thereby, it is possible to rapidly react to occurring deviations and still provide high-quality products within the acceptance criteria. As a result, the time-consuming off-line sampling and the following analytical measurements will be eliminated (Bayer, von Stosch, Melcher, Duerkop, & Striedner, 2020).

1.6 Design of experiments (DoE)

When it comes to structuring and determining the relationship between the CPPs of a pharmaceutical process, statistical DoE is the method of choice (Möller, Kuchemüller, Steinmetz, Koopmann, & Pörtner, 2019). It is a concept where the mathematical relationship between the in- and output variables of a system are investigated. The process can be characterized and optimal conditions can be identified, determining the CPPs that influence the CQAs (Patel, Parmar, & Patel, 2013).

Prior process knowledge and risk management are used to establish an DoE (Yu, 2008). In a fixed set of experiments, different levels and combinations of CPPs are examined. The set of experiments needs to be carried out in a random order to avoid systematic errors, which can occur when performing one experiment after another using similar process setups (Mandenius & Brundin, 2008).

The evaluation of experimental data is commonly carried out with the mathematical model of multiple linear regression (MLR), describing the relationship of input (independent) variables and output (dependent) variables. To correlate the variation of linear or quadratic terms interaction terms can be added to the MLR (Iordache, 2013; Mandenius & Brundin, 2008). Usually response surface models (RSM) are applied, which estimate the process optimum by using quadratic functions and representing the data in a time-resolved manner.

Moreover, in upstream bioprocess development, DoE is a widely used method for the effective understanding of a (bio)process. However, this is often challenging and complex since the CQAs of a microbial fermentation process are sensitive to many CPPs. This characterization is of high importance because the effective manipulation of a system can only be executed by understanding the impact of design and control parameters in a time resolved manner (Bayer, von Stosch, Striedner, & Duerkop, 2020). DoE varies the controllable CPPs and systematically determines their impact. The uncontrollable factors can be distinguished by replication, randomization or blocking (von Stosch, Hamelink, & Oliveira, 2016).

It is important that the impact of all CPPs and factors regarding the process itself, the product and the production host, e.g., *E. coli*, are well understood. The more CPPs shall be tested, the number of experiments increases and becomes very time-consuming, laborious and thereby also affects the economical, cost-intense aspect. Therefore, a key demand of manufacturers is to reduce the duration and the costs of process development (Mercier, 2013).

To overcome the large number of experimental runs of a classic DoE, investigating each combination of each CPP at every level, i.e., full factorial design, an efficient approach is a fractional factorial design. In a fractional factorial design, a subset of experiments are used, proceeding from selected corners in

the experimental space of the full factorial design. Thereby, less costs and effort compared to a full factorial design is required (Toms, Deardon, & Ungrin, 2017). A new promising design of DoE, aiming to reduce the number of experiments, is described by von Stosch and coworkers as iDoE (von Stosch & Willis, 2016).

1.7 Intensified design of experiments (iDoE)

A highly emphasized approach is to reduce the number of experiments, counteracting these long development times, but still gaining the same degree of process understanding by characterizing the same amount of CPP setpoints. This is where iDoE, a new concept in upstream bioprocess development, is applied. In an iDoE, the CPPs during each experiment are varied, for a fixed number of levels, in contrast to the classic DoE with static parameters. Therefore, the number of experiments can be reduced by changing the conditions during an experiment, instead of keeping them constant during the entire process. According to von Stosch 2017, the total number of experiments can be reduced by up to 40% using iDoE (von Stosch & Willis, 2016) and Bayer 2020 even demonstrated a reduction of 66% choosing the right setup (Bayer, Striedner, & Duerkop, 2020). The most important objective of all DoE in general is to gain insights into the process. iDoE additionally gathers information about the dynamic behaviour and the responses of the cells and the process. This understanding leads to the development of an appropriate dynamic process model (von Stosch & Willis, 2016). Herein, to describe this methodology, data of *E. coli* fermentations, performed in fed-batch operation mode, a widely applied operation mode for therapeutic protein production in the pharmaceutical industry (Lee et al., 1999), was used. This and other commonly applied operation modes are described below.

1.8 Process modeling

Commonly, process models are used as a representation and examination of the knowledge about a process. Thereby, they can be applied in terms of process monitoring, optimization and control. Real process data obtained under different operating conditions must be generated, to enable accurate predictions. A valid representation of the process is given when the model predictions match the analytically derived data. Under this condition, the validated model can be used for implementing and improving process control (Schubert, Simutis, Dors, Havlik, & Lübbert, 1994).

Due to the laborious experimental data generation and preparation of fermentations and bioprocesses, modelling is difficult and time-consuming. However, accurate models and computer simulations are necessary to reduce the number of experiments, which are a cost-driving factor in the biopharmaceutical industry (Luli et al., 1999).

Typically, statistic software tools like RMS are used to analyse bioprocess data (Kalil, Maugeri, & Rodrigues, 2000). With this approach, the endpoint values of the investigated process variable, e.g., the product titer, and the impact of the CPPs are examined. However, by only taking the endpoint into account, many important influences during the process are neglected and the impact of parameter deviations or temporarily failures during the process are not considered (Lundstedt et al., 1998).

One commonly used data-driven model technique is an artificial neural network (ANN). Such approaches, where the structure is only determined from data, can be classified as non-parametric models. This unstructured model is also called black-box model. Non-parametric black-box models often do not have reliable extrapolation properties and many experiments to identify the herein used factors without any physical meaning are needed to cover the whole application domain. Therefore, a disadvantage of ANN models is the lack of transparency and poor performance on new data (Luli et al., 1999; Van Can et al., 1998).

Another model approach is the white-box strategy, which is a knowledge driven approach. Testing and improving an accurate model in comparison to experiments classifies the model as a parametric one. Those mechanistic models describe the interaction between process variables by fundamental principles, i.e., the factors used do have a physical meaning. Due to the complexity of living cells and the simplicity of the used equations, it is often challenging to set up reliable white-box models for bioprocesses (Wechselberger et al., 2013).

However, there is a way to benefit from the positive aspects of black-box and white-box modelling approaches. In a so-called hybrid model (semi-parametric), the advantages of both modelling approaches are combined and make up for the respective drawbacks, i.e., utilizing process knowledge and process information from different sources in a combined model structure.

1.9 Hybrid modeling

So far, hybrid models of upstream processes have been described by numerous authors (A. Teixeira et al., 2005; A. P. Teixeira, Alves, Alves, Carrondo, & Oliveira, 2007). Knowledge and data related to the process were used to generate a model. The main focus lies in monitoring, controlling and optimization of a fermentation process (von Stosch, Oliveira, Peres, & Feyo de Azevedo, 2014). Recently hybrid models have gained more significance in downstream processing, e.g., in predicting the flux evolution and duration of cross-flow ultrafiltration processes (Krippl, Dürauer, & Duerkop, 2020).

More specifically, hybrid models combine mechanistic equations with black-box approaches in order to achieve an effective use of available information. Herein, the black-box techniques are applied to estimate the unknown parts of the system. Because of that, the hybrid model is also referred to as grey-

box (Zorzetto, Filho, & Wolf-Maciel, 2000). The application of a hybrid model enables the description of the entire process in a more accurate way. In contrast, the use of solely black-box models allows an improper prediction due to the above-mentioned and often not reliable extrapolation properties. White-box models are too general and do not consider process impacts that are not part of the applied equations. Therefore, by using a time-resolved hybrid model, deviations during the process can be explained and the change of parameters in a running process and the resulting behaviour of the process can be understood. In addition, time-consuming and expensive experiments, which are used for data input, can be reduced, which is a big benefit for manufacturers (Schubert et al., 1994; von Stosch et al., 2016).

2 AIM OF THE WORK

The aim of this thesis was to implement the innovative iDoE method for *E. coli* fed-batch fermentations to accelerate design space characterisation.

Based on a previous full factorial static DoE study of 27 CPP combination setpoints, characterized by fed-batch fermentations with static CPPs, an iDoE of nine experiments with two CPP combination changes during each fermentation was developed. Three different levels of the induction strength, the cultivation temperature and the theoretic target growth rate (controlled by the glucose feed) set up the design space of the static DoE. All CPP combinations of the static DoE were carried out in 9 intensified fermentations. The experiments were performed with an *E. coli HMS174* strain, producing recombinant *hSOD* as the model protein.

The CDM and product titer were determined over the entire course of the fermentation and compared between DoE and iDoE.

To guarantee the applicability of the iDoE setup, the occurrence of a possible memory effect of the cells due to previous CPP setpoints was investigated. Therefore, it was assessed if the starting point and the temporal order of the parameter shifts have an impact on the state of the cell, the variables of interest and the growth kinetic.

Consecutively, the data of the analytical results and the on-line measurements of the iDoE were imported into the *Novasign* software (Novasign GmbH, Vienna, Austria) and an iDoE hybrid model was created and evaluated (tested) with respect to its ability to predict the CDM of the static DoE (test set). To demonstrate the advantage of using intensified experiments for model building, the model performance of the in this thesis developed iDoE hybrid model was compared to the performance of a static hybrid model, developed using all the static experiments (derived from an earlier study).

3 MATERIALS AND METHODS

3.1 Bacterial strain and plasmid

For all experiments the host cell line HMS174(DE3) (F^- recA1 hsdR(r_{K12}^- m_{K12}^+) Rif^R), transformed with a pET11a vector (pET System Manual, 11th edition) was used. As the selectable marker, the gene encoding for ampicillin resistance was chosen. The strain was purchased from *Novagen* (Novagen, Germany) and carries a λ prophage with a gnomically integrated T7 RNA polymerase gene, enabling the expression of the target protein from T7 promoters under the control of a lac UV5 promoter. As the target protein, human Cu/Zn superoxide dismutase (hSOD) (E.C. 1.15.1.1) was chosen, by inserting its coding gene into the plasmid. The highly soluble 32 kDa protein hSOD consists of two homologous monomer subunits with a length of 153-amino acids. It is expressed in the cytoplasm and is non-toxic to the host cell. The induction of the recombinant protein production was started by adding a single pulse of isopropyl-β-D-thiogalactopyranoside (IPTG) (#1043, GERBU Biotechnik, Germany). Every fermentation was started with an inoculum of 1 mL from a deep-frozen working cell bank (WCB) to guarantee standardized starting parameters (Melcher et al., 2015; D. J. Wurm, Hausjell, Ulonska, Herwig, & Spadiut, 2017).

3.2 Cultivation in the bioreactor

3.2.1 Media preparation

A semisynthetic media was calculated to produce 22.5 g CDM in a volume of 4 L in the batch phase. During the feed phase, the additional 337.5 g CDM should be produced in synthetic media by providing 8 L feed media. In the following table (Table 1) the amounts of the ingredients are provided for generating 1 g CDM. The components were dissolved separately and added one by one. To avoid the Maillard reaction, the glucose solution was autoclaved separately and afterward added to the cooled media solution under sterile conditions. The batch media contained yeast extract to increase the initial growth of the cells.

Table 1: Composition of the semisynthetic and synthetic media for batch and fed-batch phase.

Component	Semisynthetic media	Synthetic media
KH_2PO_4	3 g/L	3 g/L
K_2HPO_4*$3H_2O$	4.58 g/L	4.58 g/L
Yeast extract	0.15 g/g CDM	-
Na_3-Citrate*$2H_2O$	0.25 g/g CDM	0.25 g/g CDM

$MgSO_4*7H_2O$	0.10 g/g CDM	0.10 g/g CDM
$CaCl_2*2H_2O$	0.02 g/g CDM	0.02 g/g CDM
trace element solution	50 µL/g CDM	50 µL/g CDM
Antifoam solution (PPG2000)	0.5 mL/L	0.5 mL/L
Glucose*H_2O	3 g/g CDM	3 g/g CDM
NH_4Cl	2.50 g/L	-
$(NH_4)_2SO_4$	2.10 g/L	-

These media compositions provided the required buffer capacity and served as sources of phosphate and potassium. According to the theoretical final amount of 360 g CDM, phosphate salts were calculated. To avoid and suppress foaming, Antifoam solution (Polypropylenglycole 2000, Sandoz) was added to the media to a final concentration of 0.05%. An 12.5% ammonia solution was used for the regulation of the pH value, the supply of nitrogen for growth and protein production.

3.2.2 Cell cultivation conditions

The fermentations were performed in a 20 L computer-controlled bioreactor (MBR, Wetzikon, CH). A total working volume of 12 L was calculated, 4 L batch volume and additionally 8 L in the feed phase. A local measuring and control system (Simatic S7-400), equipped with standard control units (Siemens PS7, Intellution iFIX) was used to measure and calculate the on-line available process parameters. To maintain the pH value at a neutral level of 7.0 ± 0.05, a 12.5% ammonium hydroxide solution (w/w) (MERCK) was added when required. The amount of ammonium was measured and calculated by a pH electrode (Mettler Toledo InPro 3250), calibrated by using commercially available buffer solutions at pH 4.0 and pH 7.0. Supervisory control and data acquisition system (SCADA/Siemens WinCC) was used on an X86 PC, running Windows XP. During the batch phase, the temperature was set to 37°C ± 0.5°C, measured with a Pt100 sensor.

To measure the DO an optical oxygen sensor (Mettler Toledo InPro 6970i) was used. To calibrate the sensor, N_2-gas (0% saturation of pO_2) and inlet air (100% process air) were applied. A level of 30% saturation of DO was stabilized through the stirrer speed (800-1200 rpm) and the aeration rate control, to avoid oxygen limitation. To suppress foaming, 0.5 mL/L of antifoam (PPG 2000, Sigma Aldrich) were added to both, media and feed medium, before starting the cultivation.

Shortly before the inoculation, the working cell bank (stored at -80°C) with an optical density of OD_{600}=1 was thawed and 1 mL was aseptically transferred to 30 mL 0.9% physiological saline solution. This solution was used for the inoculation.

3.2.3 Fed-batch phase

The fed-batch phase was calculated to perform four doubling times (generations) of *E. coli*. The batch phase ended when the culture completely consumed the batch-glucose, entering the stationary phase, with the total CDM of 22.5 g. On the online control system, the end of the batch phase was visible when the DO level rises rapidly due to the glucose exhaustion of the media and thereby decreasing pO_2 consumption. An exponential substrate feed was started at this time point. The pump speed increased according to the exponential growth algorithm (Equation 1), with X as the CDM, μ as the set growth rate and t as time.

$$X_1 = X_0 \times e^{\mu \Delta t} \qquad (1)$$

Thereby, the target growth rate was controlled. The substrate tank was stored on a balance containing a predefined concentration and has a feedback control of weight loss.

The recombinant protein expression was induced after one generation, by a single pulse of different IPTG concentrations to provide the respective amount of IPTG for the already existing CDM, followed by an additional exponential inductor feed (Marisch, Bayer, Cserjan-Puschmann, Luchner, & Striedner, 2013).

3.2.4 Design space

In all experiments, the batch phase was always carried out at 37 °C. The fed-batch phase conditions varied for each experiment, to cover all CPP combinations of the design space. The CPPs and the applied levels, namely the temperature, the target growth rate and the induction strength, used in the DoE and iDoE are listed in Table 2.

Table 2: List of varied CPPs and levels in the DoE and iDoE.

Temperature [°C]	Target growth rate [h^{-1}]	Induction strength [µmol IPTG/g CDM]
30	0.1	0.2
34	0.15	0.5
37	0.2	0.9

The fermentation setpoints of the traditional DoE are presented in Table 3, including the identification number of the respective experiment. Each experiment had static parameter conditions over the four doubling times of the feeding phase, i.e., one setpoint for the temperature, the target growth rate and the IPTG ratio was held constant for the entire fermentation.

Table 3: Parameter setpoints of the full factorial DoE comprising 27 CPP combinations.

Identification number of experiment	Temperature [°C]	Target growth rate [h^{-1}]	Induction strength [µmol IPTG/g CDM]
SOD94	30	0.1	0.2
SOD99	30	0.15	0.2
SOD102	30	0.2	0.2
SOD106	34	0.1	0.2
SOD97	34	0.15	0.2
SOD103	34	0.2	0.2
SOD104	37	0.1	0.2
SOD95	37	0.15	0.2
SOD107	37	0.2	0.2
SOD63/68	30	0.1	0.5
SOD80	30	0.15	0.5
SOD61/66	30	0.2	0.5
SOD96	34	0.1	0.5
SOD101	34	0.15	0.5
SOD105	34	0.2	0.5
SOD47	37	0.1	0.5
SOD53	37	0.15	0.9
SOD58	37	0.2	0.9

SOD79	30	0.1	0.9
SOD60	30	0.15	0.9
SOD59	30	0.2	0.9
SOD98	34	0.1	0.9
SOD93	34	0.15	0.9
SOD100	34	0.2	0.9
SOD42/44/91	37	0.1	0.9
SOD49	37	0.15	0.9
SOD56	37	0.2	0.9

To reduce the characterization time of the presented full factorial design, from 27 fermentations with static process conditions down to 9 dynamic experiments, an iDoE was implemented. Therefore, two intra-experimental shifts were performed to characterize three CPP combinations during each fermentation. After one doubling time, the pulsed induction using IPTG was performed and the exponential inducer feed started. The first CPP combination endured for the second generation of the fed-batch phase. By the end of the second generation, the first out of two parameter shifts was carried out, starting the characterization of the second CPP combination. The second parameter shift, characterizing the third CPP combination, was executed after the third doubling time. Figure 2 gives a schematic overview of an iDoE process.

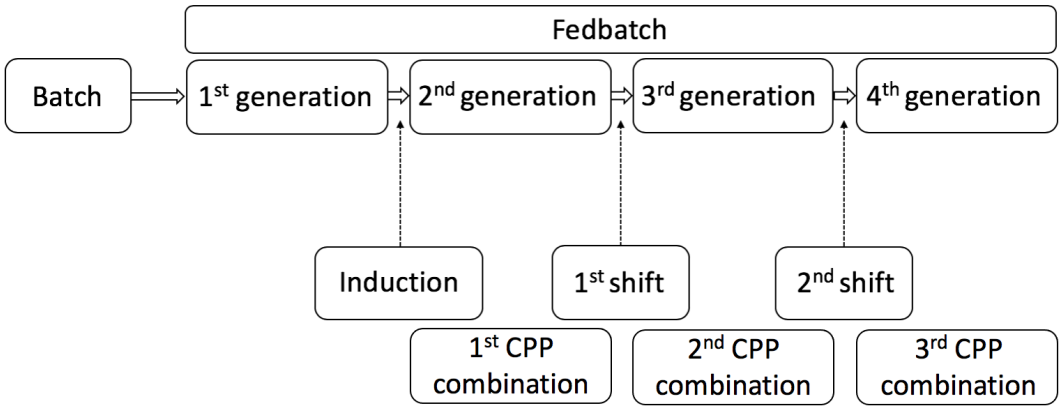

Figure 2: Schematic representation of the process of one iDoE experiment.

When designing the iDoE, important points had to be considered. On the one hand, a limiting factor is that the induction strength cannot be changed easily within a fermentation due to IPTG is not consumed. It is possible to add a higher concentration, however, it is not possible to reduce the induction strength. Therefore, only one specific concentration of IPTG was applied for each process. Due to this limitation, one can split the design space of the iDoE into three two-dimensional so-called "induction planes". Combined, three temperatures and three growth rates result in nine CPP setpoints per induction plane. Three set points are characterized within in one experiment, resulting in three intensified fermentations per induction plane. The CPP setpoints and shifts of parameters were arranged in a way that every combination is used in every state of the cell. In Table 4 the intensified fermentations, including the performed intra-experimental shifts, are listed.

Table 4: Intensified experiments with different CPP combination setpoints during one fermentation.

Identification number of experiment	Temperature [°C]	Target growth rate [h^{-1}]	Induction Strength [µmol IPTG/g CDM]
	setpoint1 → setpoint2 → setpoint3	setpoint1 → setpoint2 → setpoint3	
SOD114	37 → 34 → 34	0.1 → 0.2 → 0.1	0.2
SOD111	37 → 30 → 34	0.15 → 0.1 → 0.15	
SOD117	30 → 37 → 30	0.2 → 0.2 → 0.15	
SOD116	30 → 30 → 34	0.15 → 0.2 → 0.2	0.5
SOD112	30 → 34 → 37	0.1 → 0.1 → 0.2	
SOD110	34 → 37 → 37	0.15 → 0.15 → 0.1	
SOD115	37 → 34 → 30	0.2 → 0.15 → 0.2	0.9
SOD108	34 → 37 → 37	0.1 → 0.1 → 0.15	
SOD113	34 → 30 → 30	0.2 → 0.15 → 0.1	
SOD109	37 → 37 → 34	0.15 → 0.1 → 0.1	

3.2.5 Calculation of the time point for a shift of parameters

Based on the set target growth rate, the theoretical doubling time of the CDM was calculated. The initial CDM at the start of the fed-batch phase and the set target growth rate were used to calculate the

theoretical CDM trajectory at each time point of the fed-batch phase. Equation 1 (p. 16) shows the calculation. When the doubling of the initial CDM (X_0) of the feeding phase is reached, the recombinant protein production is induced by IPTG. After the second doubling of the cells, the first parameter shift was carried out. The second shift was performed after the third doubling of the cells. Due to different target growth rates carried out in the iDoE, the duration of each experiment varied.

3.3 Online measurements and control

Besides the above-mentioned standard process parameters, some other parameters were measured during the process:

1. Base consumption

   The base vessel was placed on a scale which continuously records values and represents information of how much base was transferred into the reactor and consumed by the cells.

2. 2D-Fluorescence spectroscopy

   A multi-wavelength online fluorescence sensor measures the 2D emission-excitation fluorescence spectrum from 270 nm to 590 nm. The used system for detection is BioView^{TM} (G111297, Delta Light & Optics, Denmark). A xenon lamp (light source), two rotating filter wheels and a detector measured the intensity of the fluorescence every 3 minutes.

3. Off-gas analysis

   The concentration of CO_2 and O_2 [%] were measured online and in real-time with the gas analyzer system BlueInOne Ferm (BlueSens gas sensor GmbH).

3.4 At-line and offline measurements

3.4.1 Sampling

By taking a control sample during the first generation, thus before induction, it was possible to check whether the CDM fits or if any contamination could be detected.

The sampling used for analytics started in each experiment at the time point of the IPTG induction. Hourly, 20 mL of cell broth were taken out of the bioreactor. The first 5 mL were discarded to purge the sampling valve. To increase the resolution and enable better insight into the behavior of the cell, the sampling frequency was increased after the parameter shift. Thereby, the effects of the shifts on the cell can be thoroughly examined. The sampling interval varied depending on the target growth rate after the shifts, as indicated in Table 5 (p. 21).

Table 5: Increased sampling interval after the parameter shifts related to the set target growth rate.

Target growth rate [h^{-1}]	Sampling frequency increased to 30 minutes
0.1	For 3h
0.15	For 2h
0.2	For 2h

The following analyses were performed out of the taken sample:

1. Optical density (OD_{600})

   The optical density was measured at a standard wavelength of λ = 600 nm (OD_{600}). The used measurement device was a photospectrometer (Amersham Biosciences Ultrospec 500 pro) and the measurements were carried out with disposable cuvettes (VWR International GmbH) in the linear range of 0.1 to 0.6. Therefore, the sample was diluted in the ratio of 1:201 with phosphate-buffered saline (PBS). With the obtained values of the OD_{600} measurement, the volume needed to transfer 1 mg CDM to an Eppendorf tube (1.5 mL) could be calculated ($3.5/OD_{600}$). The 1 mg CDM aliquots were centrifuged for 10 min at 21130 x g at 4°C, stored at -20°C and used for offline protein quantification.

2. CDM

   To determine the change in the CDM over the process duration, 2x5 mL of the sample were transferred into two 50 mL Greiner centrifuge tube (GREINER GmbH) and centrifuged for 10 min at 12697 x g (Eppendorf® Centrifuge 5804 R). The cell pellets were resuspended with 5 mL ddH_2O and subsequently centrifuged (washing step). This procedure was repeated twice. Afterward, the total sample was transferred to a pre-weighted beaker glass. It was stored at 105°C in a drying oven (Binder drying oven ED23) for a minimum of 24 h. The difference between the full and the empty beaker glass multiplied with 100 was used to calculate the amount of CDM in g/L. The total CDM was determined by multiplying this concentration with the calculated volume of the reactor, which contained the batch volume, the feed, the base and the inductor volume. The sample volume was subtracted from the total volume of the reactor.

3. Supernatant

   A volume of 1.8 mL of the supernatant from the centrifuged cell pellet were stored deep-frozen at -20°C in 2 mL Eppendorf reaction tubes and used for residual substrate analysis conducted

by high performance liquid chromatography (1100 HPLC, Agilent Technologies). The used column was an Aminex HPX-87H ion exclusion column (Biorad) and. as mobile phase 0.01 N sulfuric acid (20°C and 0.45 mL/minute) was consulted. A UV/VIS (Knauer) and a refractive index detector (Beckmann) were used for detection.

3.4.2 Gel electrophoresis

E. coli bacteria tend to form IBs, if cultivated at high temperatures and high induction strengths, therefore a sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) was used to obtain the ratio between soluble and insoluble recombinant protein and to monitor if initially IBs were formed. For SDS-PAGE a 1 mg CDM aliquot was taken. All chemicals and solutions used are listed in the appendix.

Therefore, the cells had to be disrupted first. The cell pellet was solved in 200 µL of disruption buffer and 50 µL of Lysozyme (2 mg/mL) and 50 µL of Benzonase (50 Units) were added and incubated at room temperature (RT) for 10 minutes (shaking at the TurboMix). 100 µL Triton X-100 were added and incubated for another 10 minutes (RT, shaking at the TurboMix). To separate the supernatant form the remaining cell pellet, the solution was centrifuged for 10 minutes at 19721 x g and 4°C (Eppendorf® Centrifuge 5424 R). The supernatant contained the soluble recombinant protein and was transferred to a new reaction tube and stored at 4°C until the gel was loaded. The remaining cell pellet was washed two times with 1 mL of washing buffer Tris/HCl 100 mM, pH 8.2 (resuspended and centrifuged at 19721 x g at 4°C for 10 min). After discarding the washing buffer for the second time 400 µL of the IB dissolving buffer were added. The cell pellet was solved and incubated in the dissolving buffer for 0.5 hours, softly shaking at the TurboMix (at RT). After centrifugation for 10 min at 19721 x g and 4°C, the supernatant contained the dissolved IBs and could be used to load the gel.

To load the gel 13 µL of the sample (soluble and IBs), 5 µL of NuPAGE® LDS sample buffer (4X) and 2 µL NuPAGE Reducing agent (10x) were mixed. Prior to loading the gel, the prepared solutions were heat-shocked (70°C; 10 min). A molecular weight marker, the ready to use MW-standard (Mark12™, unstained), was loaded directly onto the NuPAGE® gel (4-12% BisTris gradient gels). MES (20x) (Life Technologies) was used as running buffer. The electrophoresis was carried out at 200 V (max. 400 mA) and 50 minutes for one run.

By the end of the run, the gel was transferred to a fixing solution (40% ethanol; 50% RO-water; 10% acetic acid) for 30 minutes to prevent dispersion of the proteins. The solution denatures the protein and provides an acetic environment which enhances the interaction with the subsequent staining solution Coomassie brilliant blue R250 (45 minutes).

For de-staining, the gel was put in de-staining solution (25% acetic acid; 8% ethanol; 67% RO-water) for at least 2 h. The reaction process was terminated by relocating the gel into water.

To optically quantify the soluble protein and the IBs, the gel was scanned with Corel Photo-Paint 12 and analysed with an ImageQuant TL software. The raw volume of the single bands, obtained with the software, were taken to calculate the ratio between the soluble protein and the IB of each sample. The first lane always showed the protein ladder (MW Standard 12). hSOD migrates around the 20 kDa marker lane under reducing conditions. Herein, no standards were used to quantify the amount of hSOD, i.e., only the ratio was determined. To determine the concentration of recombinant hSOD in the samples, enzyme-linked immunosorbent assay was consulted.

3.4.3 Enzyme-linked immunosorbent assay

To determine the amount of produced recombinant hSOD in the fed-batch fermentations, enzyme-linked immunosorbent assay (ELISA) was used. The 1 mg CDM aliquots were taken for analysis and first had to be disrupted. All utilized chemicals and reagents used are listed in the appendix. The cell pellet was dissolved in 200 µL of disruption buffer and 50 µL of Lysozyme (10 mg/mL) were added, mixed well and incubated for 10 minutes at 37°C in a Thermo-shaker (350 rpm; Eppendorf® Thermomixer Compact). As the next step, 750 µL of Triton X-100 (0.5%) were added and the solution was incubated for another 10 minutes, softly shaken at 37°C (Thermomixer, 350 rpm). After the solution was centrifuged (10 min, 4°C, 19721 x g), the supernatant, containing the soluble protein, was transferred to a new reaction tube and stored at 4°C until the ELISA was performed.

The primary capture hSOD antibody (mAb IAM-SOD M05) was diluted in coating buffer at a ratio of 1:250 and 100 µL were transferred with a multichannel pipette (Thermo Scientific™, Finnpipette™ F2) in each well of a 96-well microtiter plate (Nunc® MaxiSorp™). The plate was incubated for two hours at RT or stored at 4°C overnight. Three washing steps, using a washing buffer, were carried out in a HydroFlex™ microplate washer (Tecan, Ref.: 30022011). The excess solution was removed.

In an uncoated 96-well microtiter plate (Nunc® MicroWell™) 140 µL of dilution buffer were pipetted into the wells of row A to G. The standard, with a concentration of 1.14 µg/mL, was diluted in dilution buffer in the ratio 1:11.4, reaching a working concentration of 100 ng/mL. Row H contained 300 µL of the standard (well 2-3) and four pre-diluted samples (well 4-11). The samples and the standard were diluted in 1:2 steps, from row G to A. For the samples, the expected hSOD concentration resulted from the experience of previous fermentations. The lowest dilution contained 50 ng hSOD/mL. Each standard and each sample were carried out in duplicates, at least.

As the next step 50 µL of each well were transferred to the microtiter plate, coated with the primary antibody, and incubated at RT for one hour. Afterward, the plate was washed three times as mentioned before. The enzyme-labeled secondary antibody (mAb IAM SOD A11H4 x AP) was diluted in dilution buffer in the ratio 1:1800 and 50 µL were added to each well. After an incubation time of 1 h at RT, the plate was washed three times. To detect the immunoreaction, 100 µL of staining solution was added to each well and after 7.5 min the measurement was carried out at a photometer (TECAN, Infinite® 200 Pro Series Multimode Reader) at a wavelength of 405 nm. The results were analysed using the software Magellan 6.

3.5 Process modelling

With the generated data in the iDoE fed-batch fermentations, a hybrid model was trained. Furthermore, its performance was compared to an already existing hybrid model based on the static DoE fermentations (Bayer, von Stosch, Striedner, Duerkop 2020).

3.5.1 Hybrid model

For the model building with the iDoE data set ($iDoE_9$) the Novasign hybrid modelling toolbox (Novasign GmbH, Vienna, Austria) was used. For this first approach, the chosen process variable to be modelled was the CDM concentration. The hybrid model consisted of a black-box, in this case an ANN, and a white-box in a serial structure, i.e., the target growth rate µ is estimated in the ANN using the process inputs and then transferred to the white box (Equation 2). The inputs to the ANN (black-box) were the CDM (g), the cultivation temperature (°C), the cumulative inductor mass (kg), the cumulative base (L) and the cumulative feed (L). A Levenberg–Marquardt algorithm was used to find modeling error minima to describe best the concentration profiles of the output variable, i.e., the CDM. A single hidden layer of four nodes was applied. The transferfunction for the hidden layer was set to hyperbolic while the output transfer functions were linear.

Equation 2 was used in the white box to describe the changes in the CDM:

$$\frac{dX}{dt} = \mu * X - D * X \qquad (2)$$

The concentration of the CDM (X) and the dilution rate of the fed-batch process due to feed addition (D) were provided by the white-box part, while the unknown rate expression of µ was modeled by the black-box, taking the impact of the process variables into account.

3.5.2 Model validation

To examine the quality and accuracy of the developed model, three criteria were used: the root mean square error (RMSE), the percentage error (%) as well as the standard deviation (SD). The RMSE was calculated with Equation 3, with the total number of observations (*N*), the measured values (*Y*) and the estimated (modelled) counterpart (Y_{model}) for each time point (*t*):

$$RMSE = \sqrt{\frac{1}{N} * \sum(Y_{(t)} - Y_{model\,(t)})^2} \qquad (3)$$

To facilitate the comparison of the models, the error is determined in percent by Equation 4, dividing the RMSE by the mean of the measured values ($\bar{Y}_{(t)}$):

$$Error\ [\%] = \frac{100}{N} * \sum \frac{|Y_{(t)} - Y_{model\,(t)}|}{\bar{Y}_{(t)}} \qquad (4)$$

To validate the model, the $iDoE_9$ data set was randomly split into a training a validation partition in a way that one experiment was used for validation, i.e., eight experiments for training and one experiment for validation. The model was built on the training partition and applied on the validation partition (internal validation). Once the error in the validation partition starts to increase, model-building was stopped to avoid overfitting. By doing so, the model stayed valid when it is applied on new data.

To provide a wide variety of models, this random partitioning procedure into so-called boots was repeated nine times, i.e., every possible combination of leaving an experiment out of the training partition was carried out.

To access the differences in the models and to enhance the robustness of the model, a bootstrap-aggregation (model averaging) of the performed boots was carried out, i.e., the best performing model of each boot was selected to identify the final model with the lowest error.

For external validation, the final iDoE model was applied to the test set DoE_{31}., With the averaged value achieved by the bootstrapped model ($Y_{bootstrapped}$) and the predicted value from each individual model (Y_{model}) for each time point (*t*) the SD was calculated in Equation 5:

$$SD_{(t)} = \sqrt{\frac{1}{n-1} * \sum(Y_{bootstrapped(t)} - Y_{model(i)(t)})^2} \qquad (5)$$

This equation contains the index *i*=1:5 and *n*, the number of observations. With the calculated SD, the prediction interval (PI) of the hybrid model at any time point (*t*) was calculated by the bootstrapped prediction of the model ± the SD at each time point, displaying the 68% prediction band (Equation 6).

$$PI_{(t)} = Y_{bootstrapped(t)} \pm SD_{(t)} \quad (6)$$

The performance of the iDoE hybrid model on the external validation was compared to the earlier developed hybrid model by Bayer et. al (2020), as a reference value. This static hybrid model was trained on 25 fed-batch fermentations of the DoE. For model testing, 6 fermentations of the DoE were used, chosen in a way that herein each induction plane was represented by two fermentations. More precisely, one experiment of each replicate run and three individual fermentations were selected. A detailed overview of the chosen fermentations is displayed in the Supporting Information of Bayer et al. (2020).

4 RESULTS AND DISCUSSION

4.1 Three-dimensional design space

4.1.1 iDoE setup

As an illustration, the full factorial DoE, comprising 27 CPP combination setpoints, is represented as a cube. Each axis represents a CPP at the chosen levels and each red dot herein marks respectively one fermentation set-up with one static induction strength, target growth rate and temperature during the entire duration of the run. To transform the standard DoE into an intensified one, the cube is subdivided into three so-called "induction planes". The comparative illustration of the DoE cube (Fig.3 **A**) and the iDoE cube (Fig.3 **B**) is shown in Figure 3.

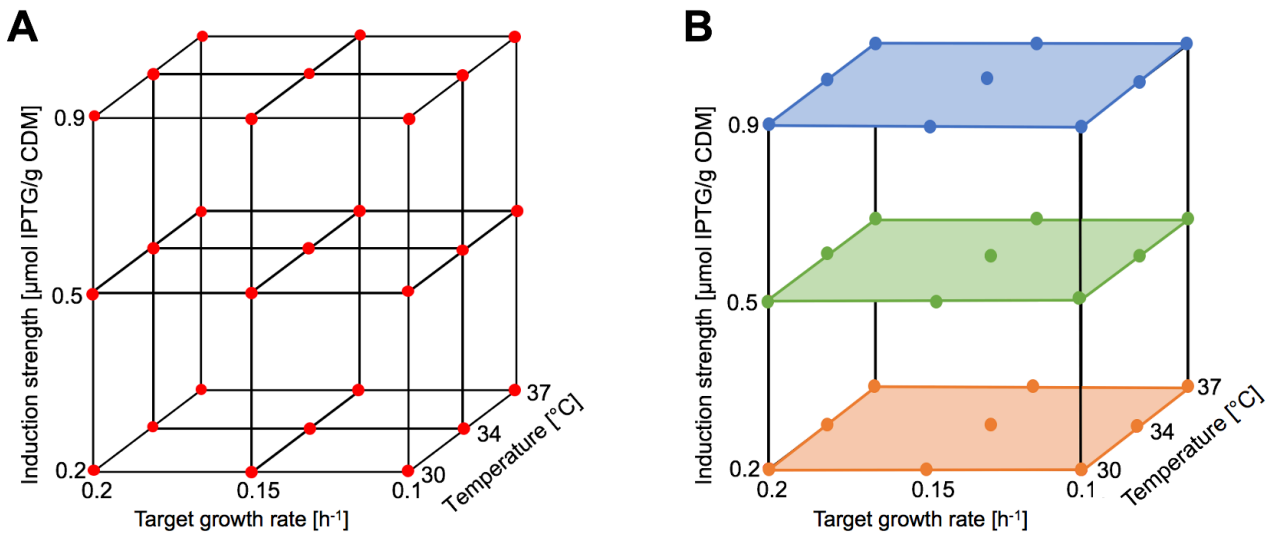

*Figure 3: DoE full factorial design cube (**A**) and iDoE cube with induction planes (**B**).*

The notional induction planes (Fig.3 **B**) are due to the fact that the induction strength was not changed during the fermentation. Within one induction plane, three different CPP combination setpoints should be represented per executed fermentation. The order of the execution of the iDoE fermentations was chosen at random. However, it was noted that the induction strength was varied after one experiment. Thereby each fermentation was chosen out of one induction plane.h^{-1}

In Figure 4 (p. 28) the detailed illustration of the three induction planes is visible. The lines indicate the order of the applied CPP combination setpoints in the respective fermentation. The combination of the parameters and the intra-experimental shifts were well thought out to make it possible that each CPP combination was examined in each point as shown in Figure 2 (p. 18). The different experiments were

labeled with the identification numbers SOD108-117, indicating the CPP setpoint at the start of the fermentation in Figure 4.

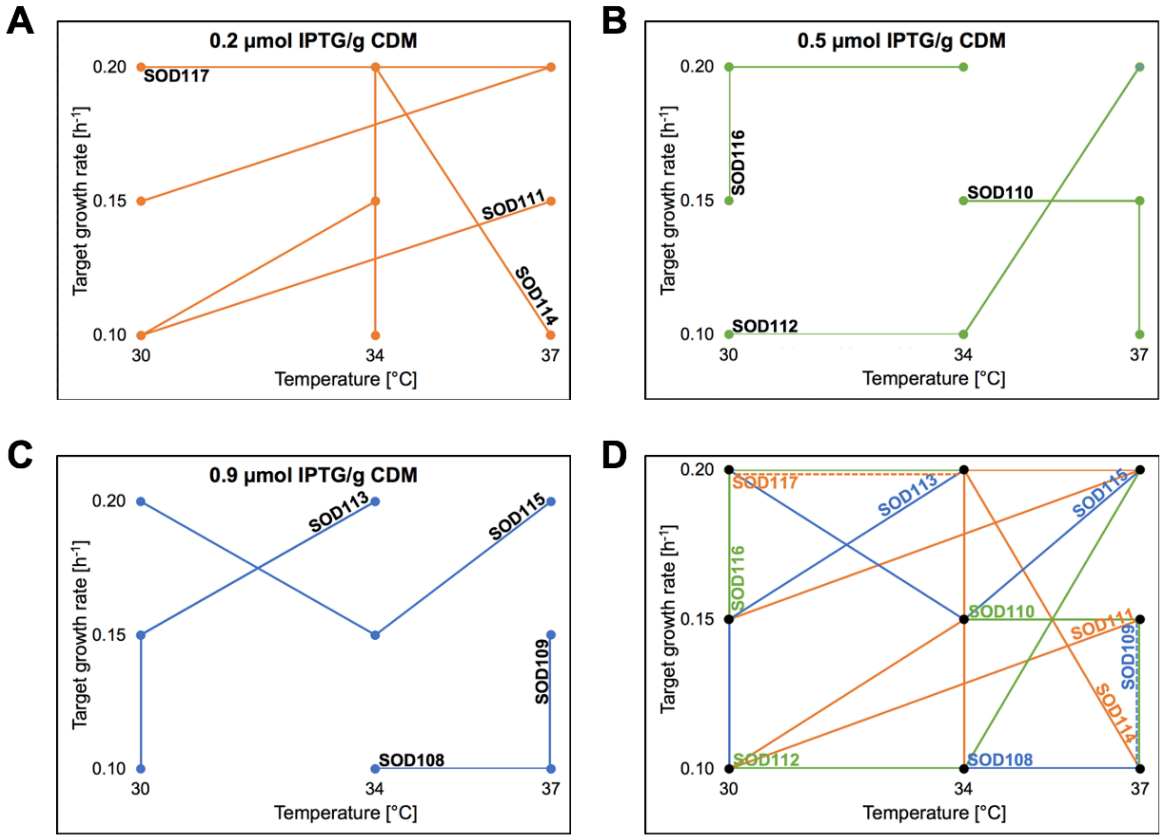

*Figure 4: Induction planes of the iDoE. The orange lines in induction plane 0.2 µmol IPTG/g CDM (**A**) connect the CPP combination setpoints in each respective fermentation. In **B** green lines connect the CPP combinations of induction plane 0.5 µmol IPTG/g CDM. In induction plane 0.9 µmol IPTG/g CDM (**C**) the blue lines connect the CPPs of the respective fermentations. In **D** all different combinations of CPP of all induction planes are overlaid. The identification number of the respective experiment (SOD108-117) marks the starting CPP setpoints.*

4.1.2 Investigation of a memory effect due to the shift direction

The first consideration was that the direction of the shifts in the iDoE might have an influence on the cell growth behaviour in the different setpoints, and therefore was examined. The experiment used for this investigation was SOD108, due to the CPP setpoints, i.e., high temperatures, a strong influence on the evolution of CDM and product was assumed and therefore is a well-suited indicator for an occurring memory effect. The fed-batch phase started with 34°C, the target growth rate 0.1 h^{-1} and the induction was carried out with 0.9 µmol/g CDM. After the first shift, the temperature was shifted to 37°C and µ=0.1 h^{-1}. The third setpoint of the CPPs was carried out at 37°C and µ=0.15h^{-1}. Another experiment (SOD109) with the same CPP combination setpoints was executed. However, the order of the CPP setpoints was reversed. The results of the cell physiology of both experiments were compared, i.e., the

CDM and product titer, and evaluated at the end of the process In Figure 5 the course of the CDM is represented.

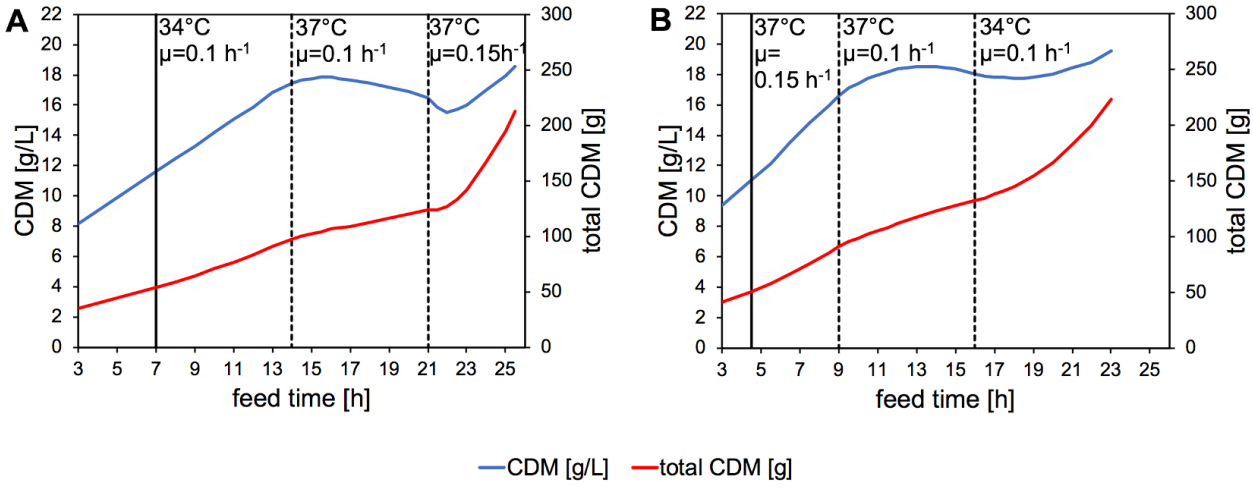

*Figure 5: Course of the CDM in g/L (blue line) and the total CDM [g] (red line) of experiment SOD108 (**A**) and SOD109 (**B**). The induction with 0.9 µmol IPTG/ g CDM is marked with the black line, the black-dotted lines represent the shift of parameter and consequently the start of a new generation. For each setpoint the temperature and target growth rate are shown.*

By increasing the target growth rate during the fermentation, the CDM showed a steep raise in the last setpoint of SOD108 (Fig.5 **A**). In contrast, SOD109 displayed a smoother course of CDM with no rapid leaps (Fig.5 **B**). Both experiments showed a similar course of the CDM in each respective setpoint of CPPs. At 34°C, µ=0.1 h^{-1} and 37°C, µ=0.15 h^{-1} the CDM increased in both experiments, while a temperature of 37°C and µ=0.1 h^{-1} led to a decrease of CDM.

The impact of the parameter shifts is also visible in the course of the product titer. In Figure 6, the course of the soluble, insoluble and total recombinant hSOD of both fermentations are shown.

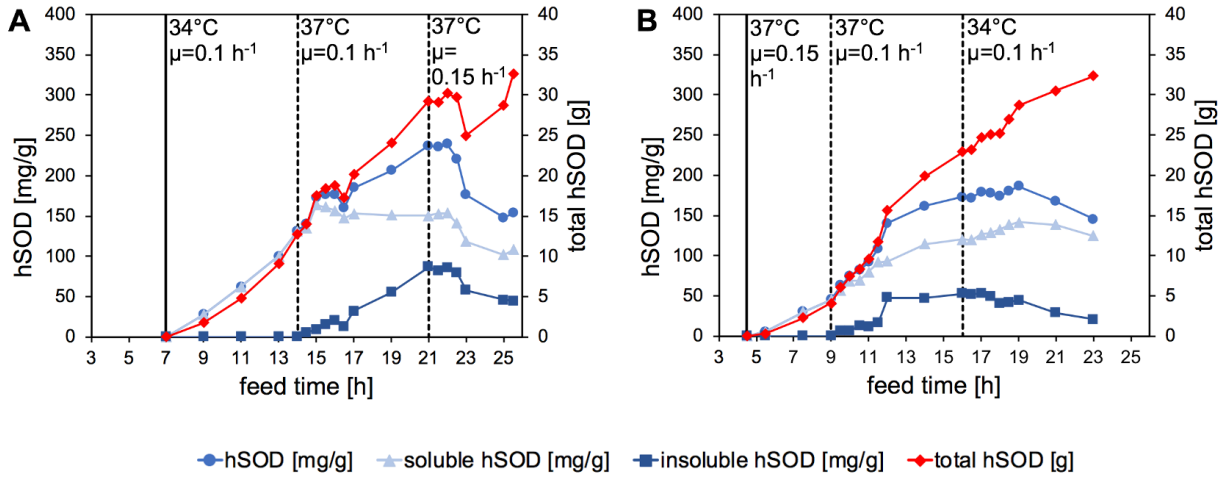

*Figure 6: Course of product titer of SOD108 (**A**) and SOD109 (**B**). Comparison of the specific product in mg/g (blue line) and the total product rate in g (red line). The ratio of soluble (light blue line) and insoluble hSOD (dark blue line) is displayed as sperate curves. The induction with 0.9 µmol IPTG/ g CDM is marked with the black line, the black-dotted lines represent the shift of parameter. For each setpoint the temperature and target growth rate are shown.*

In both experiments, an adaption phase after shifting the parameters was visible, as presented in Figure 6. The high temperature (37°C) led to the formation of insoluble hSOD.

Despite the expected slight deviation in trends of CDM and product titer, the course of the curve in the respective setpoints behave similar. As a result, it was assumed that the chronology of the parameter shifts does not matter in this setup and the iDoE could be executed as planned.

4.2 Comparison of the static and iDoE approach

To compare the offline trend and the cell physiology of both approaches, each intensified experiment was compared to the associated static experiments. This comparison is subdivided into the three induction planes. Therefore, all experiments of one plane are discussed in one section.

In addition, an overview of the endpoint values of iDoE and DoE fermentations are represented to investigate the response to the different CPP settings.

4.2.1 Comparison experiments of induction plane 0.2 µmol IPTG/g CDM

4.2.1.1 SOD111

The first setpoint had a set growth rate of µ=0.15 h^{-1} and the temperature was set to 37°C. The induction in SOD111 was conducted at a feed time of 4.5 h. The course of the CDM of SOD111 and its respective static fermentations SOD95, SOD94, SOD97 is shown in Figure 7 (**A-D**).

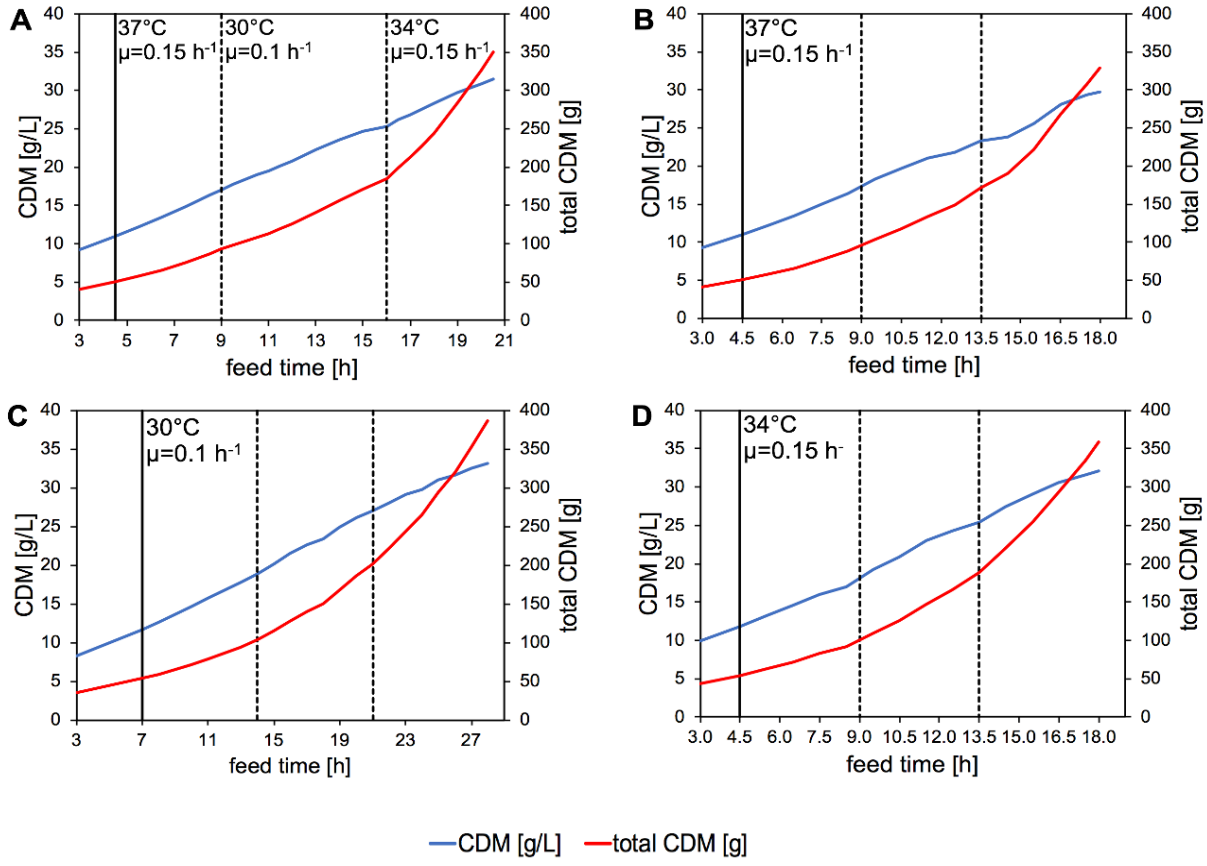

*Figure 7: Comparison of the course of the CDM of intensified experiment SOD111 (**A**) to the respective static fermentations: SOD95 (**B**), SOD94 (**C**), SOD97 (**D**). The blue line shows the course of the CDM concentration [g/L] and the red line displays the total CDM [g]. All experiments were induced with 0.2 µmol IPTG/g CDM after one generation, displayed as black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

At the first shift of parameters (9 h), to 30°C and µ=0.1 h^{-1} a concentration of 17.1 g/L CDM was reached, furthermore the value of the total CDM was 93 g. The associated static fermentation SOD95 (Fig.7 **B**) reached a CDM concentration of 17.3 g/L, i.e., 96 g. In Figure 7 **C** the static experiment SOD94, respective to the second setpoint of intensified experiment SOD111, is displayed. The static experiment SOD97 (Fig.7 **D**) is associated to the last setpoint of SOD111. Both static experiments showed a similar behaviour in their respective phase of the intensified experiment.

Figure 8 displays the course of the product titer of SOD111 and its respective experiments.

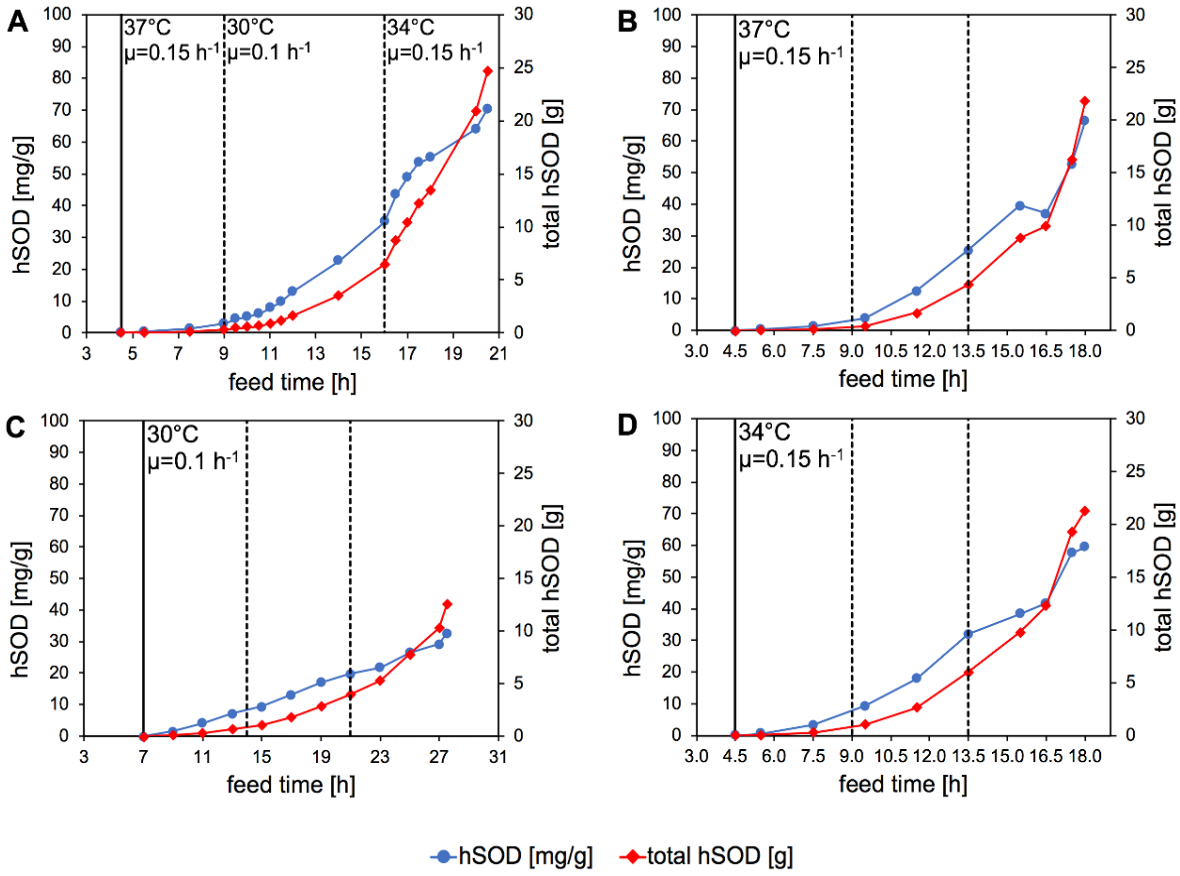

*Figure 8: Comparison of the course of the product titer of the intensified experiment SOD111 (**A**) to the respective static fermentations: SOD95 (**B**), SOD94 (**C**), SOD97 (**D**). The specific product (in mg/g) is marked as a blue line and the total product rate (in g) is displayed as red line. All experiments were induced with 0.2 µmol IPTG/g CDM after one generation, which is displayed with the black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

No IBs were formed in the intensified as well as in the associated static experiments with an induction strength of 0.2 µmol IPTG/g CDM. At a feed time of 9 h (first shift) in SOD111 (Fig.8 **A**), 3.0 mg/g hSOD and 0.3 g of total hSOD were produced. In the respective static fermentation, SOD95 (Fig.8 **B**), 2.6 mg/g and 0.3 g of hSOD were generated. In the third generation of SOD94 (Fig.8 **C**), respective for the second setpoint, a similar linear increase as in SOD111 was obtained. In the third setpoint of SOD111 a strong increase of product was identified. In the experiment SOD97 (Fig.8 **D**) a comparable strong increase could be obtained in the last generation.

4.2.1.2 SOD114

The CPPs for the first setpoint of SOD114 were 37°C and μ=0.1 h^{-1}. The induction was set to 0.2 µmol IPTG/g CDM at a feed time of 7 h. In Figure 9, the course of the CDM for each CPP setpoint of the intensified experiment SOD114 (Fig.9 **A**) is compared to the associated static fermentations SOD104, SOD103, SOD106 (Fig.8 **B**, **C**, **D**).

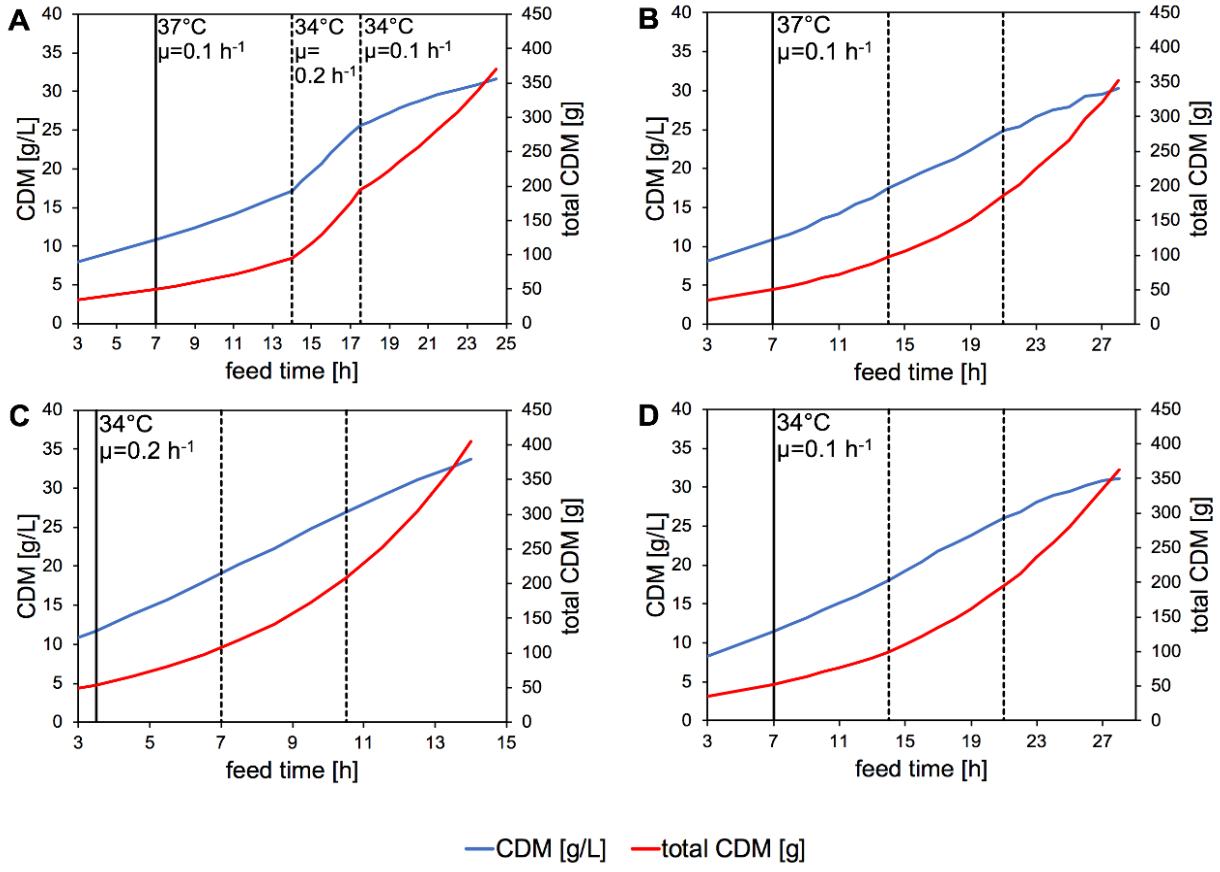

*Figure 9: Comparison of the course of the CDM of intensified experiment SOD114 (**A**) to the respective static fermentations: SOD104 (**B**), SOD103 (**C**), SOD106 (**D**). The blue line shows the course of the CDM concentration [g/L] and the red line displays the total CDM [g]. All experiments were induced with 0.2 µmol IPTG/g CDM after one generation, displayed as black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted line.*

At the first shift of parameters (14 h), a concentration of 17.2 g/L and total CDM of 95 g was reached in the experiment SOD114 (Fig.9 **A**). In the respective static fermentation, SOD104 (Fig.9 **B**), values of 17.5 g/L CDM and 97 g in total were archived at the end of the second generation (14 h). In the second setpoint of SOD114, the adaption to the high growth rate (μ=0.2 h^{-1}) could be obtained in the course of the CDM. In SOD103 (Fig.9 **C**), which is respective for the second setpoint, the increase of CDM is clearly visible. In the last setpoint of SOD114 the curve of the CDM concentration is flattening, which is also observed in SOD106 (Fig.9 **D**), the respective static fermentation.

The trend of the product titer of SOD114 and its associated fermentations is shown in Figure 10 **A-D**.

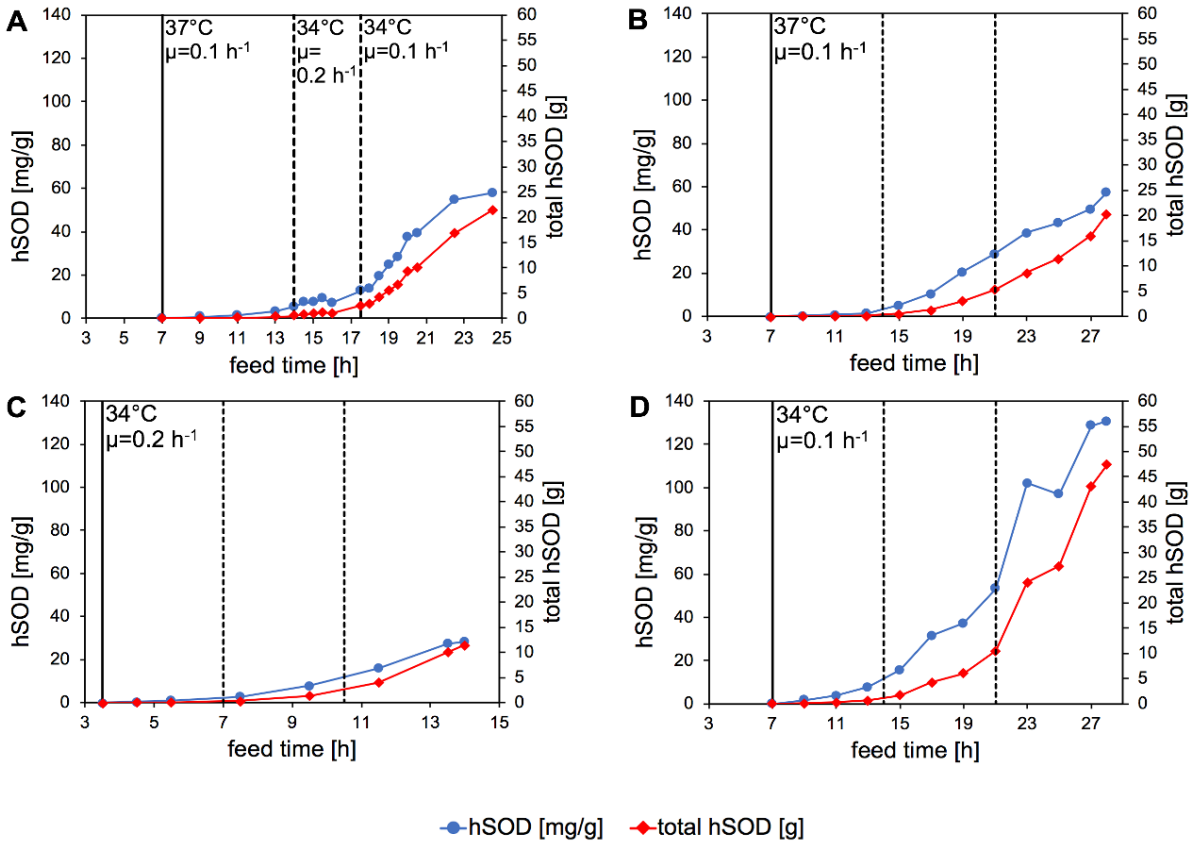

*Figure 10: Comparison of the course of the product titer of the intensified experiment SOD114 (**A**) to the respective static fermentations: SOD104 (**B**), SOD103 (**C**), SOD106 (**D**). The specific product (in mg/g) is marked as a blue line and the total product rate (in g) is displayed as red line. All experiments were induced with 0.2 µmol IPTG/g CDM after one generation, which is displayed with the black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

At the first shift, at a feed time of 14 h, in SOD114 (Fig.10 **A**), 5.5 mg/g hSOD and 0.5 g of total hSOD were produced. In SOD104 (Fig.8 **B**), the respective static fermentation, 3.4 mg/g and 0.3 g of hSOD were generated. In the second setpoint, a slight increase of the CDM could be obtained in SOD114 and the respective static experiment SOD103 (Fig.10 **C**). The product titer curve showed a drastic increase in the third setpoint of SOD114, due to favourable CPP setpoints for product formation, which is also visible in the comparable static experiment SOD106 (Fig.10 **D**).

4.2.1.3 SOD117

In SOD117, the starting CPPs were 30°C and µ=0.2 h^{-1}. The induction was performed at feed time 3.5 h, with the set ratio of 0.2 µmol IPTG/g CDM. The course of the CDM of the intensified experiment as well as its respective static fermentations are displayed in Figure 11.

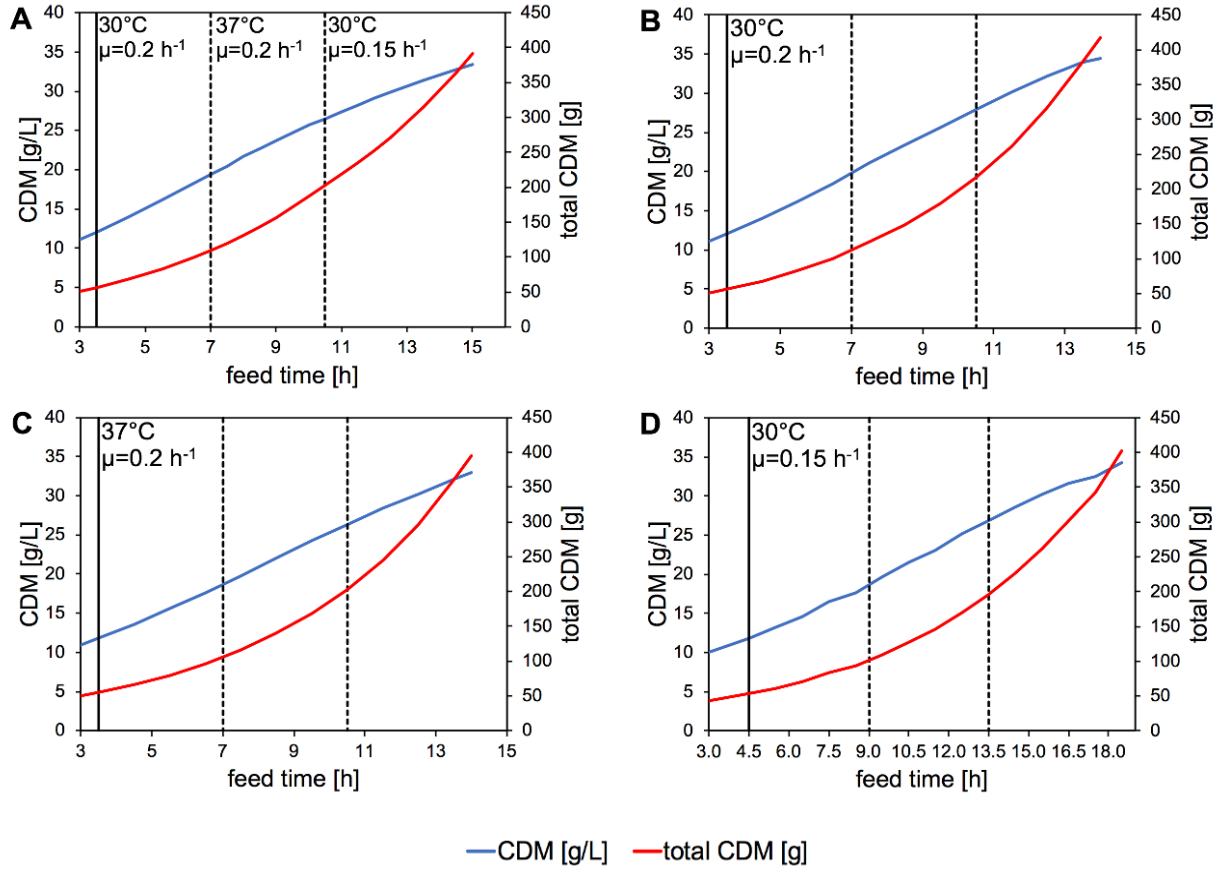

*Figure 11: Comparison of the course of the CDM of intensified experiment SOD117 (**A**) to the respective static fermentations: SOD102 (**B**), SOD107 (**C**), SOD99 (**D**). The blue line shows the course of the CDM concentration [g/L] and the red line displays the total CDM [g]. All experiments were induced with 0.2 µmol IPTG/g CDM after one generation, displayed as black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

A concentration of 19.5 g/L CDM, i.e., 109 g total CDM were reached at the first shift of parameters (7 h), to 30°C and µ=0.2 h^{-1} in SOD117 (Fig.11 **A**). The associated static fermentation SOD102 (Fig.11 **B**) reached a CDM concentration of 19.8 g/L, i.e., 112 g. The static experiment SOD107 (Fig.11 **C**), respective to the second setpoint of intensified experiment SOD111, and SOD99 (Fig.11 **D**), the respective experiment to the third setpoint, displayed a similar behaviour in their to the intensified experiment associated phase. Assimilable to experiment SOD111 (p. 31, Fig.7 **A**), the CDM concentration had a linear increase, while the total CDM showed an exponential growth.

The course of the product titer of SOD117 and its respective static fermentations is shown in Figure 12.

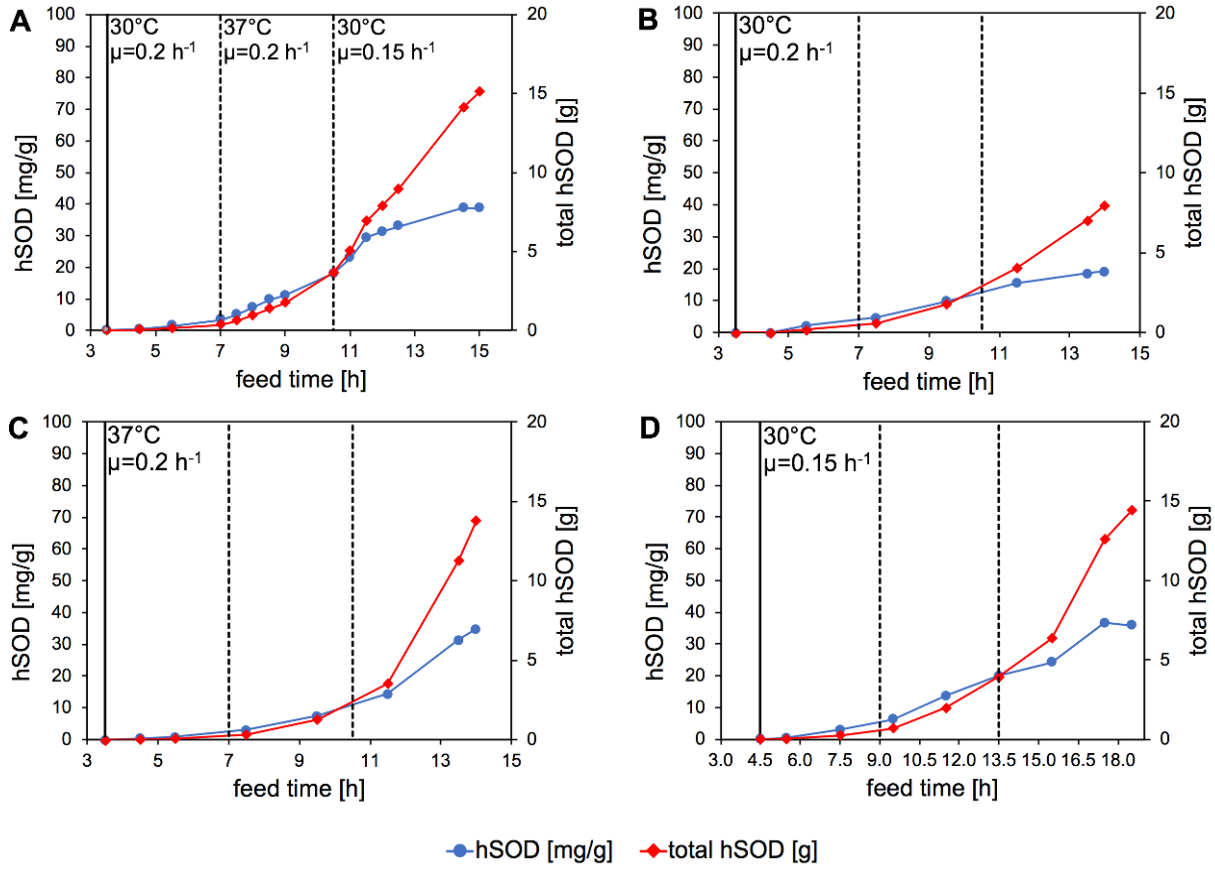

*Figure 12: Comparison of the course of the product titer of the intensified experiment SOD117 (**A**) to the respective static fermentations: SOD102 (**B**), SOD107 (**C**), SOD99 (**D**). The specific product (in mg/g) is marked as a blue line and the total product rate (in g) is displayed as red line. All experiments were induced with 0.2 µmol IPTG/g CDM after one generation, which is displayed with the black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

In SOD117, as well as in the comparable static fermentations, no IBs were formed. In SOD117 (Fig.12 **A**) 3.4 mg/g hSOD and 0.4 g of total hSOD were produced at a feed time of 7 h (first shift of parameters). In the respective static fermentation SOD102 (Fig.11 **B**) the exact same amount of hSOD was generated. Comparable to both other experiments at an induction strength of 0.2 µmol IPTG/g CDM, the CDM in SOD117 increased slightly in the second setpoint. A similar increase could be obtained in the course of the curve of SOD107 (Fig.12 **C**), the respective fermentation to the second setpoint. A drastic increase in the total CDM was obtained in the third setpoint of SOD117, similar to the associated static fermentation SOD99 (Fig.12 **D**).

4.2.2 Comparison experiments of induction plane 0.5 µmol IPTG/g CDM

4.2.2.1 SOD110

The first experiment carried out of the induction plane 0.5 µmol IPTG/g CDM was SOD110. The CPPs at the start were 34°C and µ=0.15 h^{-1} and the induction was carried out at feed time 3.5 h and. The course of the CDM is shown in Figure 13.

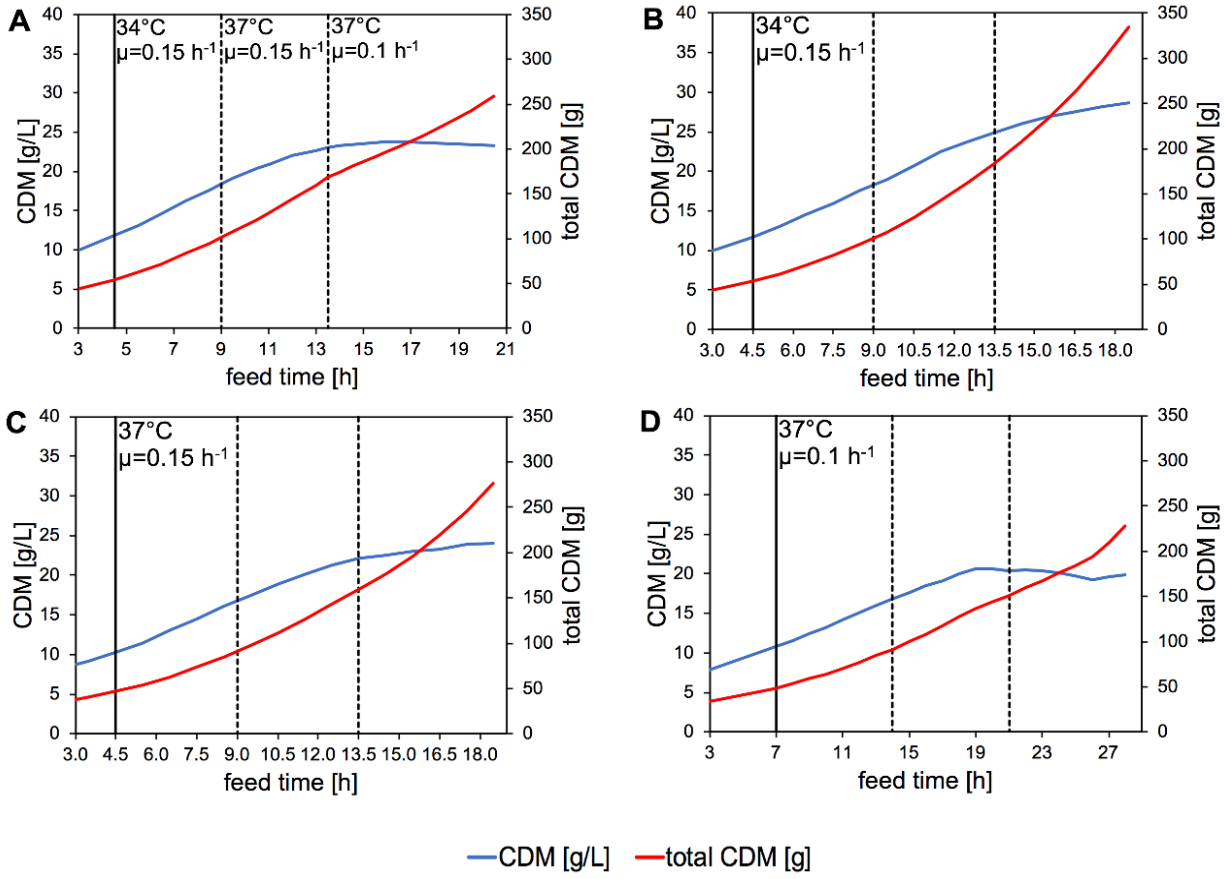

*Figure 13: Comparison of the course of the CDM of intensified experiment SOD110 (**A**) to the respective static fermentations: SOD101 (**B**), SOD53 (**C**), SOD47 (**D**). The blue line shows the course of the CDM concentration [g/L] and the red line displays the total CDM [g]. All experiments were induced with 0.5 µmol IPTG/g CDM after one generation, displayed as black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

A concentration of 18.4 g/L and total CDM of 101 g was reached at the first shift of parameters (9 h) in SOD110 (Fig.13 **A**). At the respective static fermentation of the first set point, SOD101 (Fig.13 **B**), values of 18.3 g/L and 101 g in total were achieved. The small deviation of less than 0.1 g/L between the intensified and the static experiment indicates high reproducibility and small process deviations. In the static experiment SOD53 (Fig.13 **C**), respective for the second setpoint of SOD110, a similar course of CDM was obtained. In Figure 13 **D** the course of SOD47 is displayed. Its last generation is respective to the third setpoint of SOD110. In both, the intensified and the comparable static, the flattening of the CDM concentration curve was observed, still indicating the same process behaviour.

In Figure 14, the course of the product titer of SOD110 and the respective static fermentations is shown.

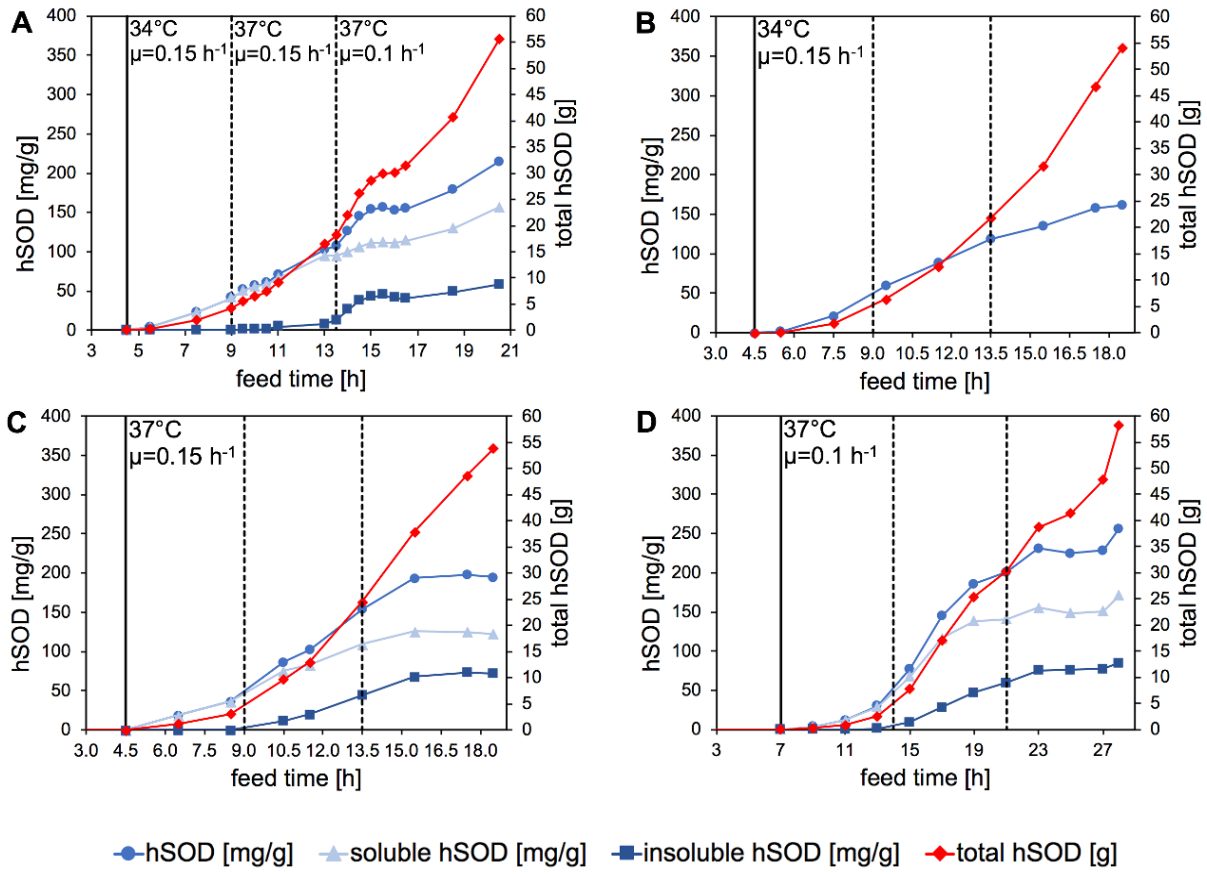

*Figure 14: Comparison of the course of the product titer of the intensified experiment SOD110 (**A**) to the respective static fermentations: SOD101 (**B**), SOD53 (**C**), SOD47 (**D**). The specific product in mg/g is marked as a blue line and the total product rate in g is displayed as red line. The ratio of soluble (light blue line) and insoluble hSOD (dark blue line) is displayed as sperate curves. All experiments were inducted with 0.5 µmol IPTG/g CDM after one generation, which is displayed with the black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

Values of 41.4 mg/g and a total hSOD of 4 g were reached at a feed time of 9 h (first shift) in SOD110 (Fig.14 **A**). In the respective static fermentation SOD101 (Fig.14 **B**) the product titer was 40.2 mg/g, i.e., 4 g in total. The high temperature of 37°C in setpoint 2 and 3, as well as in the respective static fermentations SOD53 and SOD47 (Fig.14 **C, D**) led to the formation of IBs. In the third setpoint of SOD110, the adaption to the new parameters set is visible. The last generation of the static fermentation SOD47 (Fig.14 **D**), which represents the last setpoint of SOD110, displays the same increase of the product titer as visible in SOD110 (Fig.14 **A**).

4.2.2.2 SOD112

In SOD112, the CPPs at the start were 30°C and µ=0.1 h^{-1}. The induction with 0.5 µmol IPTG/g CDM was performed at feed time 7 h The course of the CDM of the intensified experiment and its respective static fermentations are displayed in Figure 15.

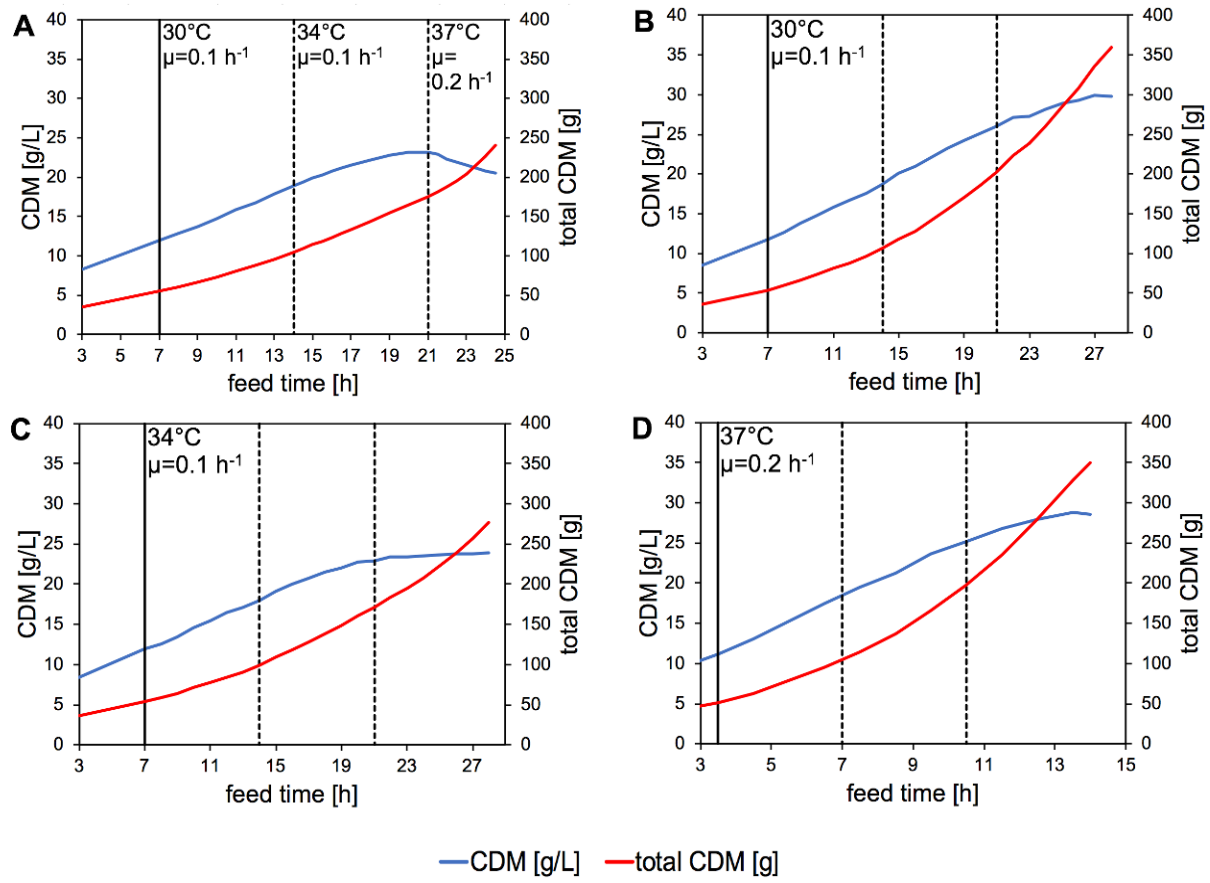

*Figure 15: Comparison of the course of the CDM of intensified experiment SOD112 (**A**) to the respective static fermentations: SOD63 (**B**), SOD96 (**C**), SOD58 (**D**). The blue line shows the course of the CDM concentration [g/L] and the red line displays the total CDM [g]. All experiments were induced with 0.5 µmol IPTG/g CDM after one generation, displayed as black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

In SOD112 a concentration of 18.9 g/L CDM, i.e., 105 g total CDM was reached at the first shift of parameters (14 h), to 34°C and µ=0.1 h^{-1} (Fig.15 **A**). In the respective static fermentation SOD63 (Fig.15 **B**) a concentration of 18.7 g/L, i.e., 106 g was obtained. The static experiment SOD96 (Fig.15 **C**), respective to the second setpoint of intensified experiment SOD112, showed a similar curve behaviour. SOD58 (Fig.15 **D**), the associated experiment to the third setpoint, a slight difference to SOD112 in the course of the CDM concentration is visible. After the second shift in SOD112, the CDM concentration started to decrease, due to the unfavourable set CPPs, while in the static experiment SOD58 (Fig.15 **D**) a slower flattening of the curve was obtained.

In Figure 16 the course of the product titer of SOD112 and the associated static fermentations is shown.

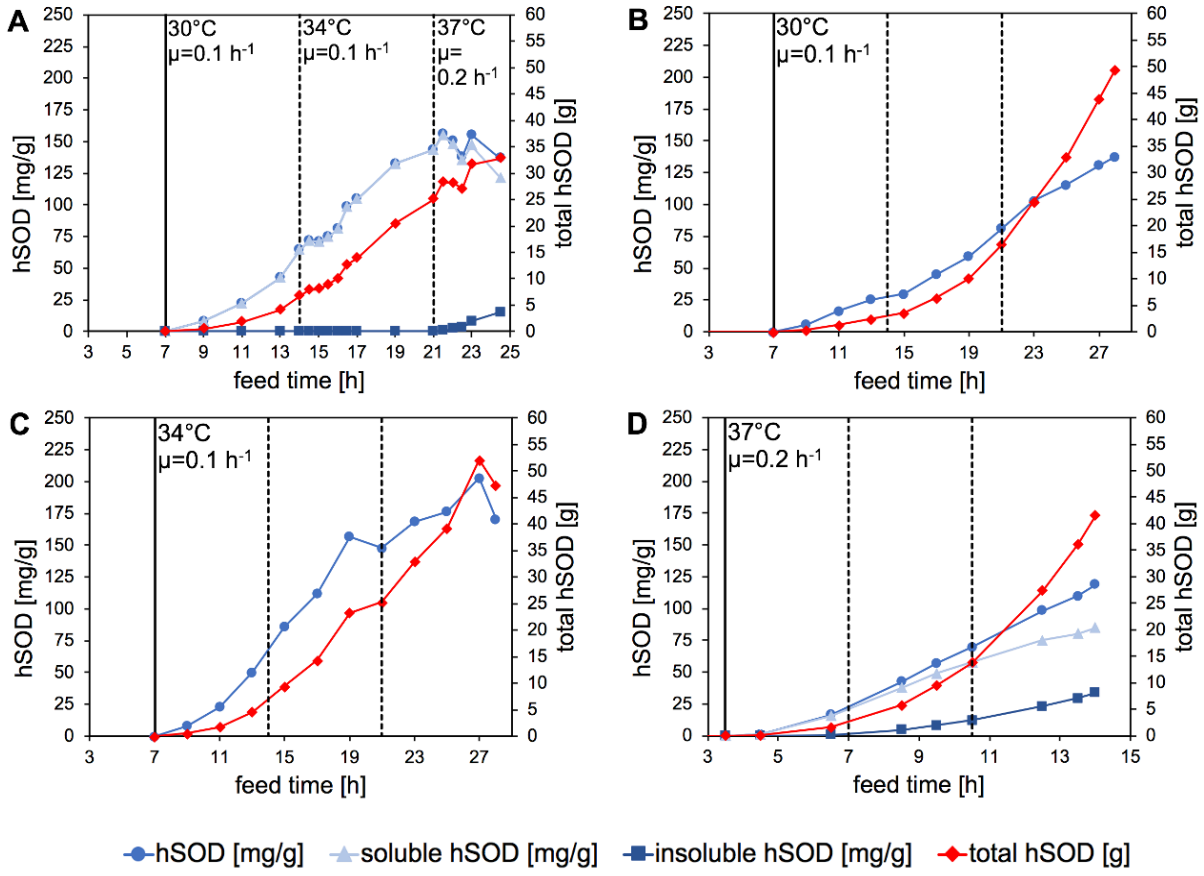

*Figure 16: Comparison of the course of the product titer of the intensified experiment SOD112 (**A**) to the respective static fermentations: SOD63 (**B**), SOD96 (**C**), SOD58 (**D**). The specific product in mg/g is marked as a blue line and the total product rate in g is displayed as red line. The ratio of soluble (light blue line) and insoluble hSOD (dark blue line) is displayed as sperate curves. All experiments were inducted with 0.5 µmol IPTG/g CDM after one generation, which is displayed with the black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

In Figure 16 **A** the lag phase of the product titer to adapt after the shifts in the intensified fermentation is clearly visible. Regarding the particular setpoints, values of 64.3 mg/g and a total hSOD of 7 g were reached at a feed time of 14 h (first shift) in SOD112 (Fig.16 **A**). In the respective static fermentation SOD63 (Fig.16 **B**) the product titer was 27.6 mg/g, i.e., 3 g in total. Half of the amount of product titer produced in SOD112 was reached in the static experiment SOD63, however the slight increase of the curve was obtained in both. Due to the temperature change from 30°C to 34°C in the second setpoint of SOD112, the product titer increased after a short adaptation phase. The steep curve was also observed in the associated experiment SOD96 (Fig.16 **C**). In SOD112, the formation of hSOD in IBs was, as expected, only detected in the third setpoint (Fig.16 **A**) due to the high temperature and the high growth rate. The formation of insoluble protein was also observed in the for the third setpoint respective static fermentation SOD58 (Fig.16 **D**).

4.2.2.3 SOD116

The CPP setpoint in SOD116 at the start were 30°C and a growth rate of μ=0.15 h^{-1}. It was induced with 0.5 µmol IPTG/g CDM at a feed time of 4.5 h. In Figure 17, the intensified experiment SOD116 and the respective static fermentations, SOD80, SOD66, SOD105, are displayed.

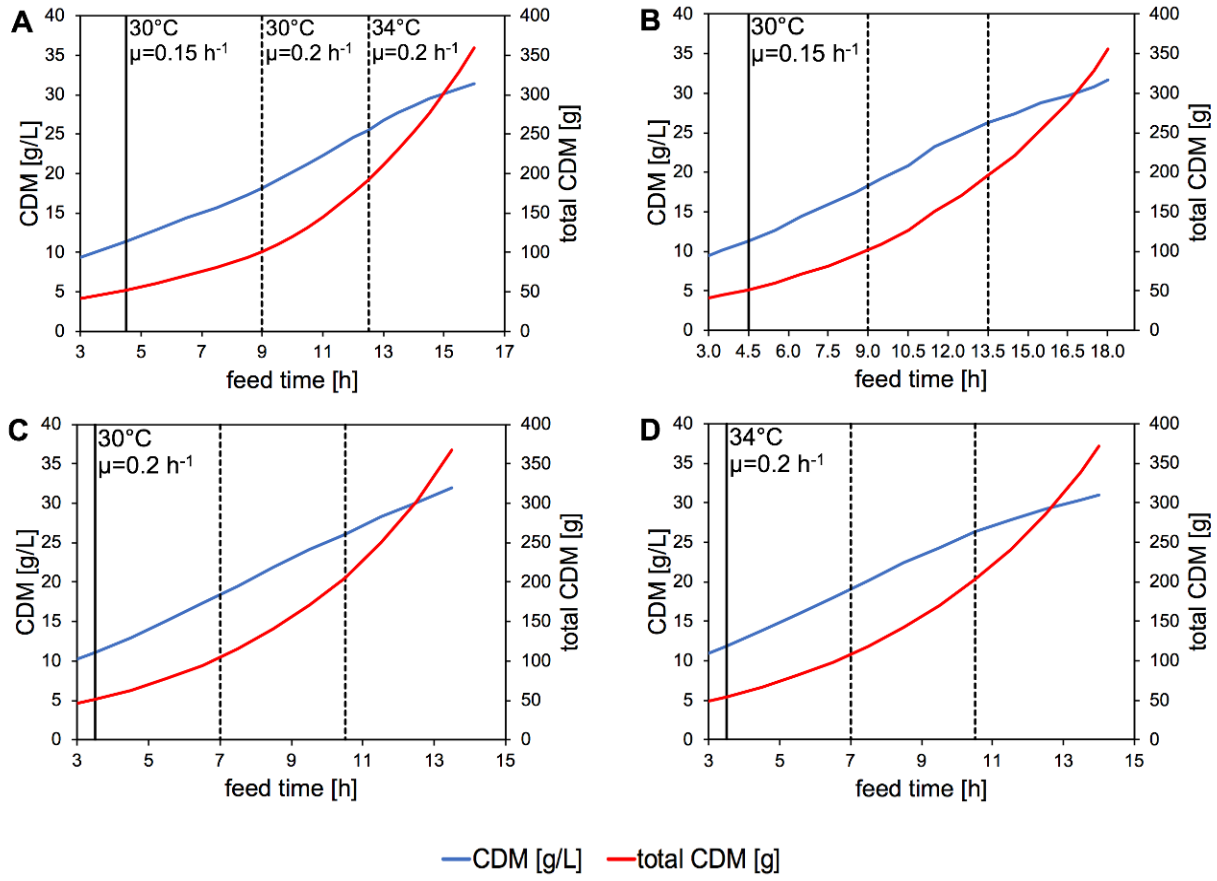

*Figure 17: Comparison of the course of the CDM of intensified experiment SOD116 (**A**) to the respective static fermentations: SOD80 (**B**), SOD66 (**C**), SOD105 (**D**). The blue line shows the course of the CDM concentration [g/L] and the red line displays the total CDM [g]. All experiments were induced with 0.5 µmol IPTG/g CDM after one generation, displayed as black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

The values at the first shift of parameters (9 h) were 18.2 g/L and 101 g total CDM (Fig.17 **A**). In SOD80 (Fig.17 **B**), the comparable static fermentation to the first setpoint, the values of CDM were 18.4 g/L and 102 g total CDM, i.e., fewer than a difference of 0.2 g/L was observed. In the second setpoint of SOD116, the course of the curve increased, due to the change of the growth rate from μ=0.15 h^{-1} to μ=0.2 h^{-1}. The similar course of the curve was obtained in SOD66 (Fig.17 **C**). Due to the shift to 34°C in the last setpoint of SOD116, the total CDM increased while the concentration started to flatten, which was also observed in SOD105 (Fig17 **D**).

In Figure 18 the course of the product titer in fermentation SOD116 and the associated static experiments is shown.

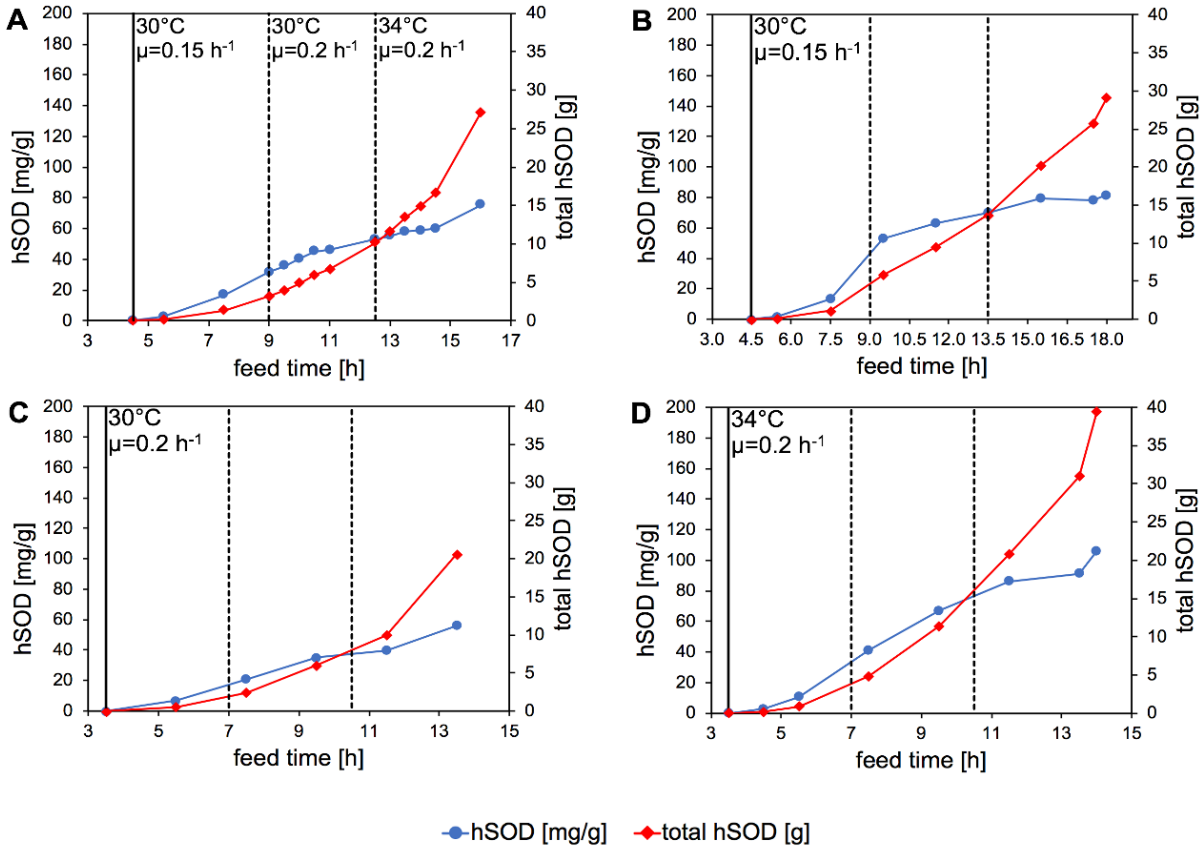

*Figure 18: Comparison of the course of the product titer of the intensified experiment SOD116 (**A**) to the respective static fermentations: SOD80 (**B**), SOD66 (**C**), SOD105 (**D**). The specific product in mg/g is marked as a blue line and the total product rate in g is displayed as red line. The ratio of soluble (light blue line) and insoluble hSOD (dark blue line) is displayed as sperate curves. All experiments were inducted with 0.5 µmol IPTG/g CDM after one generation, which is displayed with the black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

Similar to the CDM, the course of the product titer of the intensified and the respective static experiments showed a similar behaviour (Figure 18). No IBs were formed during all shown fermentations. At the first shift of parameters (9 h), the values were 31.6 mg/g and 3 g of total hSOD, while in the respective static fermentation, SOD80 (Fig.18 **B**), values of 33.2 mg/g and 4 g of total hSOD were obtained. In SOD66 (Fig.18 **C**), respective to setpoint two of SOD116, the same slight increase of the curve was obtained. In the last setpoint of SOD116, the total product titer had a considerable increase while the hSOD concentration showed a flat curve with a steam increase at the last values. A similar behaviour was obtained in SOD105 (Fig.18 **D**), the respective static experiment to the last setpoint of SOD116.

4.2.3 Comparison experiments of induction plane 0.9 µmol IPTG/g CDM

4.2.3.1 SOD108

SOD108 started with a temperature of 34°C and µ=0.1 h^{-1} and was induced at a feed time of 7 h with 0.9 µmol IPTG/g CDM. In Figure 19, the course of the CDM of the intensified experiment SOD108 and the three respective static experiments are shown.

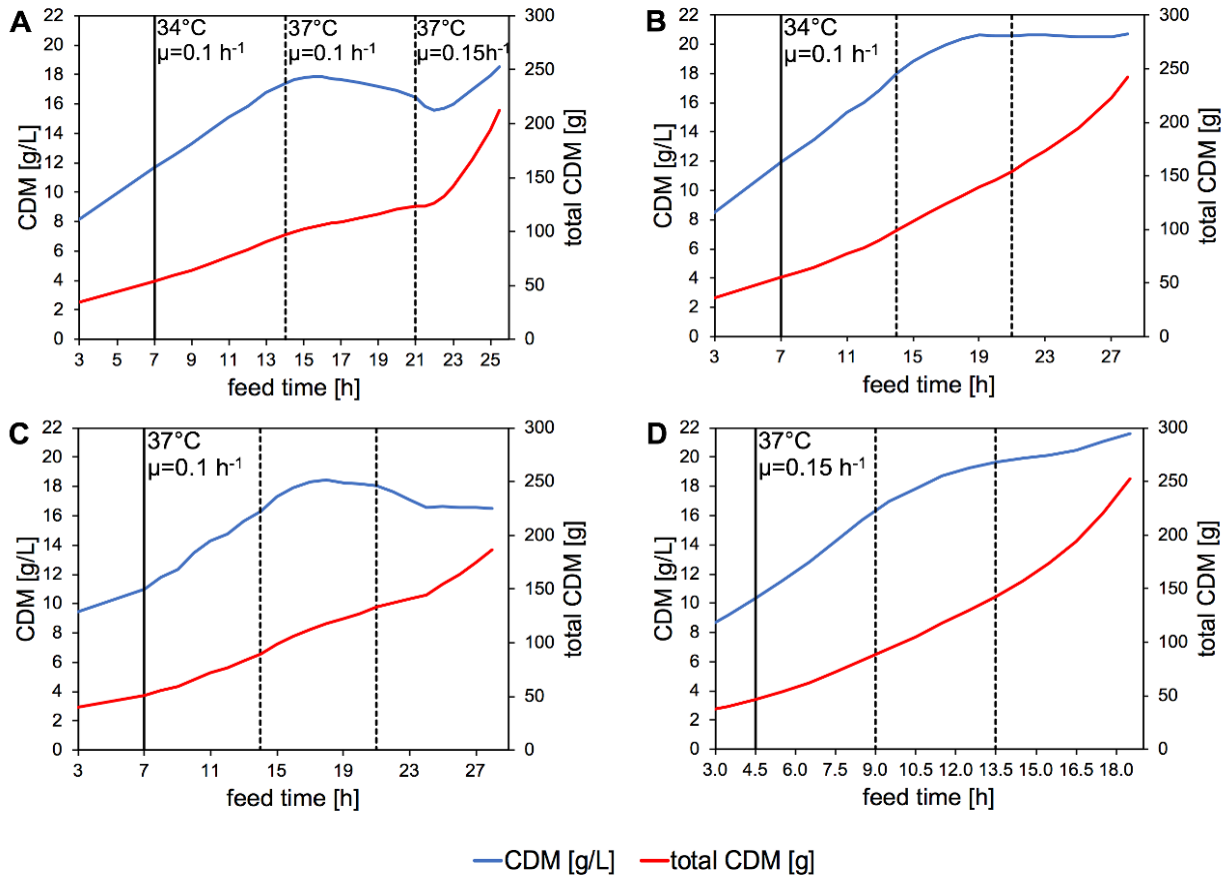

*Figure 19: Comparison of the course of the CDM of intensified experiment SOD108 (**A**) to the respective static fermentations: SOD98 (**B**), SOD92 (**C**), SOD49 (**D**). The blue line shows the course of the CDM concentration [g/L] and the red line displays the total CDM [g]. All experiments were induced with 0.9 µmol IPTG/g CDM after one generation, displayed as black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

In SOD108 a concentration of 17.4 g/L CDM, i.e., 97 g total CDM was reached at the first shift of parameters (14 h), to 37°C and µ=0.1 h^{-1} (Fig.19 **A**). SOD98 (Fig.19 **B**) represents the first setpoint of the intensified experiment and therefore showed similar values, with 18.1 g/L CDM, i.e., 100 g total CDM, and a similar trend for this phase. After a short lag phase due to the first CPP shift to 37°C in SOD108, the CDM concentration decreased due to the higher temperature (Fig.19 **A**). The second setpoint is represented by the static experiment SOD92 (Fig.19 **C**). The growth behaviour of the third generation is similar to the intensified experiment at the second setpoint. In the third setpoint of SOD108, both, concentration and total CDM, showed a steep increase till the process endpoint. The third setpoint is

represented by the static experiment SOD49 (Fig.19 **D**). The high temperature (37°C) and target growth rate (μ=0.15 h^{-1}) led to a strong increase of CDM during the entire experiment, which is also visible in setpoint three of the intensified experiment.

The course of the product titer of SOD108 and the respective static fermentations are shown in Figure 20.

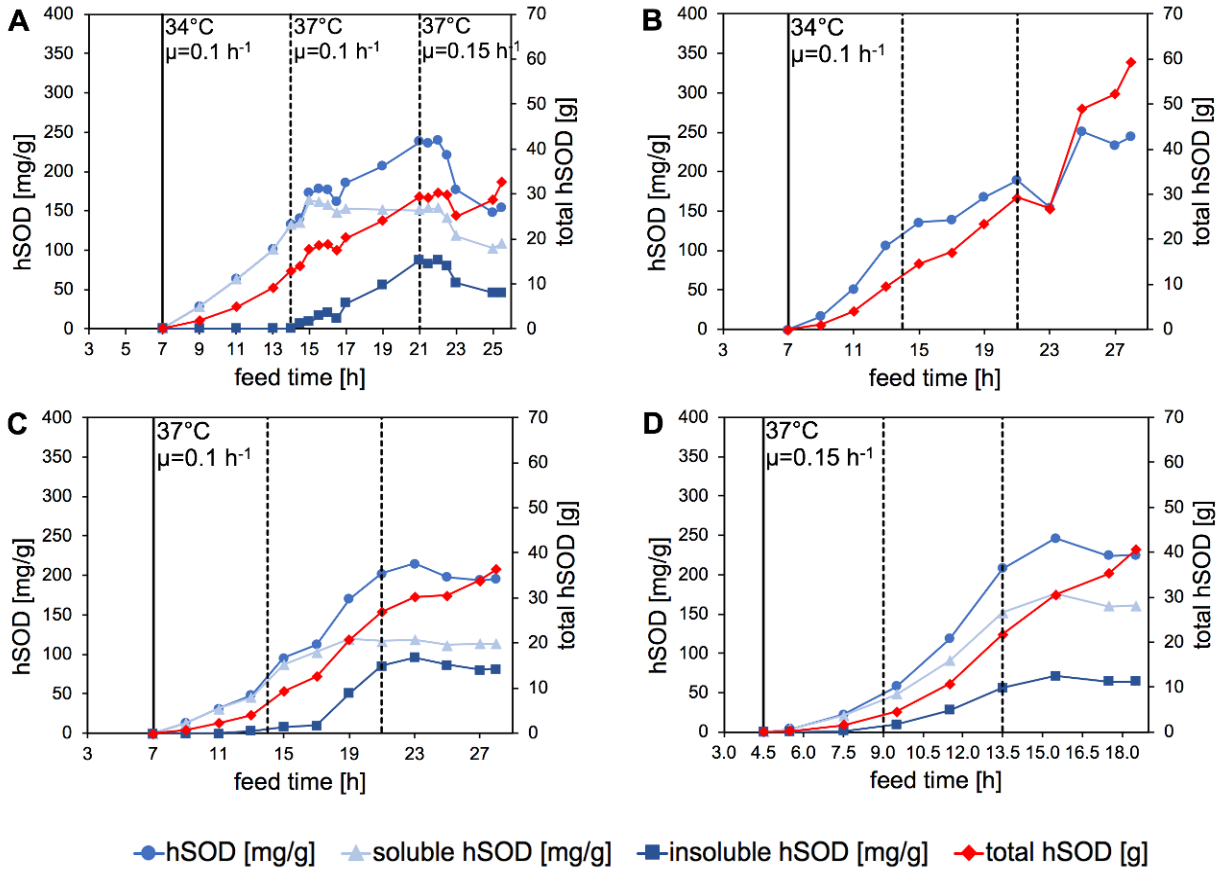

*Figure 20: Comparison of the course of the product titer of the intensified experiment SOD108 (**A**) to the respective static fermentations: SOD98 (**B**), SOD92 (**C**), SOD49 (**D**). The specific product in mg/g is marked as a blue line and the total product rate in g is displayed as red line. The ratio of soluble (light blue line) and insoluble hSOD (dark blue line) is displayed as sperate curves. All experiments were inducted with 0.9 μmol IPTG/g CDM after one generation, which is displayed with the black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

In the first setpoint of SOD108 (Fig.20 **A**), the product titer increased continuously and values of 131.5 mg/g and 13 g of total hSOD were obtained at the first shift of parameters (14 h). In the respective static fermentation, SOD98 (Fig.20 **B**), values of 120.7 mg/g and 12 g of total hSOD were reached. In setpoint two of SOD108 (Fig.20 **A**), after a lag phase due to the shift, the total hSOD increased, while the specific hSOD decreased. In the associated static fermentation SOD92 (Fig.20 **C**) showed a similar behaviour in both curves. In static fermentation SOD49 (Fig.20 **D**), respective for the third setpoint, the trend of the product titer (except the total hSOD) decreased in the last generation,

which was also observed in SOD108. IBs were formed in setpoint two and three of SOD108, which was also obtained in the respective static fermentations, SOD92 and SOD49 (Fig.20 **C+D**).

4.2.3.2 SOD113

The second experiment carried out at the induction strength 0.9 µmol IPTG/g CDM was SOD113. It started with a temperature of 34°C and µ=0.2 h^{-1} and was induced at a feed time of 4.5 h. The course of the growth behaviour of CDM of SOD113 and its respective fermentations is shown in Figure 21.

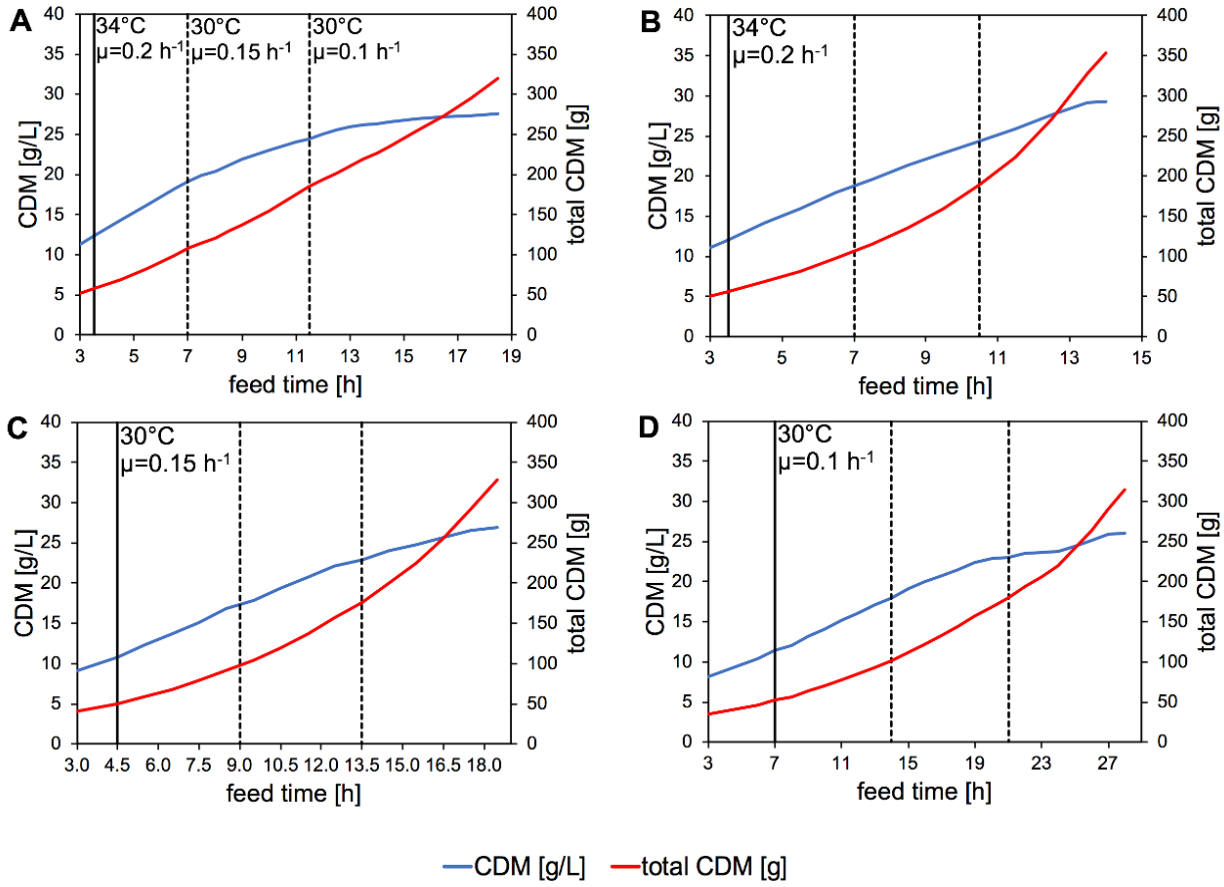

*Figure 21: Comparison of the course of the CDM of intensified experiment SOD113 (**A**) to the respective static fermentations: SOD100 (**B**), SOD60 (**C**), SOD79 (**D**). The blue line shows the course of the CDM concentration [g/L] and the red line displays the total CDM [g]. All experiments were induced with 0.9 µmol IPTG/g CDM after one generation, displayed as black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

At the first shift of parameters (7 h) of SOD113 (Fig.21 **A**) a concentration of 19.2 g/L and total CDM of 108 g was reached. Similar values were achieved at 7 h feed time in the respective static fermentation for the first set point, SOD100 (Fig.21 **B**), values of 18.8 g/L and 107 g. The curve in the second setpoint of SOD113 showed a comparable behaviour like the one in the respective generation in the static fermentation SOD60 (Fig.21 **C**). Due to the lower growth rate in setpoint three of the intensified experiment, the curve of the CDM concentration reached a plateau while the total CDM still increased, which was also likewise observed in the last generation of the static experiment SOD79 (Fig.21 **D**).

The course of the product titer of SOD113 and the respective static experiments is shown in Figure 22.

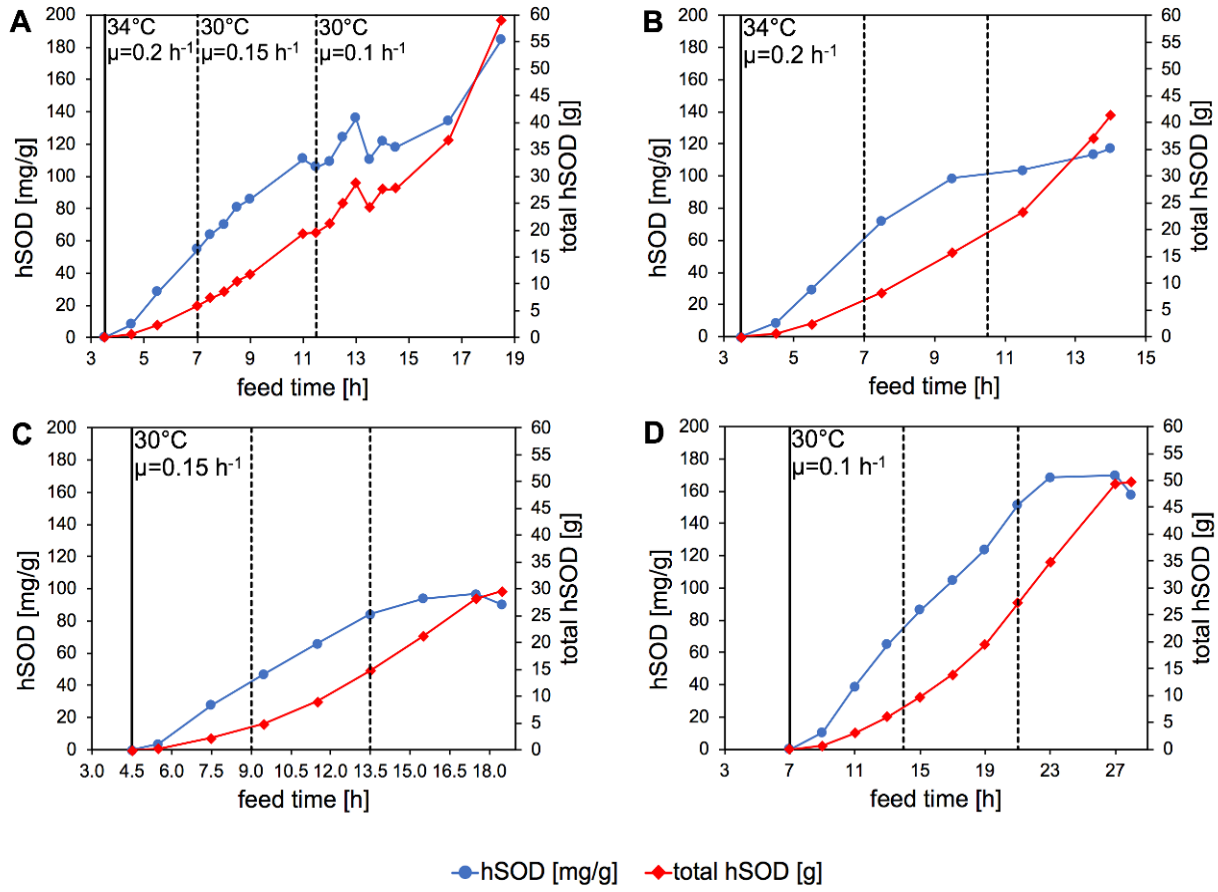

*Figure 22: Comparison of the course of the product titer of the intensified experiment SOD113 (**A**) to the respective static fermentations: SOD100 (**B**), SOD60 (**C**), SOD79 (**D**). The specific product in mg/g is marked as a blue line and the total product rate in g is displayed as red line. The ratio of soluble (light blue line) and insoluble hSOD (dark blue line) is displayed as sperate curves. All experiments were inducted with 0.9 µmol IPTG/g CDM after one generation, which is displayed with the black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

In this intensified experiment, as well as in the respective static fermentations no IBs were formed. In setpoint one and two of SOD113 (Fig.22 **A**), the product titer increased continuously until the second shift of parameters. A value of 54.6 mg/g, i.e., 5.9 g of hSOD was produced at feed time 7 h (first shift), while in the associated static experiment SOD100 (Fig.22 **B**) 50.8 mg/g and 5 g hSOD were obtained at 7h. The respective generations of the associated static fermentations to the first and second setpoint, SOD100 (Fig.22 **B**) and SOD60 (Fig.22 **C**), showed a similar increase of the product titer as in the intensified experiment. In setpoint three of SOD113, after a lag phase, the product titer increased while fluctuations were visible (Fig.22 **A**). In static fermentation SOD49 (Fig20 **D**), respective for the third setpoint, the course of the total product titer increased while the specific hSOD reached a plateau.

4.2.3.2 SOD115

The CPPs at the beginning of the intensified experiment SOD115 were 37°C and a set growth rate of μ=0.2 h^{-1}. It was performed with an induction of 0.9 µmol IPTG/g CDM at a feed time of 3.5 h. The course of the CDM curve of SOD115 and its respective static fermentations is shown in Figure 23.

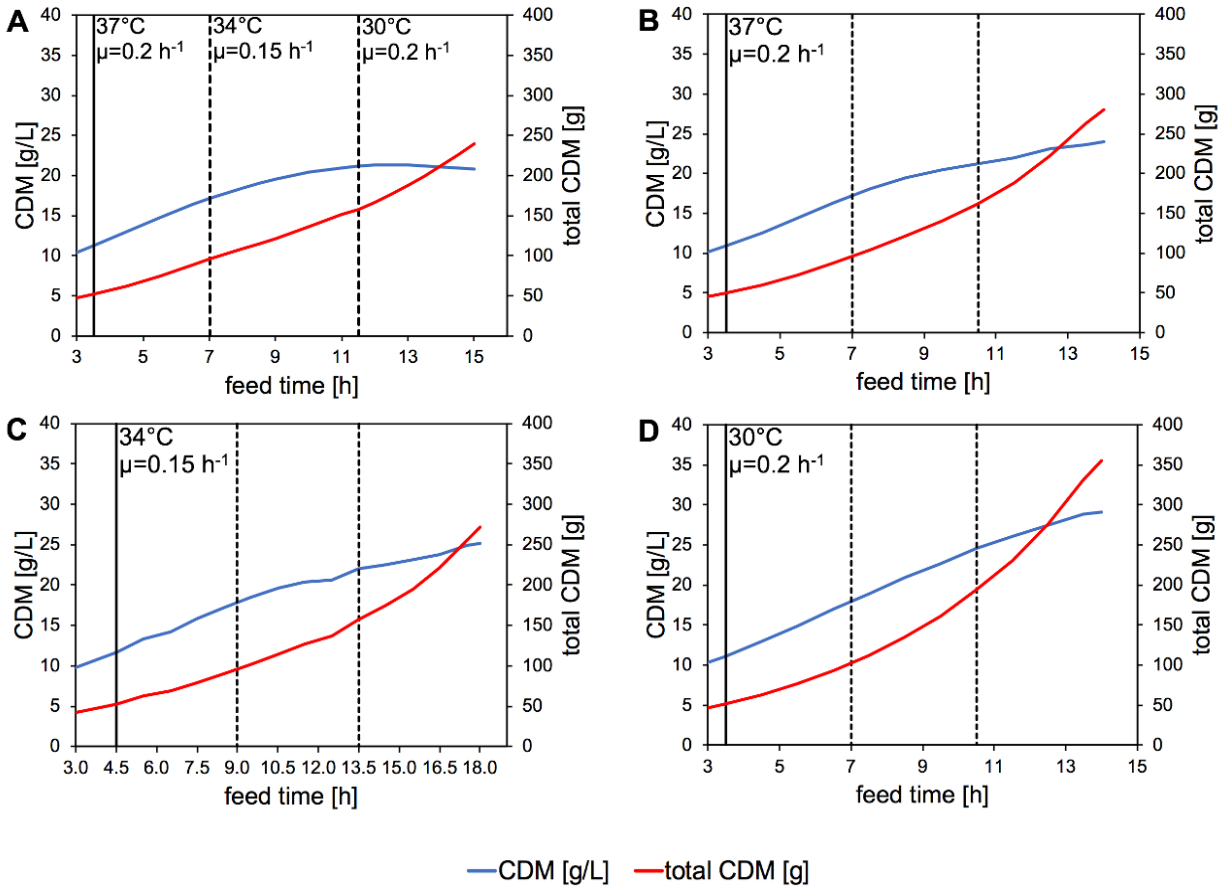

*Figure 23: Comparison of the course of the CDM of intensified experiment SOD115 (**A**) to the respective static fermentations: SOD56 (**B**), SOD93 (**C**), SOD59 (**D**). The blue line shows the course of the CDM concentration[g/L] and the red line displays the total CDM [g]. All experiments were induced with 0.9 µmol IPTG/g CDM after one generation, displayed as black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

A concentration of 17.2 g/L CDM, i.e., 96 g total CDM was reached at the first shift of parameters (7 h) in SOD115 (Fig.23 **A**). The associated static fermentation SOD56 (Fig.23 **B**) reached the same CDM concentration and in total. The static experiment SOD93 (Fig.23 **C**), respective to the second setpoint of intensified experiment SOD115, showed similar behaviour in its to the intensified experiment associated phase. In the last setpoint of SOD115, the CDM concentration slightly decreased, while in the static experiment SOD59 (Fig.23 **D**), which represents the third setpoint, the concentration still increased.

The course of the product titer of SOD115 and its respective static experiments is displayed in Figure 24.

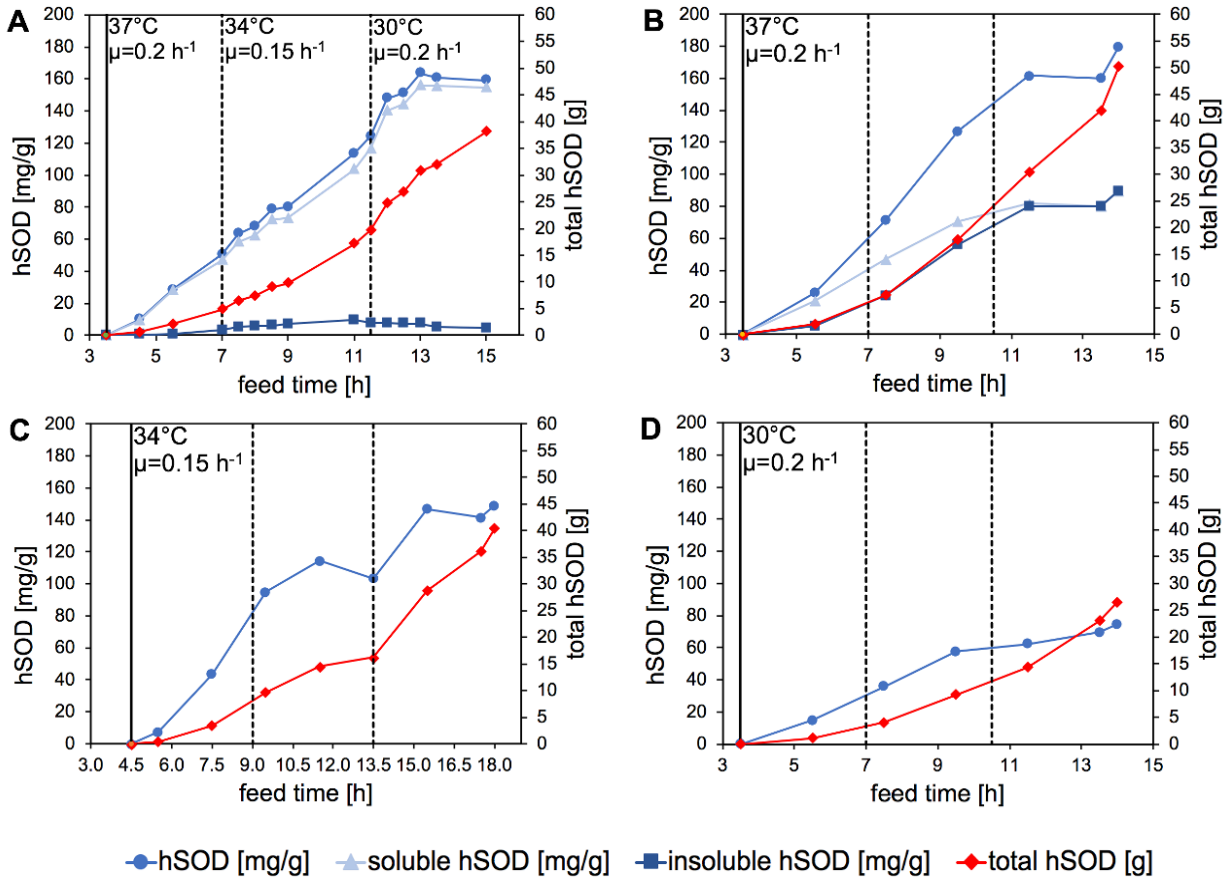

*Figure 24: Comparison of the course of the product titer of the intensified experiment SOD115 (**A**) to the respective static fermentations: SOD56 (**B**), SOD93 (**C**), SOD59 (**D**). The specific product in mg/g is marked as a blue line and the total product rate in g is displayed as red line. The ratio of soluble (light blue line) and insoluble hSOD (dark blue line) is displayed as sperate curves. All experiments were inducted with 0.9 µmol IPTG/g CDM after one generation, which is displayed with the black line. The shifts in **A** and consequently the start of a new generation (**B-D**) are marked as dotted lines.*

In the first setpoint of SOD115 (Figure 24 **A**), at 7 h feed time, first shift of parameters, 50.4 mg/g, i.e., 5 g hSOD were produced. In the respective static fermentation, SOD56 (Fig.24 **B**), 48.8 mg/g, i.e., 5 g hSOD were obtained. In the second setpoint of SOD115, the product titer increased nearly linearly. In SOD93 (Fig.24 **C**), respective to setpoint two, the curve of hSOD increased, except the value at feed time 13.5 h (a possible outliers). In the last setpoint of SOD115, the total product titer had a considerable increase while the curve of the hSOD concentration reached a plateau. A similar behaviour was obtained in SOD59 (Fig.24 **D**), the respective static experiment to the last setpoint of SOD115. The formation of IBs was obtained in setpoint one of SOD115 and was observable during the further setpoints. In the respective static experiment SOD56 the formation of IBs was detected as well.

4.2.4 Response to CPP settings

Table 6 gives an overview of the analytically obtained endpoint values in the iDoE. The experiments were separated into the induction planes of 0.2, 0.5 and 0.9 µmol IPTG/g CDM. The values of the CDM (concentration and in total) and product titer (specific, in total and formed as IB) are displayed.

Table 6: Overview of the endpoint values of the iDoE, split up in the three different induction planes: 0.2, 0.5, 0.9 µmol IPTG/g CDM. The different setpoints (temperature, target growth rate) for each experiment and the results at the endpoint observed for CDM (concentration and in total) and product titer (specific, in total and IB) are displayed.

Induction Strength [µmol IPTG/g CDM]	FermCode	Temperature [°C]	Specific Growth Rate [h^{-1}]	CDM [g/L]	total CDM [g]	hSOD [mg/g]	total hSOD [g]	IB [%]
0.2	SOD114	37 -> 34 -> 34	0.1 -> 0.2 -> 0.1	31.6	369	57.9	21.4	0
	SOD111	37 -> 30 -> 34	0.15 -> 0.1 -> 0.15	31.5	351	70.3	24.6	0
	SOD117	30 -> 37 -> 30	0.2 -> 0.2 -> 0.15	33.5	391	38.6	15.1	0
0.5	SOD116	30 -> 30 -> 34	0.15 -> 0.2 -> 0.2	31.4	360	75.4	27.1	0
	SOD112	30 -> 34 -> 37	0.1 -> 0.1 -> 0.2	20.5	240	136.5	32.8	11
	SOD110	34 -> 37 -> 37	0.15 -> 0.15 -> 0.1	23.3	259	214.4	55.5	27
0.9	SOD115	37 -> 34 -> 30	0.2 -> 0.15 -> 0.2	20.9	240	159.0	38.1	3
	SOD108	34 -> 37 -> 37	0.1 -> 0.1 -> 0.15	18.6	213	141.4	32.6	29
	SOD113	34 -> 30 -> 30	0.2 -> 0.15 -> 0.1	27.6	319	184.5	58.9	0

With respect to the CDM (Table 6), the experimental values, obtained at each endpoint of the individual iDoE fermentations, decreased towards higher induction strength. In contrast, the higher the induction strength, the more the product titer (specific and in total) increased. The formation of IBs was observed in those experiments performed with induction strength of 0.5 and 0.9 µmol IPTG/g CDM and the temperature of 37°C.

Table 7 (p. 50) contains the analytically obtained endpoint values for the CDM (concentration and in total) and product titer (specific, in total and IB) in the DoE fermentations. As well as in Table 6, the experiments were separated into the induction planes 0.2, 0.5, 0.9 µmol IPTG/ g CDM.

Table 7: Overview of the endpoint values of the DoE, split up in the three different induction planes: 0.2, 0.5, 0.9 µmol IPTG/g CDM. The different setpoints (temperature, target growth rate) for each experiment and the results at the endpoint observed for CDM (concentration and in total) and product titer (specific, in total and IB) are displayed.

Induction Strength [µmol IPTG/g CDM]	FermCode	Temperature [°C]	Specific Growth Rate [h^{-1}]	CDM [g/L]	total CDM [g]	hSOD [mg/g]	total hSOD [g]	IB [%]
0.2	SOD94	30	0.1	33.2	386	32.6	12.6	0
	SOD99		0.15	34.3	402	35.9	14.4	0
	SOD102		0.2	34.5	417	19.2	8.0	0
	SOD106	34	0.1	31.1	363	130.6	47.4	0
	SOD97		0.15	32.1	358	59.5	21.3	0
	SOD103		0.2	33.7	404	28.3	11.4	0
	SOD104	37	0.1	30.3	353	57.5	20.3	0
	SOD95		0.15	29.7	328	66.6	21.9	0
	SOD107		0.2	32.9	395	34.9	13.8	0
0.5	SOD63/68	30	0.1	29.8	359	137.6	49.5	0
	SOD80		0.15	31.7	356	81.7	29.1	0
	SOD61/66		0.2	31.5	387	49.4	19.1	0
	SOD96	34	0.1	24.0	277	170.7	47.3	0
	SOD101		0.15	28.7	335	161.7	54.1	0
	SOD105		0.2	31.0	372	106.2	39.5	0
	SOD47	37	0.1	19.9	228	255.9	58.2	33
	SOD53		0.15	24.1	277	194.7	54.0	37
	SOD58		0.2	28.6	349	118.9	41.5	29
0.9	SOD79	30	0.1	26.1	315	157.7	49.7	0
	SOD60		0.15	26.9	328	90.3	29.6	0
	SOD59		0.2	29.1	355	74.7	26.5	0
	SOD98	34	0.1	20.7	242	244.9	59.3	0
	SOD93		0.15	25.2	272	149.0	40.5	0
	SOD100		0.2	29.3	353	117.4	41.4	0
	SOD42/44/91	37	0.1	16.3	190	199.6	38.0	44
	SOD49		0.15	21.6	252	160.5	40.5	40
	SOD56		0.2	24.0	280	179.8	50.3	50

In the static experiments of the full factorial DoE, the endpoint values of the CDM concentration and in total decreased towards higher induction strengths (Table 7). The product titer (specific and in total) showed increasing endpoint values when the experiment was performed with a higher induction strength. Concluding, higher values were obtained in the fermentation with an induction strength of 0.9 µmol IPTG/g CDM than in those performed with 0.2 µmol IPTG/g CDM. The formation of IBs was only observed in the fermentations with an induction strength of 0.5 and 0.9 µmol IPTG/g CDM, and only with the set temperature of 37°C.

These endpoint values of the iDoE and DoE fermentations in Table 6 and Table 7 could not be compared directly, due to the shift of parameters in the iDoE and consequently varying process conditions during each experiment. However, in both, iDoE and DoE, the induction strength had a similar direct influence on the amount of CDM and product titer at the endpoints (Table 6 and 7), i.e., the CDM decreased and the product titer increased with a higher concentration of the inducer. Therefore, the iDoE experiments with the induction strength of 0.9 µmol IPTG/g CDM displayed the lowest values of the CDM and the highest product titer at the endpoints, compared to the experiments in the induction plane 0.2 and 0.5 µmol IPTG/g CDM. For the iDoE, the formation of IBs also was only observed with an induction strength of 0.5 and 0.9 µmol IPTG/g CDM and at 37°C.

4.2.5 Overview of the trends for the iDoE and DoE experiments

As a summary overview of all trends of the iDoE and the DoE, the values of the CDM concentration and in total are displayed in Figure 25 (p. 52). The experiments were separated in the three different growth rates, 0.1 h^{-1}, 0.15 h^{-1} and 0.2 h^{-1}. The course of the CDM concentration and in total, of all intensified experiments are shown in panel **A** and **B**. Panel **C**-**H** show the respective static experiments of the full factorial DoE, separated in the three growth rates.

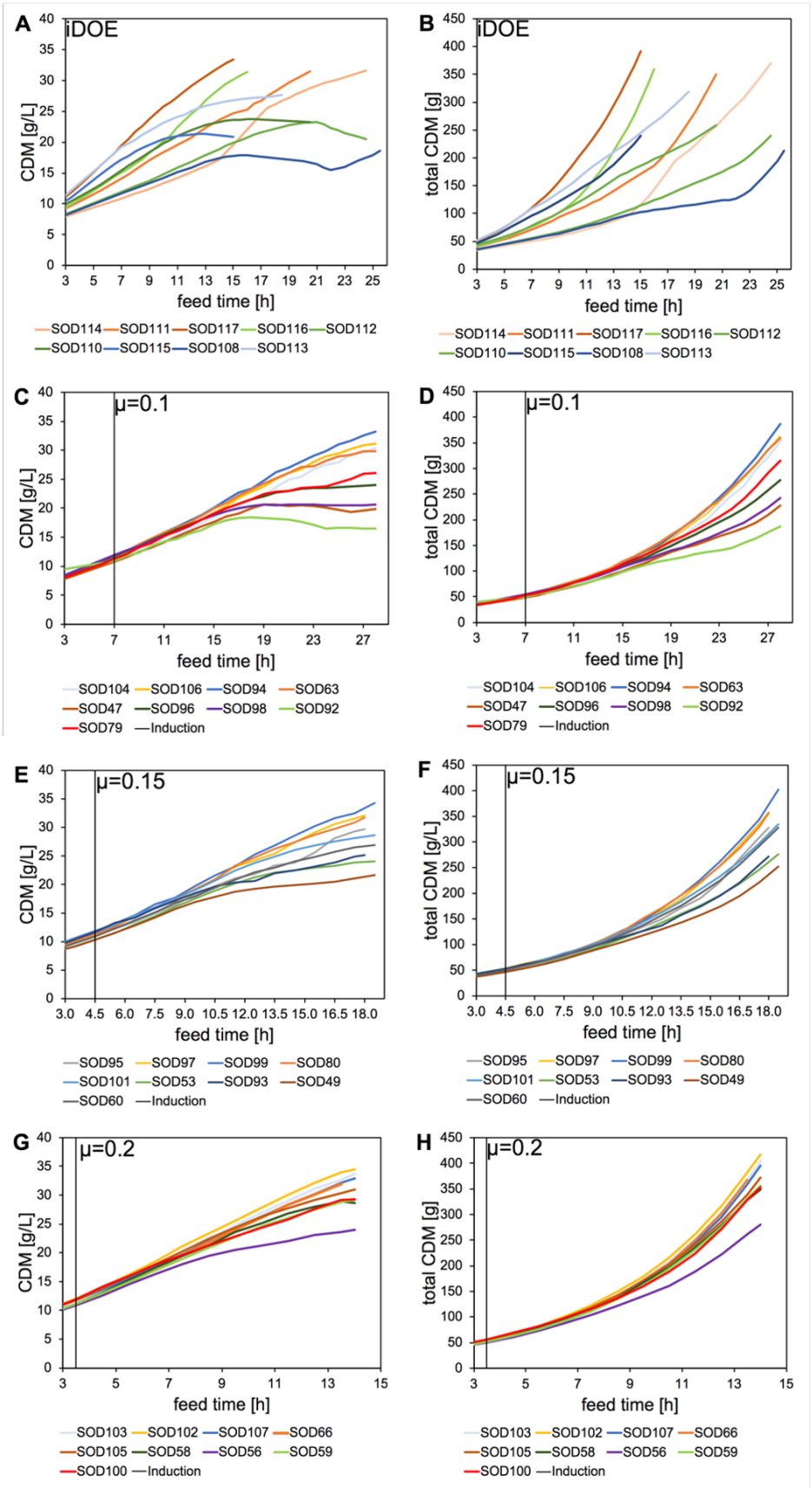

*Figure 25: Overview of the trends of CDM concentration and in total of the iDoE (**A-B**) and the DoE (**C-H**), subdivided in the three growth rates, 0.1 h^{-1}, 0.15 h^{-1}, 0.2 h^{-1}. The induction in **C-H** is marked as black line. Each fermentation represented is displayed in a different colour.*

In Figure 26 (p. 54) the course of the specific product titer and the total product, of all intensified experiments of the iDoE are shown in panel A and B. Panel C-H show the respective static experiments of the full factorial DoE, separated into the different growth rates.

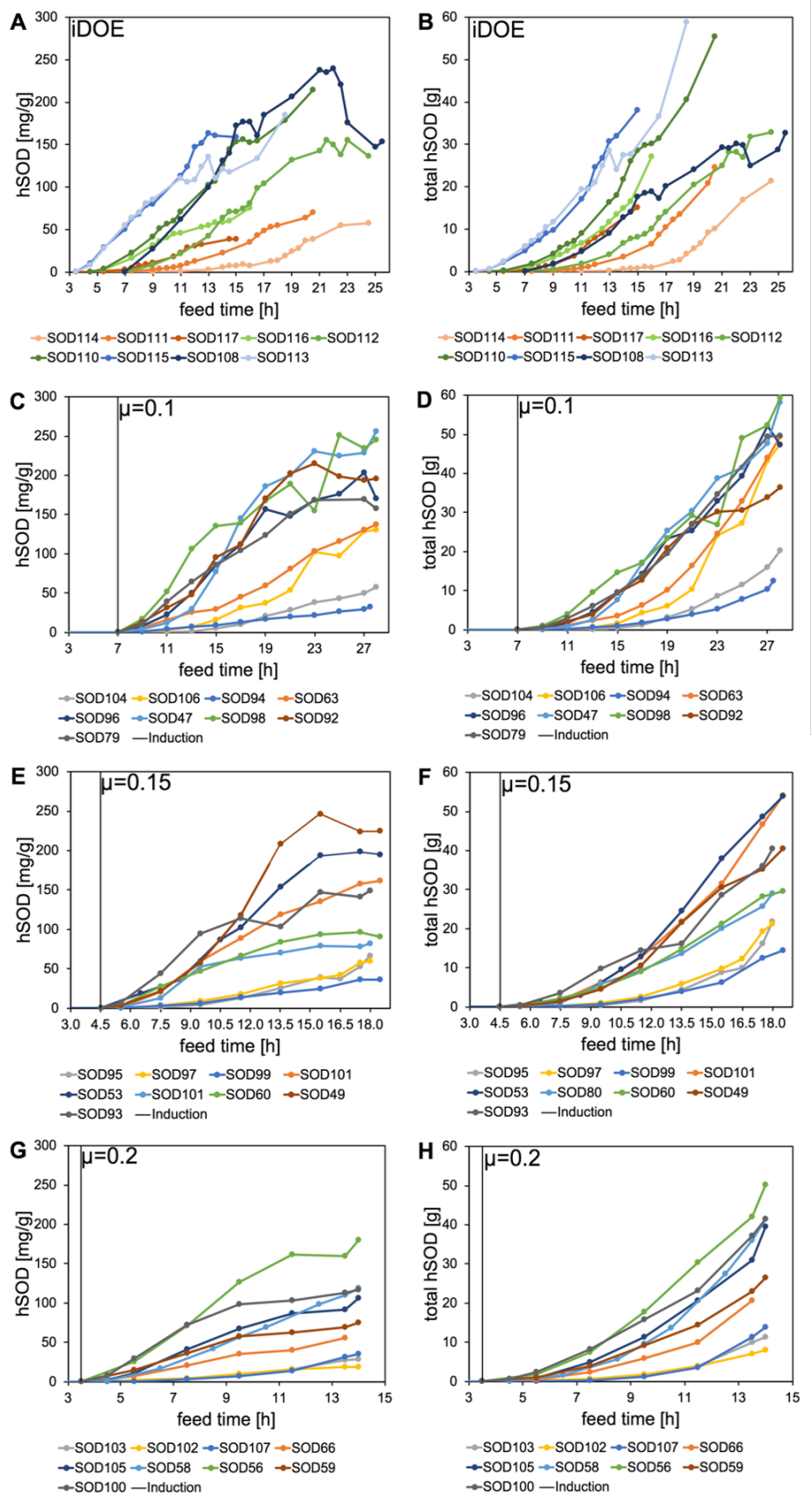

*Figure 26: Overview of the trends of specific product titer and in total (soluble and insoluble fraction) of the iDoE (**A-B**) and the DoE (**C-H**), subdivided in the three growth rates, 0.1 h^{-1}, 0.15 h^{-1}, 0.2 h^{-1}. The induction in **C-H** is marked as black line. Each fermentation represented is displayed in a different colour.*

By comparing the trends of the CDM and the product for the nine intensified fermentations to the static experiments in detail, no significant difference in response to the CPP setting was obtained (Section 4.2.). All the intensified fermentations display a similar trend for the CDM as well as the recombinant product, compared to the respective phase in the static fermentations. After a lag phase, i.e., a short adaptation phase after the shift of parameters, the trend of the current CPP setpoint adapted to the trend observed in the associated static fermentation. This lag phase is clearly visible, for the product titer. Regarding the response to the CPP settings of the CDM and the product, both, iDoE and DoE fermentations, had the same pattern in increasing and decreasing endpoint values due to the three different induction concentrations (0.2, 0.5, 0.9 µmol IPTG/g CDM). The endpoint values could not be compared directly due to the changing setpoints in the iDoE. Nevertheless, the trend of the CDM and product in the intensified experiments was highly similar to the respective phase in the static fermentation.

Due to the similar behaviour of the cells ins each respective setpoint, a memory effect of the cells due to the intra-experimental CPP shifts and therefore a limitation of the applicability of this iDoE approach was excluded.

Subsequently, to investigate the possibility of using iDoE to generate process models to predict static experiments, a hybrid model was developed by utilizing the generated iDoE data. This iDoE hybrid model should be able to predict the CDM concentration of all static fed-batch fermentations, which would reduce the number of required experiments for model building by two thirds (>66 %). Such a high reduction of experiments and therefore raw materials, time and money would be highly beneficial from an economical point of view. As a reference for the performance of this iDoE hybrid model, a previously developed hybrid model, based on the full factorial static DoE was used.

4.3 iDoE hybrid model

For the development of the iDoE hybrid model, two data sets were assessed. The data set used for training the iDoE hybrid model consisted of the nine iDoE experiments ($iDoE_9$). The historical static data set contained 31 static fed-batch fermentations, received from the full factorial DoE (DoE_{31}). It included 31 experiments which covered all 27 CPPs, the remaining four were two duplicates and one triplicate experiment. This data set was used for external validation (test set). Both data sets consisted of the following process variables for modelling: the analytical CDM concentration (g/L) and the online available process variables, e.g., accumulated feed and base (g), accumulated inducer (mg) stirrer speed (rpm), inlet air (L/min) and temperature (°C).

The hybrid model was developed to accurately predict the CDM concentration during the entire process for all static DoE fermentations of the historical data set. The earlier derived time-resolved hybrid model, trained on the static DoE experiments, was developed to overcome the limitation of only endpoint predictions, ignoring the majority of the process including process deviations, as it is the state of the art in upstream process characterization.

The within this work developed iDoE hybrid model incorporated the process dynamics due to the intra-experimental CPP shifts and thereby also required fewer experiments to characterize the same space was compared to the full factorial hybrid model trained on 3 times the experiments performed in a classical statical way. Figure 27 displays an overview of the bootstrap-aggregated iDoE hybrid model, predicting the CDM, i.e., the nine intensified experiments (training data, Fig.27 **A+B**), as well as all 31 static experiments (test data, Fig.27 **C+D**), are displayed. In the scatter plots, the analytically obtained (measured) values versus the predicted values are shown and the SD of each prediction is indicated. For the time-resolved plots, the analytical values versus the respective prediction from the iDoE hybrid model and the PI for each fermentation are illustrated. Each fermentation is displayed in a different colour.

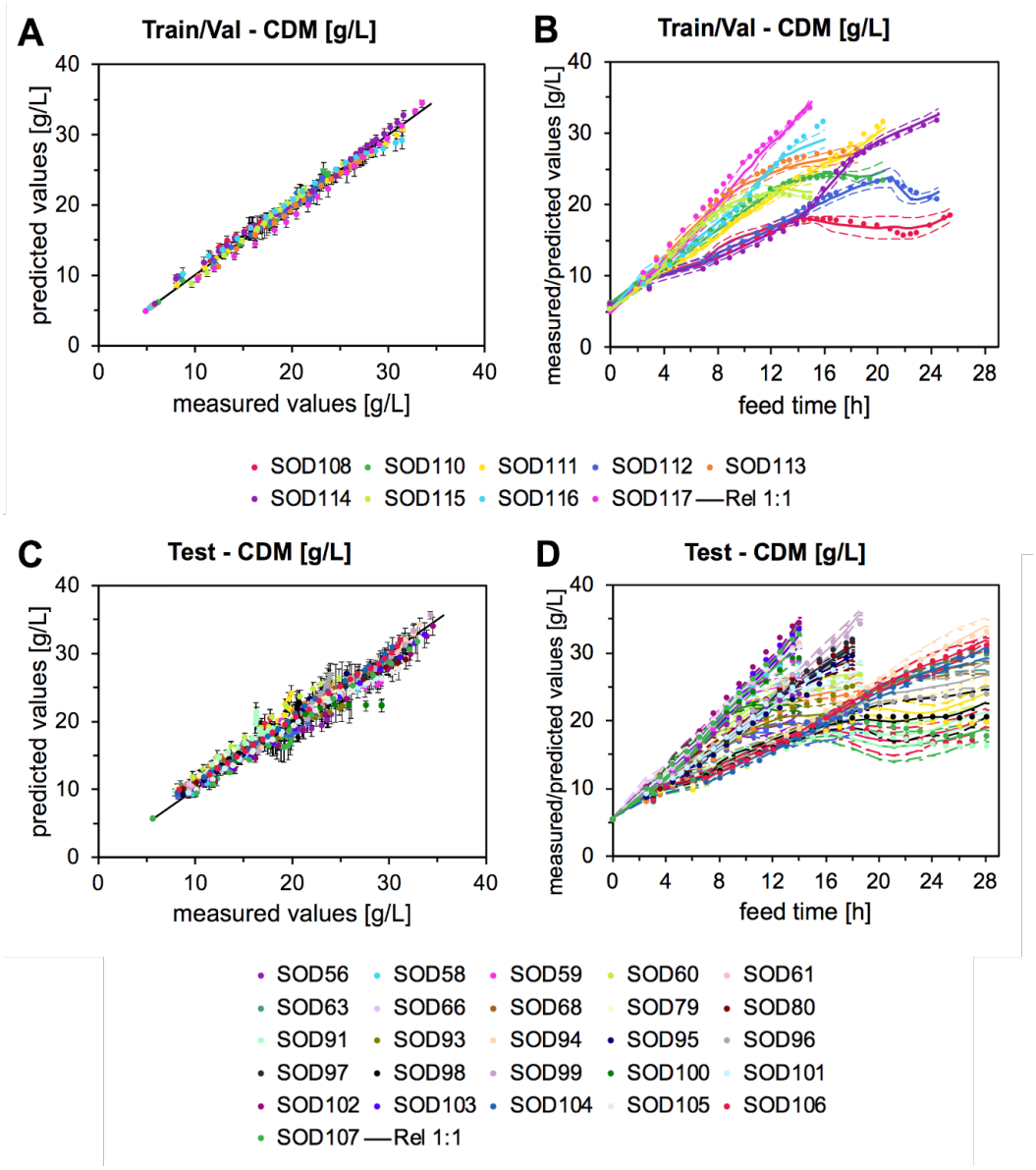

*Figure 27: Overview of the results of the hybrid model for the CDM evolution. Each fermentation is displayed in a different color. Panel **A** and **B** show the results for the training data set illustrated as scatter and a time-resolved plot. In panel **C** and **D** the modelling results for the test dataset (static experiments) is given as scatter and time-resolved plot. In the scatter plots the respective SD is shown and the PI for each experiment is given in the time-resolved plots.*

The bootstrap-aggregated iDoE hybrid model showed an outstanding performance for predicting the CDM. Tight distribution of the PI, due to small SDs, indicates a small risk of a model misprediction, as displayed for the training data (Fig.27 **A**) as well as in the test data (Fig.27 **C**). With respect to the training data, the model struggled with the prediction of SOD108, clearly illustrated in Figure 27 **B** and indicated by the broad PI.

4.3.1 Model performance on the training data

A more detailed illustration of the training/validation data set of the iDoE hybrid model is given in Figure 28. Figure 28 **A-I** shows the measured values at each time point of the intensified fermentations and

the respective prediction of the model. For each prediction, the PI and the calculated RMSE value for each individual fed-batch fermentation are displayed.

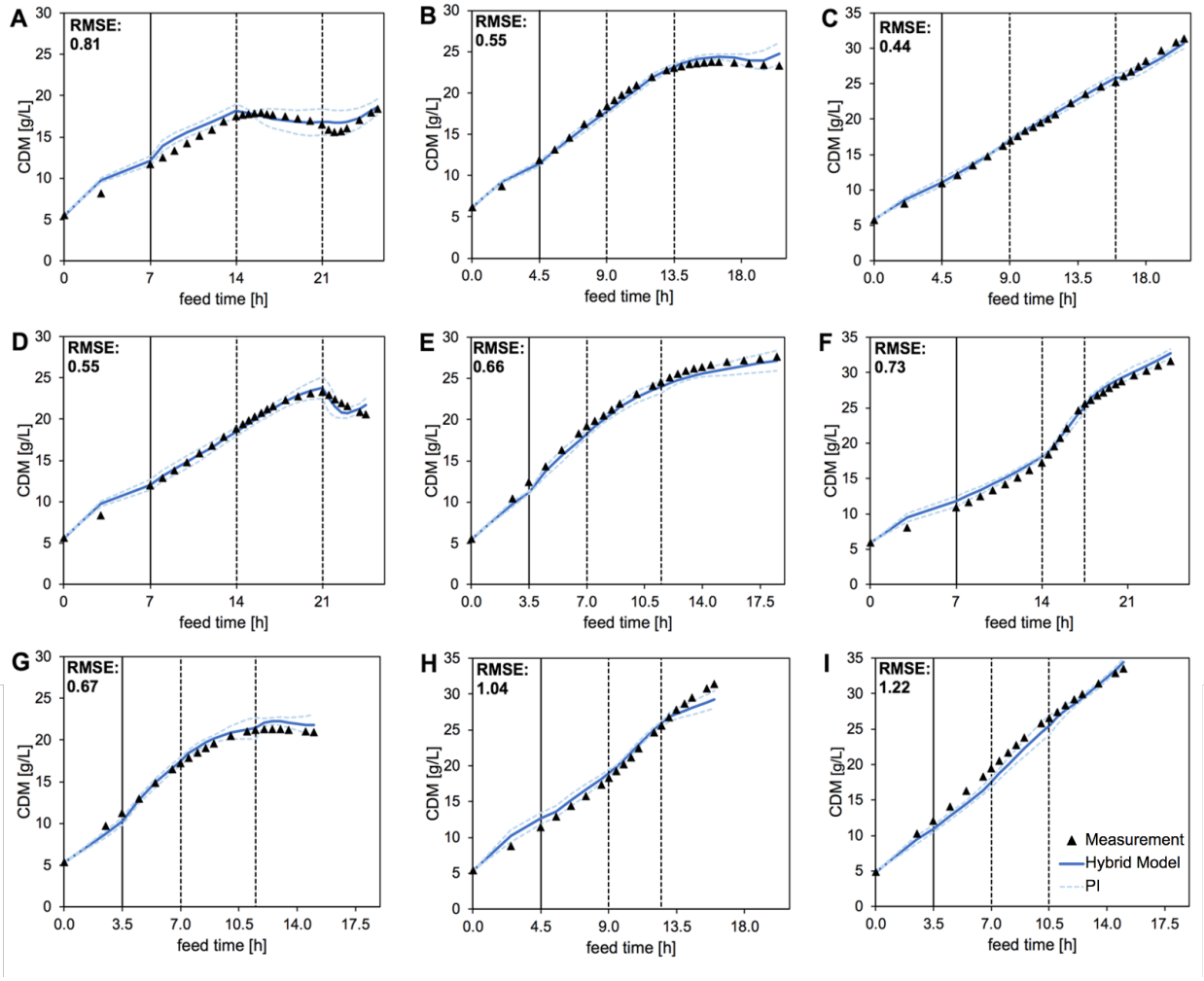

*Figure 28: Training data for the hybrid model of all intensified fermentations: SOD108 (**A**) – SOD117 (**I**). The time-resolved development of the measured values and the modelled CDM concentration are displayed. The RMSE value and the PI for each model are shown in the respective panel. The induction is marked as a continuous black line, while the shift of parameters is labeled as dashed lines.*

By evaluating the performance of the model on the training data, i.e., for each individual intensified fed-batch fermentation, SOD111 (Fig.28 **C**) displayed the best result: an RMSE value of 0.44 g/L, and a small PI. Considering that the intensified experiment SOD111 had a nearly linear trend of the analytically measured values, it was not astonishing that it was rather easy for the hybrid model to predict the respective values. However, in respect to the other experiments, the adaption of the analytical values of the CDM to the parameter shifts and the hereby increased process dynamic are clearly observable, e.g., in SOD108 (Fig.28 **A**) and SOD114 (Fig.28 **F**). In Figure 28 **A-I** the model predicted the CDM trends highly accurate and with small PIs. Except for SOD108 (Fig.28**A**), which showed a broad PI. However, the RMSE of 0.81 g/L is on average similar to the others.

Summarising, the training data set is accurately predicted by the model, displaying only slight deviations from the measured values. The calculated f R^2, RMSE and the error are displayed in Table 8 (p. 63/64). The generated iDoE hybrid model was applied to the test set DoE_{31} to evaluate the performance on predicting unknown data.

4.3.2 Model performance on the test data

The results of the iDoE hybrid model applied to new data are displayed in the following Figures (Fig.29-31). The different static fermentations were separated into the three growth rates 0.1 h^{-1} (Fig.29), 0.15 h^{-1} (Fig.30) and 0.2 h^{-1} (Fig.31). For each experiment, the measured analytical values, the prediction of the iDoE hybrid model as well as the PI and the RMSE are presented in a time-resolved plot.

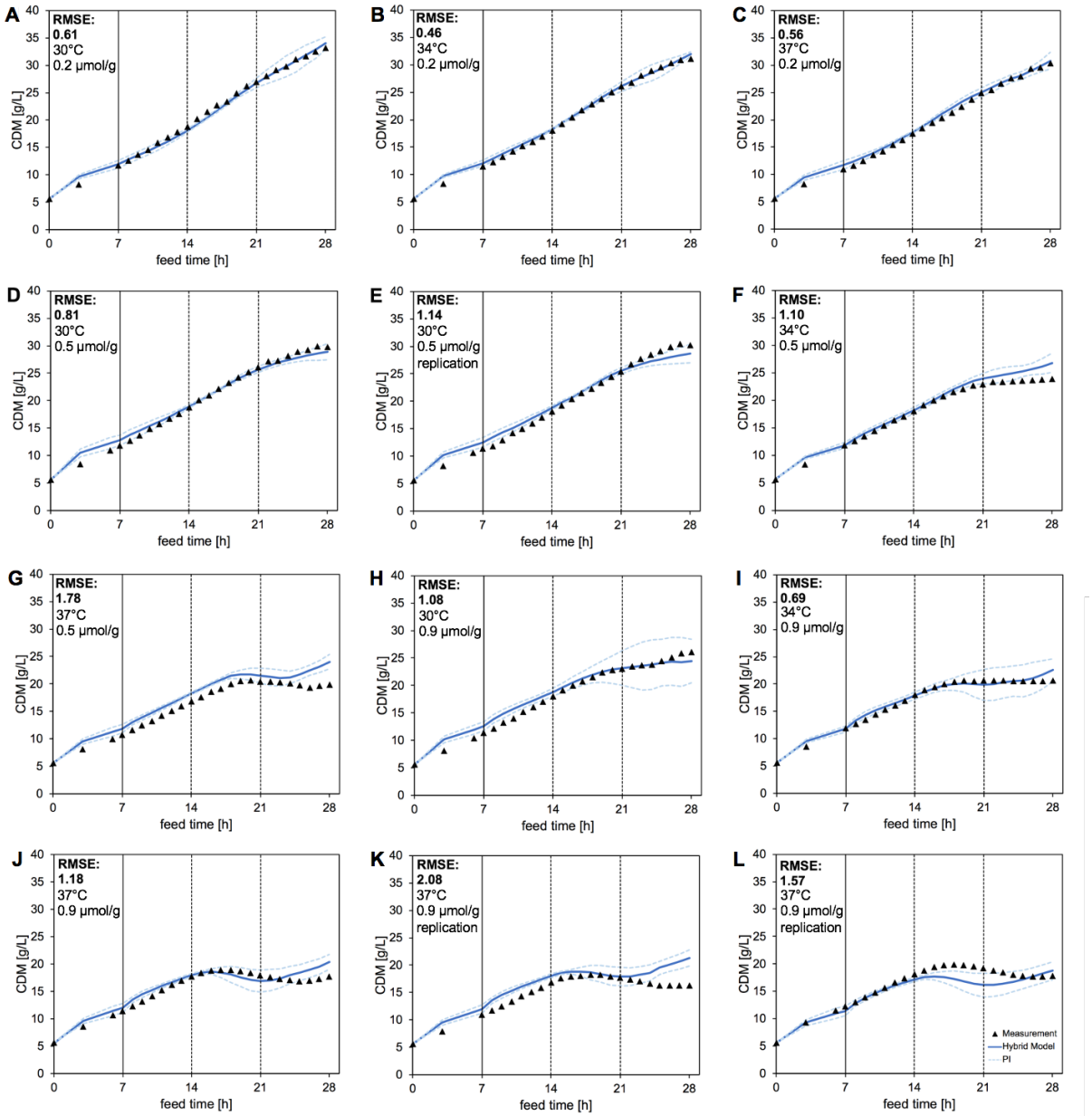

*Figure 29: Results of the hybrid model applied to the test set DoE_{31}. All experiments of the test set with a growth rate of 0.1 h^{-1} are displayed: SOD94 (**A**), SOD106 (**B**), SOD104 (**C**), SOD63 (**D**), SOD68 (replication of SOD63) (**E**), SOD96 (**F**), SOD47 (**G**), SOD79 (**H**), SOD98 (**I**), SOD42 (**J**), SOD91 (replication of SOD42) (**K**), SOD44 (replication of SOD42) (**L**). The time-resolved development of the measured values and the modelled CDM concentration are displayed. The RMSE value, the temperature (°C), the induction strength (in µmol/g) and the PI for each model are shown in the respective plot. The induction is marked as a continuous black line, while the dashed lines label the start of a new generation.*

Regarding the performance of the iDoE hybrid model for each individual fed-batch fermentation of the test set with a growth rate 0.1 h^{-1}, SOD94 (Fig.29 **A**), SOD104 (Fig.29 **C**) and SOD106 (Fig.29 **B**) showed the best results, displaying an RMSE value of 0.61 g/L, 0.56 g/L and 0.46 g/L and a small PI. For fermentation SOD47 (Fig.29 **G**) and SOD91 (Fig.29 **K**) the model predictions displayed a slight deviation from the analytical values during the end of the process, which is reflected in the RMSE values of

1.78 g/L and 2.08 g/L. A broad PI was obtained in SOD79 (Fig.29 **H**) and SOD98 (Fig.29 **I**) in the respectively last generation.

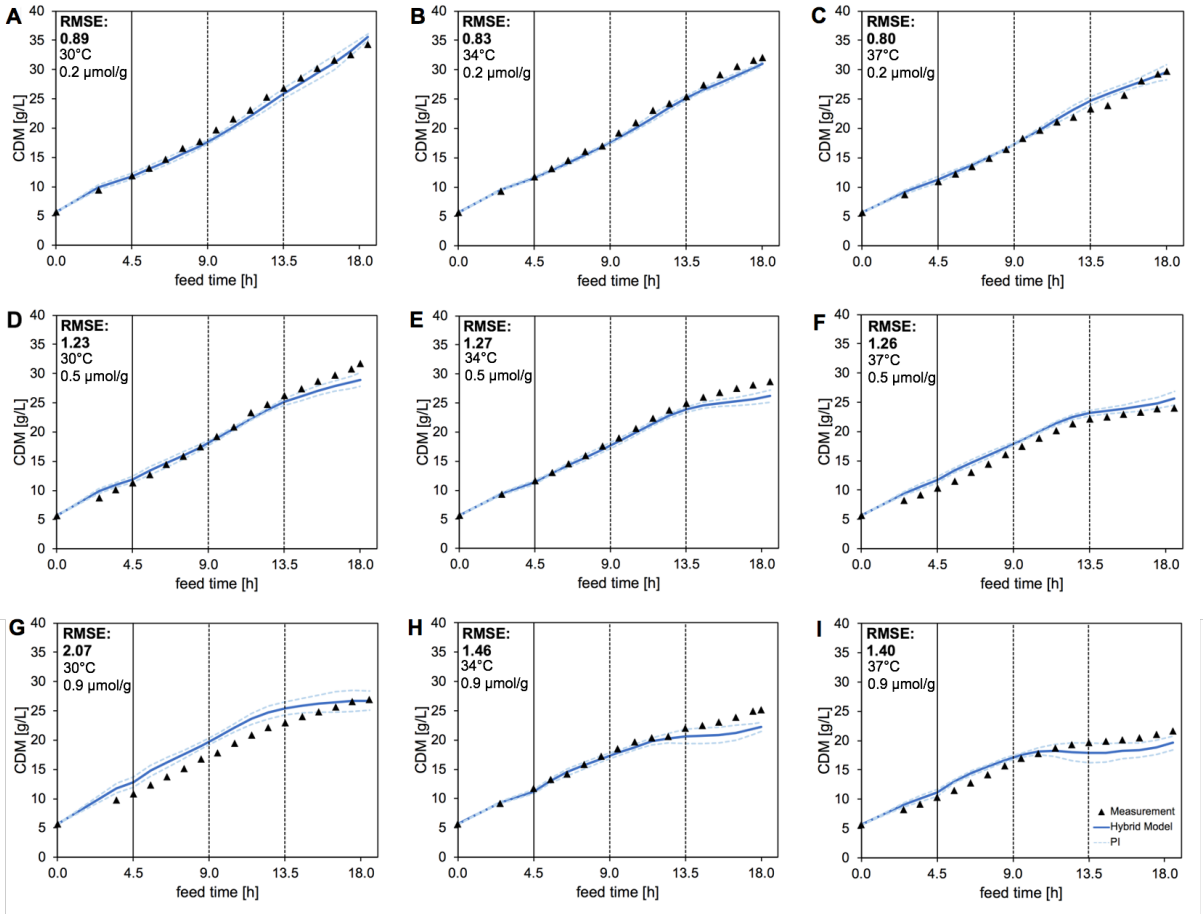

*Figure 30: Results of the hybrid model applied to the test set DoE_{31}. All experiments of the test set with a growth rate of 0.15 h^{-1}.are displayed: SOD99 (**A**), SOD97 (**B**), SOD95 (**C**), SOD80 (**D**), SOD101 (**E**), SOD53 (**F**), SOD60 (**G**), SOD93 (**H**), SOD49 (**I**). The time-resolved development of the measured values and the modelled CDM concentration are displayed. The RMSE value, the temperature (°C), the induction strength (in µmol/g) and the PI for each model are shown in the respective plot. The induction is marked as a continuous black line, while the dashed lines label the start of a new generation.*

Predicting the static experiments with the growth rate 0.15 h^{-1} the model performed the best in SOD95 (Fig.30 **C**) and SOD97 (Fig.30 **B**). RMSE values of 0.80 g/L and 0.83 g/L are displayed as well as small PIs. For fermentation SOD60 (Fig.30 **G**), the model struggled with the prediction of the values, resulting in a high RMSE value of 2.07 g/L.

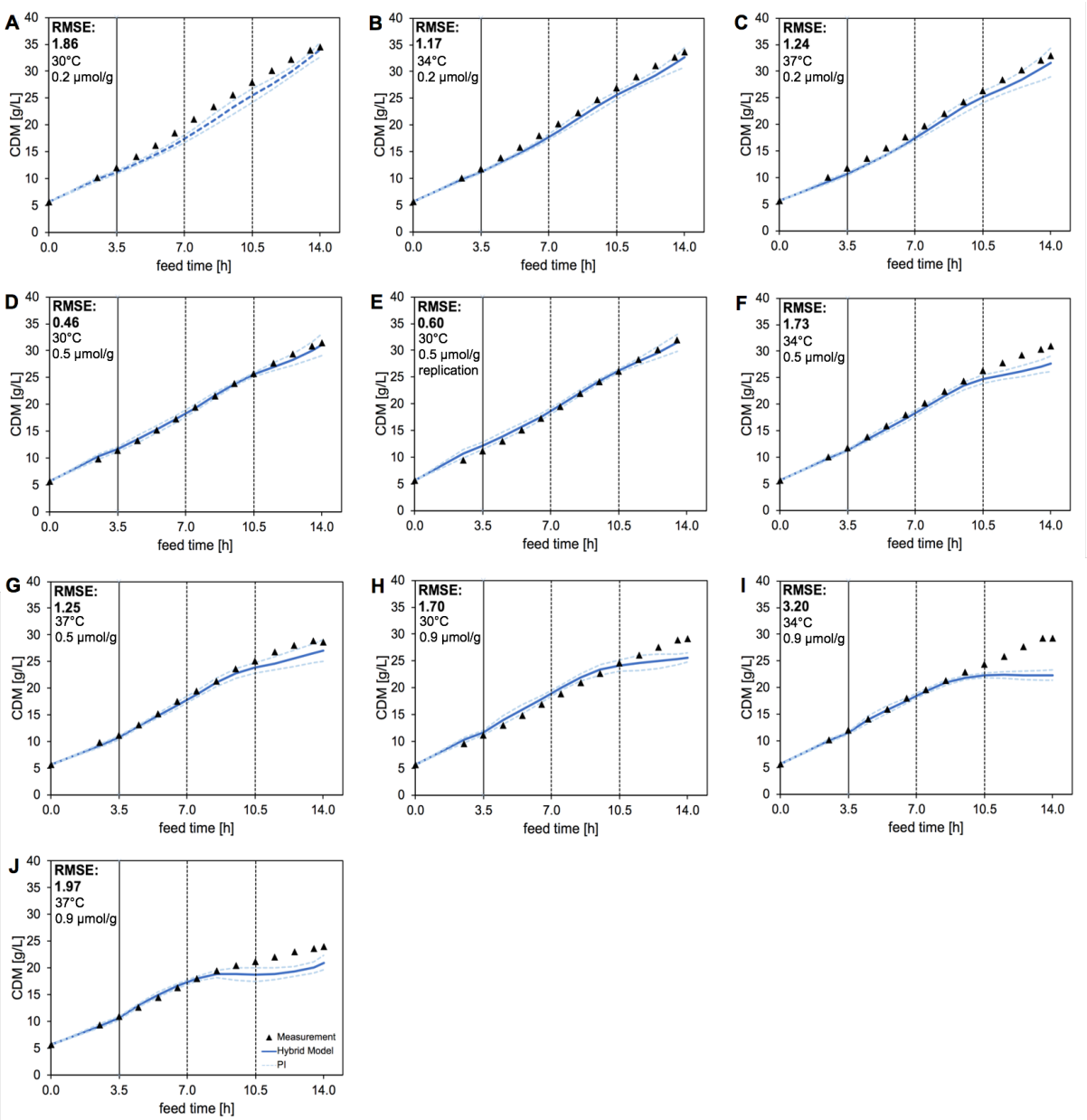

*Figure 31: Results of the hybrid model applied to the test set DoE$_{31}$. All experiments of the test set with a growth rate of 0.2 h^{-1}.are displayed: SOD102 (**A**), SOD103 (**B**), SOD107 (**C**), SOD61 (**D**), SOD66 (replication of SOD61) (**E**), SOD105 (**F**), SOD58 (**G**), SOD59 (**H**), SOD100 (**I**), SOD56 (**J**). The time-resolved development of the measured values and the modelled CDM concentration are displayed. The RMSE value, the temperature (°C), the induction strength (in µmol/g) and the PI for each model are shown in the respective plot. The induction is marked as a continuous black line, while the dashed lines label the start of a new generation.*

Regarding the performance of the iDoE hybrid model for the experiments of the test set with growth rate 0.2 h^{-1}, the best results were obtained in SOD61 (Fig.31 **D**) and SOD66 (Fig.31 **E**). RMSE values of 0.46 g/L and 0.60 g/L are displayed. The model struggled in the prediction of fermentation SOD56 and SOD100 (Fig.31 **J+I**) in the last generation. A high RMSE value of 3.20 g/L was obtained in SOD100.

The calculated results for the RMSE, R^2 and the error of each fed-batch fermentation of the entire test set are displayed in Table 8 (p. 63/64).

Table 8 summarizes the calculated values of the performance criteria, which indicate the quality and accuracy of the developed bootstrap-aggregated iDoE hybrid model. A 10% error limit was set as a threshold value for good model performance. Herein, the fermentations marked green were predicted with an error below 10%. The fed-batch fermentations marked red displayed a calculated error which exceeded the threshold.

Table 8: Calculated results of the hybrid model for the training and test partition. For each fermentation, the respective RMSE, R^2 and the error of the model are indicated. The green boxes indicate the fed-batch fermentations with a good model performance. The red boxes mark the fed-batch fermentations were the predicted values exceed the threshold.

Train/Val data			
RMSE/Error total	0.74 g/L	3.44%	
Fermentation number	RMSE [g/L]	Error [%]	R^2
SOD108	0.81	4.68	0.96
SOD110	0.55	2.43	0.99
SOD111	0.44	1.26	0.99
SOD112	0.55	2.5	0.99
SOD113	0.66	3.05	0.99
SOD114	0.73	3.56	0.99
SOD115	0.67	3.3	0.98
SOD116	1.04	4.78	0.98
SOD117	1.22	5.39	0.99
Test data			
RMSE/Error total	1.27 g/L	5.31%	
Experiments with μ=0.1	RMSE [g/L]	Error [%]	R^2
SOD42	1.18	4.13	0.91
SOD44	1.57	6.43	0.91
SOD47	1.78	9.65	0.97
SOD63	0.81	3.98	0.99
SOD68	1.14	5.77	0.99
SOD79	1.08	5.85	0.99
SOD91	2.08	10.93	0.86
SOD94	0.61	2.77	0.99
SOD96	1.10	3.88	0.99
SOD98	0.69	3.27	0.98
SOD104	0.56	2.99	0.99
SOD106	0.46	2.31	0.99
Experiments with μ=0.15	RMSE [g/L]	Error [%]	R^2
SOD49	1.40	7.78	0.95
SOD53	1.26	7.94	0.99
SOD60	2.07	11.16	0.98
SOD80	1.23	4.62	0.99
SOD93	1.46	4.89	0.98
SOD95	0.80	2.66	0.99
SOD97	0.83	3.06	0.99
SOD99	0.89	3.55	0.99
SOD101	1.27	3.98	0.99
Experiments with μ=0.2	RMSE [g/L]	Error [%]	R^2
SOD56	1.97	6.74	0.96
SOD58	1.25	4.34	0.99
SOD59	1.70	6.32	0.97
SOD61	0.46	1.62	0.99
SOD66	0.60	3.27	0.99
SOD100	3.20	7.58	0.92
SOD102	1.86	7.38	0.99
SOD103	1.17	4.77	0.99

SOD105	1.73	5.33	0.99
SOD107	1.24	5.81	0.99

Except for two fed-batch fermentations, all model predictions displayed results below the threshold SOD60 and SOD91 slightly exceed the threshold value of 10%, with errors of 11.16% and 10.93%.

To sum up, Figure 29-31 show that the model can accurately predict the trends of the measured analytical values of the static test set. However, the range of the PI widened with an increasing feed time. Nevertheless, the calculated performance criteria (RMSE, percentage error and R^2) indicate good quality and accuracy.

4.3.3 Performance comparison of the iDoE and DoE hybrid model

In Table 9, the results obtained with the iDoE hybrid model trained with the iDoE data set ($iDoE_9$) are compared to the results of the hybrid model trained with the full factorial static DoE data set, developed by Bayer et al. (2020). The performance criteria, namely, the RMSE value, the Error and R^2 are shown.

Table 9: Comparison of the results of the hybrid model trained with the iDoE, e.g., the data of nine experiments, and the hybrid model characterized with 31 static fermentations of the DoE. The values for the RMSE, the Error and R^2 are displayed.

Hybrid model	iDoE	DoE
used fermentations	9	31
RMSE [g/L]	1.27	1.10
Error [%]	5.31	4.24
R^2	0.97	0.98

Considering the iDoE hybrid model trained with nine intensified fermentations, nearly identical results to those received with the static DoE hybrid model, trained with 31 experiments, were obtained. The RMSE value differs in 0.17 g/L and for the error, a difference of 1.07% was obtained. With values of 0.97 and 0.98 of R^2, both models showed high quality of matching the analytical results. This comparison shows that the iDoE hybrid model performs as reliable and accurate as the hybrid model trained with the full-factorial static DoE, i.e., the CDM can be predicted with a total error of 5.31% (RMSE 1.27 g/L). This indicates that the highly beneficial iDoE concept for an accelerated process characterization in upstream processing, to reduce this time-consuming task, is possible by developing a hybrid model, based on fewer intensified experiments.

5 CONCLUSION

Since design space characterization is time-consuming and expensive, an innovative approach to accelerate this process task was established. The overall aim of this study was to set up an iDoE for *E. coli* fed-batch fermentations (HMS174), expressing recombinant hSOD and investigate the process dynamic and behaviour compared to the static experiments and potential limitations.

A design space with three CPPs, each at three levels, i.e., 27 CPP combination setpoints, was characterized by utilizing static fed-batch fermentations in an earlier study and a hybrid model predicting the CDM concentration of these processes was developed. Based on this design space, an iDoE, consisting of nine experiments with two CPP shifts during one fermentation was developed and the usability of this approach for accelerated process characterization investigated.

The first consideration was the importance of the direction and the order of the intra-experimental CPP shifts. Therefore, the order of the CPP combination setpoints of the first iDoE experiment was contrary performed in a new experiment. Based on the growth behaviour of the CDM and the product titer, it has been shown that the cells rapidly adapt to the changing parameters. Despite the expected slightly different trends of the CDM and the product, due to the different order of the performed CPP setpoints, the course of the curve of each respective setpoint were highly comparable. As a result, it was assumed that the order of the shift of parameters does not affect the outcome and the iDoE can be executed as planned.

At the end of each executed intensified experiment, the off-line trend and the cell physiology were compared with the associated static experiment of each CPP setpoint. Especially for the product titer trend, the shift of parameters is clearly visible, by means of a short adaption phase due to the new setpoint. Overall, both, the CDM as well as the product, showed comparable results in the trend to each respective static fermentation. Therefore, these investigations show that, using this particular setup, a memory effect of the cells can be excluded, due to the adaptability of the iDoE experiments to the different CPP setpoints.

The iDoE hybrid model with the *Novasign* hybrid modelling toolbox was developed to predict the CDM concentration of the test set. It enabled precise modeling of the CDM concentration for the static fed-batch fermentations in a time-resolved manner. This iDoE hybrid model performed with an overall error of 5.31% (RMSE 1.27 g/L) for predicting the CDM concentration of the test set. This demonstrates that a hybrid model, based on iDoE, is able to reliably and accurately predict the outcome of static fed-batch fermentations.

In comparison, an earlier derived hybrid model, based on the full factorial static DoE, performed with an overall error of 4.24% (RMSE 1.10 g/L) for predicting the CDM concentration. Even though the error in the iDoE hybrid model slightly increased, comparable model performance was maintained and only nine experiments were required to completely characterize the same design space as for the static hybrid model. This demonstrates the potential timesaving by utilizing iDoE and highlights this highly beneficial concept for improved and accelerated upstream process characterization.

In the ongoing investigations, not only the accurate prediction of CDM concentration but also the additional prediction of the product titer by this iDoE hybrid model should be implemented. Thereby, both values could be modeled with this time- and cost-saving iDoE hybrid model in the future.

To potentially gain superior results in possible future iDoE studies, other inputs to train the model or a different model structure should be considered, e.g., one opportunity would be to exclude the base as an input, which would lead to three CPPs as remaining inputs. Using only these three easily controllable CPPs, namely, induction strength, glucose feeding rate and temperature, as inputs to the hybrid model, the possibility to set up a controllable model is enabled. This allows not only accurate predictions but also to interfere with the process to generate a stable quality output with repetitive accuracy, e.g., model predictive control (MPC), in future applications.

6 ABBREVATIONS

ddH_2O	double-distilled water
CDM	Cell Dry Mass
cGMP	Current Good Manufacturing Practice
CHO	Chinese Hamster Ovary
CPPs	Critical Process Parameters
CQAs	Critical Quality Attributes
DoE	Design of Experiment
DO	Dissolved Oxygen
FDA	U.S. Food and Drug Administration
hSOD	recombinant human superoxide dismutase
IB	Inclusion bodies
iDoE	intensified Design of Experiments
IPTG	Isopropyl-β-D-thiogalactopyranoside
LPS	Lipopolysaccharides
MLR	Multiple linear regression
MPC	Model predictive control
OD_{600}	Optical density measured at a wavelength of 600 nm
PBS	Phosphate-buffered saline
PI	Prediction interval
POI	Protein of Interest
QbD	Quality by Design
RMSE	Root mean square error
RSM	Response surface model
RT	Room temperature

SCADA	Supervisory control and data acquisition
SD	Standard deviation
SDS-PAGE	Sodium dodecyl sulphate polyacrylamide gel electrophoresis
Tris	Tris(hydroxymethyl)aminomethane
WCB	Working cell bank

7 LIST OF FIGURES

8 LIST OF TABLES AND EQUATIONS

9 REFERENCES

Abu-Absi, N. R., Kenty, B. M., Cuellar, M. E., Borys, M. C., Sakhamuri, S., Strachan, D. J., … Li, Z. J. (2011). Real time monitoring of multiple parameters in mammalian cell culture bioreactors using an in-line Raman spectroscopy probe. *Biotechnology and Bioengineering*. https://doi.org/10.1002/bit.23023

Alzari, P. M., Berglund, H., Berrow, N. S., Blagova, E., Busso, D., Cambillau, C., … Owens, R. J. (2006). Implementation of semi-automated cloning and prokaryotic expression screening: The impact of SPINE. *Acta Crystallographica Section D: Biological Crystallography*, *62*(10), 1103–1113. https://doi.org/10.1107/S0907444906029775

Amelio, A., Van der Bruggen, B., Lopresto, C., Verardi, A., Calabro, V., & Luis, P. (2016). Pervaporation membrane reactors: Biomass conversion into alcohols. In *Membrane Technologies for Biorefining*. https://doi.org/10.1016/B978-0-08-100451-7.00014-1

Bayer, B., Striedner, G., & Duerkop, M. (2020). Hybrid Modeling and Intensified DoE: An Approach to Accelerate Upstream Process Characterization. *Biotechnology Journal*. https://doi.org/https://doi.org/10.1002/biot.202000121

Bayer, B., von Stosch, M., Melcher, M., Duerkop, M., & Striedner, G. (2020). Soft sensor based on 2D-fluorescence and process data enabling real-time estimation of biomass in Escherichia coli cultivations. *Engineering in Life Sciences*, *20*(1–2), 26–35. https://doi.org/10.1002/elsc.201900076

Bayer, B., von Stosch, M., Striedner, G., & Duerkop, M. (2020). Comparison of Modeling Methods for DoE-Based Holistic Upstream Process Characterization. *Biotechnology Journal*, *1900551*. https://doi.org/10.1002/biot.201900551

Brass, J. M., Hoeks, F. W. J. M. M., & Rohner, M. (1997). Application of modelling techniques for the improvement of industrial bioprocesses. *Journal of Biotechnology*, *59*(1–2), 63–72. https://doi.org/10.1016/S0168-1656(97)00165-X

FDA. (2004). Guidance for Industry, PAT-A Framework for Innovative Pharmaceutical Development, Manufacturing and Quality Assurance. *U.S. Department of Health and Human Services, Food and Drug Administration*. Retrieved from http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/ucm070305.pdf

FDA. (2006). Guidance for Industry. Pharmaceutical CGMP Regulations Guidance for Industry. *U.S.*

Department of Health and Human Services, Food and Drug Administration.

FDA. (2018). Current Good Manufacturing Practice (CGMP) Regulations | FDA. *9*.

Ferrer-Miralles, N., Domingo-Espín, J., Corchero, J., Vázquez, E., & Villaverde, A. (2009). Microbial factories for recombinant pharmaceuticals. *Microbial Cell Factories*, *8*, 1–8. https://doi.org/10.1186/1475-2859-8-17

Finkler, C., & Krummen, L. (2016). Introduction to the application of QbD principles for the development of monoclonal antibodies. *Biologicals*. https://doi.org/10.1016/j.biologicals.2016.07.004

Garcia, T., Cook, G., & Nosal, R. (2008). PQLI key topics - Criticality, design space, and control strategy. *Journal of Pharmaceutical Innovation*. https://doi.org/10.1007/s12247-008-9032-4

Gnoth, S., Jenzsch, M., Simutis, R., & Lübbert, A. (2008). Control of cultivation processes for recombinant protein production: A review. *Bioprocess and Biosystems Engineering*. https://doi.org/10.1007/s00449-007-0163-7

Gomes, A. R., & Byregowda. (2016). An Overview of Heterologous Expression Host Systems for the Production of Recombinant Proteins. *Advances in Animal and Veterinary Sciences*. https://doi.org/10.14737/journal.aavs/2016/4.7.346.356

Gräslund, S., Nordlund, P., Weigelt, J., Hallberg, B. M., Bray, J., Gileadi, O., … Gunsalus, K. C. (2008). Protein production and purification. *Nature Methods*, *5*(2), 135–146. https://doi.org/10.1038/nmeth.f.202

Hassan, S., Simaria, A. S., Varadaraju, H., Gupta, S., Warren, K., & Farid, S. S. (2015). Allogeneic cell therapy bioprocess economics and optimization: Downstream processing decisions. *Regenerative Medicine*, *10*(5), 591–609. https://doi.org/10.2217/rme.15.29

Huang, C. J., Lin, H., & Yang, X. (2012). Industrial production of recombinant therapeutics in Escherichia coli and its recent advancements. *Journal of Industrial Microbiology and Biotechnology*. https://doi.org/10.1007/s10295-011-1082-9

Iordache, O. (2013). Design of Experiments and Analysis. *Polytope Projects*, 181–190. https://doi.org/10.1201/b15863-11

Kalil, S. J., Maugeri, F., & Rodrigues, M. I. (2000). Response surface analysis and simulation as a tool for bioprocess design and optimization. *Process Biochemistry*. https://doi.org/10.1016/S0032-9592(99)00101-6

Krippl, M., Dürauer, A., & Duerkop, M. (2020). Hybrid modeling of cross-flow filtration: Predicting the

flux evolution and duration of ultrafiltration processes. *Separation and Purification Technology*. https://doi.org/10.1016/j.seppur.2020.117064

Krishnan, V., Eppinger, S. D., & Whitney, D. E. (1997). A model-based framework to overlap product development activities. *Management Science*, *43*(4), 437–451. https://doi.org/10.1287/mnsc.43.4.437

Kumar, V., Bhalla, A., & Rathore, A. S. (2014). Design of experiments applications in bioprocessing: Concepts and approach. *Biotechnology Progress*, *30*(1), 86–99. https://doi.org/10.1002/btpr.1821

Lee, J., Lee, S. Y., Park, S., & Middelberg, A. P. J. (1999). Control of fed-batch fermentations. *Biotechnology Advances*. https://doi.org/10.1016/S0734-9750(98)00015-9

Luli, G. W., Strohl, W. R., Kumar, V., Bhalla, A., Rathore, A. S., Herwig, C., … Hanus, R. (1999). A new training method for hybrid models of bioprocesses. *Bioprocess Engineering*, *21*(5), 539–550. https://doi.org/10.1007/s004490050697

Lundstedt, T., Seifert, E., Abramo, L., Thelin, B., Nyström, Å., Pettersen, J., & Bergman, R. (1998). Experimental design and optimization. *Chemometrics and Intelligent Laboratory Systems*, *42*(1–2), 3–40. https://doi.org/10.1016/S0169-7439(98)00065-3

Mandenius, C.-F., & Brundin, A. (2008). REVIEW : BIOCATALYSTS AND BIOREACTOR DESIGN, Bioprocess Optimization, Using Design-of-experiments Methodology. *Biotechnol. Progr.*, *24*(6), 1191–1203. https://doi.org/10.1021/bp.67

Marisch, K., Bayer, K., Cserjan-Puschmann, M., Luchner, M., & Striedner, G. (2013). Evaluation of three industrial Escherichia coli strains in fed-batch cultivations during high-level SOD protein production. *Microbial Cell Factories*, *12*(1), 1. https://doi.org/10.1186/1475-2859-12-58

Marques, M. P. C., Cabral, J. M. S., & Fernandes, P. (2010). Bioprocess scale-up: Quest for the parameters to be used as criterion to move from microreactors to lab-scale. *Journal of Chemical Technology and Biotechnology*, *85*(9), 1184–1198. https://doi.org/10.1002/jctb.2387

Mattanovich, D., Branduardi, P., Dato, L., Gasser, B., Sauer, M., & Porro, D. (2012). Recombinant protein production in yeasts. *Methods in Molecular Biology*. https://doi.org/10.1007/978-1-61779-433-9_17

Melcher et al., M. (2015). *Melcher et al. - 2015 - The potential of random forest and neural networks for biomass and recombinant protein modeling in Escherichi.pdf*.

Mercier, S. M. (2013). Multivariate data analysis as a PAT tool for early bioprocess development data. *Journal of Biotechnology*, *167*(3), 262–270. https://doi.org/10.1016/j.jbiotec.2013.07.006

Meyer, C., & Beyeler, W. (1984). Control strategies of continuous bioprocesses based on biological activities. *Biotechnology and Bioengineering*, *26*(8), 916–925. https://doi.org/10.1002/bit.260260816

Möller, J., Kuchemüller, K. B., Steinmetz, T., Koopmann, K. S., & Pörtner, R. (2019). Model-assisted Design of Experiments as a concept for knowledge-based bioprocess development. *Bioprocess and Biosystems Engineering*, *42*(5), 867–882. https://doi.org/10.1007/s00449-019-02089-7

Neubauer, P., Cruz, N., Glauche, F., Junne, S., Knepper, A., & Raven, M. (2013). Consistent development of bioprocesses from microliter cultures to the industrial scale. *Engineering in Life Sciences*. https://doi.org/10.1002/elsc.201200021

Ni, Y., & Chen, R. (2009). Extracellular recombinant protein production from Escherichia coli. *Biotechnology Letters*, *31*(11), 1661–1670. https://doi.org/10.1007/s10529-009-0077-3

Nickel, D. B., Cruz-Bournazou, M. N., Wilms, T., Neubauer, P., & Knepper, A. (2017). Online bioprocess data generation, analysis, and optimization for parallel fed-batch fermentations in milliliter scale. *Engineering in Life Sciences*. https://doi.org/10.1002/elsc.201600035

Omasa, T., Onitsuka, M., & Kim, W.-D. (2010). Cell Engineering and Cultivation of Chinese Hamster Ovary (CHO) Cells. *Current Pharmaceutical Biotechnology*, *11*(3), 233–240. https://doi.org/10.2174/138920110791111960

Opalka, N., Brown, J., Lane, W. J., Twist, K. A. F., Landick, R., Asturias, F. J., & Darst, S. A. (2010). Complete structural model of Escherichia coli RNA polymerase from a Hybrid Approach. *PLoS Biology*. https://doi.org/10.1371/journal.pbio.1000483

Patel, H., Parmar, S., & Patel, B. (2013). A comprehensive review on quality by design (QbD) in pharmaceuticals. *International Journal of Pharmaceutical Sciences Review and Research*, *21*(1), 223–236.

Peuker, T., Riedel, M., Kaiser, C., Ellert, A., Lenz, K., Elshols, O., & Luttmann, R. (2004). At-line determination of glucose, ammonium and acetate in high cell density cultivations of Escherichia coli. *Engineering in Life Sciences*, *4*(2), 138–143. https://doi.org/10.1002/elsc.200401910

Pourmir, A., & Johannes, T. W. (2012). Directed evolution: Selection of the host organism. *Computational and Structural Biotechnology Journal*. https://doi.org/10.5936/csbj.201209012

Rao, G., Moreira, A., & Brorson, K. (2009). Disposable bioprocessing: The future has arrived. *Biotechnology and Bioengineering*. https://doi.org/10.1002/bit.22192

Rathore, A. S., Bhambure, R., & Ghare, V. (2010). Process analytical technology (PAT) for biopharmaceutical products. *Analytical and Bioanalytical Chemistry*, *398*(1), 137–154. https://doi.org/10.1007/s00216-010-3781-x

Read, E. K., Park, J. T., Shah, R. B., Riley, B. S., Brorson, K. A., & Rathore, A. S. (2010). Process analytical technology (PAT) for biopharmaceutical products: Part I. Concepts and applications. *Biotechnology and Bioengineering*, *105*(2), 276–284. https://doi.org/10.1002/bit.22528

Rohe, P., Venkanna, D., Kleine, B., Freudl, R., & Oldiges, M. (2012). An automated workflow for enhancing microbial bioprocess optimization on a novel microbioreactor platform. *Microbial Cell Factories*. https://doi.org/10.1186/1475-2859-11-144

Scheper, T., Hitzmann, B., Stärk, E., Ulber, R., Faurie, R., Sosnitza, P., & Reardon, K. F. (1999). Bioanalytics: Detailed insight into bioprocesses. *Analytica Chimica Acta*. https://doi.org/10.1016/S0003-2670(99)00612-1

Schubert, J., Simutis, R., Dors, M., Havlik, I., & Lübbert, A. (1994). Bioprocess optimization and control: Application of hybrid modelling. *Journal of Biotechnology*, *35*(1), 51–68. https://doi.org/10.1016/0168-1656(94)90189-9

Sommeregger, W., Sissolak, B., Kandra, K., von Stosch, M., Mayer, M., & Striedner, G. (2017). Quality by control: Towards model predictive control of mammalian cell culture bioprocesses. *Biotechnology Journal*. https://doi.org/10.1002/biot.201600546

Teixeira, A., Cunha, A. E., Clemente, J. J., Moreira, J. L., Cruz, H. J., Alves, P. M., … Oliveira, R. (2005). Modelling and optimization of a recombinant BHK-21 cultivation process using hybrid grey-box systems. *Journal of Biotechnology*. https://doi.org/10.1016/j.jbiotec.2005.04.024

Teixeira, A. P., Alves, C., Alves, P. M., Carrondo, M. J. T., & Oliveira, R. (2007). Hybrid elementary flux analysis/nonparametric modeling: Application for bioprocess control. *BMC Bioinformatics*. https://doi.org/10.1186/1471-2105-8-30

Terpe, K. (2006). Overview of bacterial expression systems for heterologous protein production: From molecular and biochemical fundamentals to commercial systems. *Applied Microbiology and Biotechnology*. https://doi.org/10.1007/s00253-006-0465-8

Theisen, M., & Liao, J. C. (2016). Industrial Biotechnology: Escherichia coli as a Host . *Industrial*

Biotechnology, 149–181. https://doi.org/10.1002/9783527807796.ch5

Toms, D., Deardon, R., & Ungrin, M. (2017). Climbing the mountain: Experimental design for the efficient optimization of stem cell bioprocessing. *Journal of Biological Engineering*, *11*(1), 1–10. https://doi.org/10.1186/s13036-017-0078-z

Van Can, H. J. L., Te Braake, H. A. B., Dubbelman, S., Hellinga, C., Luyben, K. C. A. M., & Heijnen, J. J. (1998). Understanding and applying the extrapolation properties of serial gray-box models. *AIChE Journal*, *44*(5), 1071–1089. https://doi.org/10.1002/aic.690440507

Varsakelis, C., Dessoy, S., von Stosch, M., & Pysik, A. (2019). Show me the money! Process modeling in pharma from the investor's point of view. *Processes*, *7*(9). https://doi.org/10.3390/pr7090596

Verbelen, P. J., De Schutter, D. P., Delvaux, F., Verstrepen, K. J., & Delvaux, F. R. (2006). Immobilized yeast cell systems for continuous fermentation applications. *Biotechnology Letters*. https://doi.org/10.1007/s10529-006-9132-5

Verma, R., Boleti, E., & George, A. J. T. (1998). Antibody engineering: Comparison of bacterial, yeast, insect and mammalian expression systems. In *Journal of Immunological Methods* (Vol. 216).

von Stosch, M., Hamelink, J. M., & Oliveira, R. (2016). Hybrid modeling as a QbD/PAT tool in process development: an industrial E. coli case study. *Bioprocess and Biosystems Engineering*, *39*(5), 773–784. https://doi.org/10.1007/s00449-016-1557-1

von Stosch, M., Oliveira, R., Peres, J., & Feyo de Azevedo, S. (2014). Hybrid semi-parametric modeling in process systems engineering: Past, present and future. *Computers & Chemical Engineering*, *60*, 86–101. https://doi.org/10.1016/J.COMPCHEMENG.2013.08.008

von Stosch, M., & Willis, M. J. (2016). Intensified design of experiments for upstream bioreactors. *Engineering in Life Sciences*. https://doi.org/10.1002/elsc.201600037

Waegeman, H., & Soetaert, W. (2011). Increasing recombinant protein production in Escherichia coli through metabolic and genetic engineering. *Journal of Industrial Microbiology and Biotechnology*, *38*(12), 1891–1910. https://doi.org/10.1007/s10295-011-1034-4

Warikoo, V., Godawat, R., Brower, K., Jain, S., Cummings, D., Simons, E., … Konstantinov, K. (2012). Integrated continuous production of recombinant therapeutic proteins. *Biotechnology and Bioengineering*, *109*(12), 3018–3029. https://doi.org/10.1002/bit.24584

Wechselberger, P., Sagmeister, P., & Herwig, C. (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. https://doi.org/10.1007/s00449-012-0848-4

Wurm, D. J., Hausjell, J., Ulonska, S., Herwig, C., & Spadiut, O. (2017). Mechanistic platform knowledge of concomitant sugar uptake in Escherichia coli BL21(DE3) strains. *Scientific Reports*. https://doi.org/10.1038/srep45072

Wurm, F. M. (2004). Production of recombinant protein therapeutics in cultivated mammalian cells. *Nature Biotechnology*, *22*(11), 1393–1398. https://doi.org/10.1038/nbt1026

Yamanè, T., & Shimizu, S. (2005). Fed-batch techniques in microbial processes. In *Bioprocess Parameter Control*. https://doi.org/10.1007/bfb0006382

Yang, Y., & Sha, M. (2020). A beginner's guide to bioprocess modes-batch, fed-batch, and continuous fermentation. *Application Note No. 408*, (408), 1–16.

Yu, L. X. (2008). Pharmaceutical quality by design: Product and process development, understanding, and control. *Pharmaceutical Research*. https://doi.org/10.1007/s11095-007-9511-1

Zorzetto, L. F. M., Filho, R. M., & Wolf-Maciel, M. R. (2000). Processing modelling development through artificial neural networks and hybrid models. *Computers & Chemical Engineering*, *24*(2–7), 1355–1360. https://doi.org/10.1016/S0098-1354(00)00419-1

9.1 Manuals and protocols

IBs: Zellaufschluss und IB lösen: Cserjam-Puschmann, M.; Austrian Centre of Biopharmaceutical Technology; MO-Fermentation internal document; 27.12.2005

NuPAGE Elektrophorese: Cserjan-Puschmann, M.; Austrian Centre of Biopharmaceutical Technology; MO-Fermentation internal document; 07.10.2008

SOD-ELISA: Nemecek, S.; Austrian Centre of Biopharmaceutical Technology; MO Fermentation internal document; 20.12.2005

10 APPENDIX

10.1 Solutions for fed-batch fermentations

10.1.1 Trace element solution

The trace element solution was prepared in 5 N HCl.

Table 10: Composition of Trace element solution

Component	Amount
$FeSO_4*7H_2O$	40 g/L
$MnSO_4*H_2O$	10 g/L
$AlCl_3*6H_2O$	10 g/L
$CoCl_2$	4 g/L
$ZnSO_4*7H_2O$	2 g/L
$Na_2MoO4*2H_2O$	2 g/L
$CuCl_2*2H_2O$	1 g/L
H_3BO_3	0.5 g/L

10.1.2 IPTG solution

For the IPTG solutions, the different amounts of IPTG (20 mg, 50 mg, 90 mg) were dissolved in 300 g double-distilled water (ddH_2O) and sterile filtered afterward.

10.1.3 Sodium chloride solution 0.9%

An amount of 9 g NaCl were dissolved in 1 L ddH_2O. After autoclaving, the solution was stored at RT.

10.1.4 PBS buffer

The following components were solved in ddH_2O and filled up to 1 L.

Table 11: Components for PBS buffer.

Component	Amount
NaCl	8 g

KCl	0.2 g
Na_2HPO_4	1.42 g
KH_2PO_4	0.27 g

10.2 Solutions for SDS Page

10.2.1 Solutions for cell disruption for SDS Page

Table 12: Solutions and their composition for cell disruption

Component	Composition
Tris/HCl 100mM, pH 8.2	1.2114 g/100 mL Tris was solved in ddH_2O and adjusted with 2 M HCl to pH 8.2 and filled up with ddH_2O to 100 mL
Tris/HCl 30 mM, pH 8.2	To 30 mL of 100 mM Tris/HCl 70 mL ddH_2O were added.
EDTA 0.5 M, pH 8.2	7.44 g EDTA were solved in ddH_2O. The pH of 8.2 is adjusted with 5 M KOH and was filled up to 100 mL.
Lysozyme 2 mg/mL	With ddH_2O the stock solution was diluted to a concentration of 2 mg/mL. Aliquots were stored at -20°C.
Benzonase	50 Units/mL solved in Benzonase buffer.
Triton X-100 6%	Triton X-100 stock solution was diluted to a 6% solution with Tris/HCl 20 mM, pH 8.2.
NuPAGE™ Sample Reducing Agent (0.5 M DTT)	Ready to use, stored at 4°C.

10.2.1 Cell disruption buffer

Table 13: Components for cell disruption buffer for a total volume of 3 mL.

Component	Amount

Tris/HCl 30 mM	2.7 mL
EDTA	150 µL
$MgCl_2x6H_2O$	150 µL
Reducing Agent	6 µL

10.2.2 IB dissolving buffer

Table 14: Components for IB dissolving buffer.

Component	Composition
Urea solution 8 M	48.05 g Urea and 63.98 g Tris/HCl 100 mM mixed together and stored in 5 mL aliquots at -20°C.
Reducing Agent	28 µL per mL Urea solution

10.2.3 Components for SDS gel electrophoresis

Table 15: Stock solutions and reagents for SDS gel electrophoresis.

Component	Composition
NuPAGE™ Sample Reducing Agent (10X)	Ready to use, stored at 4°C.
NuPAGE™ LDS Sample Buffer (4X)	Ready to use, stored at 4°C.
Invitrogen™ Mark12™ Unstained Standard	Ready to use, stored at 4°C.
NuPAGE™ 4-12% BisTris Protein Gels	Ready to use, stored at 4°C.
NuPAGE™ MES SDS Running Buffer (20X)	For one chamber 800 mL buffer were used. The running buffer can be reused in the lower chamber. For the upper chamber 10 mL MES (20X) running buffer and 190 mL ddH_2O were combined (dilution in a ration 1:20).
Fixing solution	500 mL ethanol, 100 mL acetic acid were filled up to 1000 mL ddH_2O.

Staining solution	1.16 g Coomassie Brilliant Blue R250, 250 mL ethanol and 80 mL acetic acid were filled up to 1000 mL with ddH_2O.
Destaining solution	250 mL ethanol, 80 mL acetic acid were filled up to 1000 mL ddH_2O.

10.3.1 Protein ladder for SDS Page

To determine the protein weight on a gel, the protein ladder Mark12™ unstained standard was used. 7 µL were loaded onto the gel.

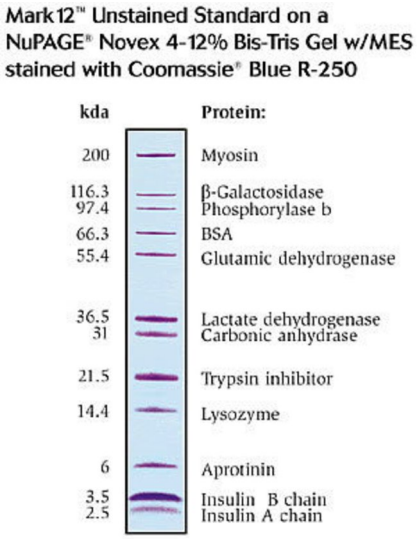

Figure 32: Mark12™ unstained standard to determine the amount of protein weight on SDS Page. (Illustration taken from https://www.thermofisher.com/order/catalog/product/LC5677#/LC5677; 07.11.2019)

10.4 Reagents for ELISA

Table 16: Reagents used for cell disruption for ELISA.

Component	Composition
Tris/HCl 20 mM, pH 8.2	1.2114 g Tris were filled up to 500 mL ddH_2O and a pH of 8.2 is adjusted with 2 M HCl.
EDTA 0.5 M, pH 8.2	7.44 g EDTA were solved in ddH_2O. The pH of 8.2 is adjusted with 5 M KOH and is filled up to 100 mL.
NuPAGE™ Sample Reducing Agent (0.5 M DTT)	Ready to use, stored at 4°C.

Lysozyme 10 mg/mL	With ddH_2O the stock solution was diluted to a concentration of 10 mg/mL. Aliquots are stored at -20°C.
Triton X-100 0.5%	Triton X-100 stock solution was diluted to a 0.5% solution with Tris/HCl 20 mM, pH 8.2.
Cu/Zn solution	0.2 M $CuCl_2$ (1.7048 g) and 0.02 M $ZnSO_4$ (0.2875 g) were filled up with ddH_2O to 50 mL.

10.4.1 Cell disruption buffer for ELISA

Table 17: Amount of components used for ELISA cell disruption buffer. Calculation for 23 samples.

Component	Amount
Tris/HCl 20 mM, pH 8.2	4.5 mL
EDTA	250 µL
Cu/Zn solution	125 µL
Reducing Agent	50 µL

10.4.2 Coating buffer for ELISA

0.1 N $NaHCO_3$ buffer, pH 9.6-9.8. The components were solved in ddH_2O.

Table 18: Amount of components used for 200 mL ELISA coating buffer. Stable for one week stored at 20°C.

Component	Amount
$NaHCO_3$	8.4 g/L → 1.68 g/200 mL
Na_2CO_3	4.2 g/L → 0.84 g/200 mL

10.4.3 Washing buffer for ELISA

PBS buffer, pH 7.2-7.4. The components were solved in ddH_2O.

Table 19: Amount of components used for 5 L ELISA washing buffer. Stable for one week stored at 20°C.

$Na_2HPO_4*2H_2O$	1.15 g/L; 5.75 g/5 L

(or Na_2HPO_4)	(4.6 g/5 L)
KH_2PO_4	0.2 g/L; 1 g/5 L
KCl	0.2 g/L; 1 g/5 L
NaCl	8 g/L; 40 g/5 L
Tween 20	1 mL/L; 5 mL/5 L

10.4.4 Dilution buffer for ELISA

1% of Bovine Serum Albumin was filled up with washing buffer. For 4 ELISA plates, 500 mL of dilution buffer were prepared. Stable for one day at 20°C.

10.4.5 Staining solution for ELISA

P-Nitrophenyl phosphate (PNPP) solution with a concentration of 100 mg/mL were stored in aliquots at -20°C. For one ELISA plate 150 µL were added to 11 mL coating buffer.

10.4.6 Antibodies used for ELISA

Table 20: Antibodies and standard used for ELISA.

Primary caption antibody SOD PA225/C2H2; Lot.: 280105	Dilution 1:250 in coating buffer
Enzyme-labeled antibody SOD A11H4 x AP; Lot.: 180511	Dilution 1:1800 in dilution buffer
IAM SOD standard 1.14 µg/mL	Dilution 1:11.4 in dilution buffer