Combinatorial analysis of host-leader-Fab can greatly improve integrated process design

Mathias Fink[1], Clemens Schimek[2], Monika Cserjan-Puschmann[2], Daniela Reinisch[3], Cécile Brocard[3], Rainer Hahn[4], and Gerald Striedner[2]

[1]Christian Doppler Laboratory for production of next-level biopharmaceuticals in E. coli,University of Natural Resources and Life Sciences
[2]Christian Doppler Laboratory for production of next-level biopharmaceuticals in E. coli,University of Natural Resources and Life Sciences , Department of Biotechnology Vienna, AT
[3]Boehringer Ingelheim RCV GmbH & Co KG, Dr. Boehringer-Gasse 5-11, A-1120 Vienna, Austria
[4]Christian Doppler Laboratory for production of next-level biopharmaceuticals in E. coli, Department of Biotechnology, University of Natural Resources and Life Sciences, A-1190 Vienna, Austria

November 11, 2020

Abstract

Bioprocess development and optimization is a challenging, costly, and time-consuming effort. In this multidisciplinary task, upstream processing (USP) and downstream processing (DSP) are conventionally considered distinct disciplines. This consideration fosters "one-way" optimization without considering interdependencies between unit operations; thus, the full potential of the process chain cannot be achieved. Therefore, it is necessary to fully integrate USP and DSP process development to provide balanced biotechnological production processes. The aim of the present study was to investigate how different host/leader/antigen binding fragment (Fab) combinations in E. coli expression systems influence USP and primary recovery performance and the final product quality. We ran identical fed-batch cultivations with 16 different expression clones to study growth and product formation kinetics, as well as centrifugation efficiency, viscosity, extracellular DNA, and endotoxin content, which are important parameters in DSP. We observed a severe influence on cell growth, product titer, extracellular product, and cell lysis, accompanied by a significant impact on the analyzed parameters of DSP performance. Our results provide the basis for establishing integrated process development considering interdependencies between USP and DSP. These interdependencies need to be understood for rational decision-making and efficient process development.

Research Article

Combinatorial analysis of host-leader-Fab can greatly improve integrated process design

Mathias Fink[1]

Clemens Schimek[1]

Monika Cserjan-Puschmann[1]

Daniela Reinisch[2]

Cécile Brocard[2]

1

Rainer Hahn[1]

Gerald Striedner[1]

[1]Christian Doppler Laboratory for production of next-level biopharmaceuticals in *E. coli* , Department of Biotechnology, University of Natural Resources and Life Sciences, A-1190 Vienna, Austria

[2]Boehringer Ingelheim RCV GmbH & Co KG, Dr. Boehringer-Gasse 5-11, A-1120 Vienna, Austria

Correspondence: Mrs. Monika Cserjan-Puschmann, Department of Biotechnology, University of Natural Resources and Life Sciences, Vienna Muthgasse 18, A-1190 Vienna, Austria.

E-mail : monika.cserjan@boku.ac.at

Keywords: *E. coli* ; Fab; integrated process development; Up- & Downstream processing; Downstream manufacturability

Abbreviations: CDM, cell dry mass; **Fab,** antigen binding fragment; **HCD,** high cell density; **DSP** , downstream processing; **μ-βιορεαςτορ,** micro-bioreactor; **QbD,** quality by design; **scFv,** single-chain fragment variable; **USP,** upstream processing

Abstract

Bioprocess development and optimization is a challenging, costly, and time-consuming effort. In this multidisciplinary task, upstream processing (USP) and downstream processing (DSP) are conventionally considered distinct disciplines. This consideration fosters "one-way" optimization without considering interdependencies between unit operations; thus, the full potential of the process chain cannot be achieved. Therefore, it is necessary to fully integrate USP and DSP process development to provide balanced biotechnological production processes. The aim of the present study was to investigate how different host/leader/antigen binding fragment (Fab) combinations in *E. coli* expression systems influence USP and primary recovery performance and the final product quality. We ran identical fed-batch cultivations with 16 different expression clones to study growth and product formation kinetics, as well as centrifugation efficiency, viscosity, extracellular DNA, and endotoxin content, which are important parameters in DSP. We observed a severe influence on cell growth, product titer, extracellular product, and cell lysis, accompanied by a significant impact on the analyzed parameters of DSP performance. Our results provide the basis for establishing integrated process development considering interdependencies between USP and DSP. These interdependencies need to be understood for rational decision-making and efficient process development.

Introduction

Traditionally, upstream processing (USP) and downstream processing (DSP) studies are performed separately, focusing on their respective functions, though efficient and economical processes comprise both disciplines. Therefore, it is essential to develop processes in a holistic and integrated manner taking both USP and DSP into account [1]. For example, in the course of process development, upstream conditions are often optimized only in view of the final product concentrations, which cannot be equated to high purity and yield in DSP. In addition, during early USP development, low fermentation volumes are used that are not adequate for bench-scale DSP equipment, which makes it necessary to develop ultra-microscale systems [2, 3]. Thus, the significant advances gained in USP coupled with a growing number of therapeutic proteins cause increasing challenges in DSP. For example, high cell densities (HCDs) lead to a bottleneck in primary recovery and purification. Higher protein titers can exceed present DSP capacities and throughput, raising costs [4-6]. Furthermore, HCDs are associated with more fragile cells and exceed periplasmic protein load capacity by translocating recombinant protein during late stage fermentation. The higher number of cells and increased leakage of chromosomal DNA and intracellular contents result in higher viscosities, which can go hand in hand with product leakage [7]. Early stage process development starts after strain engineering and the selection of production strains, which is decisive for the whole bioprocess, as it determines subsequent expression levels and overall process productivity. Select production strains are further characterized and evaluated during process optimization. The feeding strategy is an attractive target for

2

optimization, as growth rates can significantly influence product yields and quality [8]. Higher growth rates can lead to higher total productivity but may also cause higher leakage of periplasmic products. Excessive use of the protein translocation pathway for the transfer of target protein into the periplasm can lead to restricted transportation of important cellular membrane proteins and adversely affect membrane stability [9, 10]. Further fermentation parameters that influence product formation include pH, temperature, carbon and nitrogen sources, and other media components. For example, the use of glucose as carbon source can trigger high acetate formation, which negatively affects growth and product formation. This problem can for example be solved by accurate medium feed strategies or glucose substitution by glycerol. Nevertheless, the carbon source can have a direct impact on the product yield in upstream and downstream unit operations [1, 11, 12]. The physical and biological state of the cells due to cultivation conditions and recombinant protein expression can directly affect DSP performance and product yields. Cells stressed by production exhibit an increased tendency for lysis, which in turn leads to unfavorable particle size distribution, enhanced broth viscosity, poor clearance performance, and higher product losses. One method for avoiding the problems caused by cell lysis is to adjust the time of harvest to the physiological state of the cells. There is also the possibility to positively influence cell fitness and integrity by reducing the culture temperature and weaker induction, which can significantly improve the processability of the cell broth and the efficiency of primary recovery steps [4, 13, 14]. When the optimum time for harvesting has been determined, it is important to ensure rapid processing of the fermentation broth. Interactions between cells and dissolved substances in the broth can negatively influence broth processing, potentially leading to higher resistance in membrane filtration [15]. A different approach aims at periplasmic expression and extracellular secretion to simplify DSP and improve product quality. The latter can be achieved by leaky strains in which cell wall and membrane biosynthesis-associated genes are knocked out and/or by the addition of chemicals promoting the permeability of periplasmic proteins. However, this approach necessitates the optimization of cultivation conditions, as such strains are often affected in growth and robustness [16, 17]. These complexities are even more pronounced when challenging proteins need to be expressed.

Such proteins are mostly produced as low-yield soluble forms or in the form of inclusion bodies (IBs), and they can be toxic to cells. Their DNA sequence can contain rare codons, resulting in low translational efficiency. Furthermore, these proteins can overload host cell folding capacities or be exposed to proteolysis by host proteases [18, 19]. Low expression yields additionally challenge the scale-up and DSP, significantly affecting production costs [20].

Examples of such challenging proteins are antibody fragments, such as antigen binding fragments (Fabs) or single-chain variable fragments (scFvs), which can be produced in *E. coli* [21, 22]. We used Fabs as model proteins in this study. Fabs are small immunoglobulin G (IgG)-derived molecules that constitute monovalent binding arms composed of a heavy and a light chain connected by disulfide bonds [23]. Despite intensive research and development efforts, the production of Fabs in *E. coli* still poses a great challenge. Fab-specific properties strongly interfere with cell vitality, even at low expression levels, preventing efficient production. The formation of disulfide bonds, which are required for correct folding, is another issue. On the one hand, thioredoxin mutant strains were established that change the environment in the cytoplasm from reducing to oxidizing. On the other hand, secretory signals are used to secrete Fabs into the extracellular space or translocate it to the oxidizing periplasm [14, 24, 25]. An additional parameter that impacts the expression of Fabs in the periplasm is inefficient translocation and incomplete processing of signal peptides, again causing low yields, challenging USP and DSP [26]. Another possibility is to express Fabs as insoluble IBs in the cytosol. However, this latter approach has some drawbacks. After separation of IBs from the disrupted cells, they need to be solubilized and the Fab must be refolded. A number of chemicals and unit operations are required, further challenging DSP activities [27, 28]. Therefore, it is obvious that changes in USP directly affect the operability and performance of DSP. Thus, to develop productive and economic processes, USP and DSP must be investigated simultaneously [29]. This is in line with a quality-by-design (QbD) approach in bioprocessing, which demands a thorough understanding of the whole process [1, 30].

The aim of the present study was to investigate how different Fabs and leader combinations integrated into two expression hosts influence both the expression and downstream operability. The influence of these

3

variations was investigated at the level of cell physiology, product yield, and parameters of early stage DSP. The results show that the specific Fab and host strain had the greatest influence when using a predetermined fermentation strategy. Our results will serve as a basis to further evaluate the direct impact of USP variations on DSP.

Materials and methods

Production clones

All of the production clones used during this study and their generation were described in detail previously [31]. We used genome-integrated *E. coli* BL21(DE3) and *E. coli* HMS174(DE3), expressing the four different Fabs BIBH1, BIWA4, FTN2, and Fabx. These Fabs were translocated to the periplasm using the ompA and dsbA signal sequences (SSs), resulting in a total of 16 expression clones.

Fermentation

All clones were cultivated as biological triplicates in a fed-batch process using a DASGIP® parallel bioreactor system (Eppendorf AG) equipped with four parallel operable vessels holding a maximum working volume of 1.8 L. For online monitoring, each system was equipped with a pH and optical DO probe (Hamilton Bonaduz AG, Switzerland), as well as a DASGIP® GA4X-module (Eppendorf AG) for off gas monitoring. The bioreactors were inoculated with a pre-culture obtained from 500 mL baffled shaker flasks filled with 50 mL semi-synthetic medium at 180 rpm at 37°C in an orbital shaker. For inoculation of the pre-culture, a deep-frozen working cell bank was used after thawing. When the pre-culture reached an OD_{600} of 2.8-3.8, an appropriate amount was transferred to the bioreactor to reach an initial OD_{600} of 0.04. The batch volume was 600 mL, and the medium was prepared to achieve a final cell dry mass (CDM) of 6 g. After the batch phase, an exponential feed maintaining a constant growth rate of 0.1 h^{-1} was started and maintained for 19 h to gain a calculated final biomass of 40 g CDM. The batch phase was performed at 37°C, whereas the feed phase was performed at 30°C. During the whole process, the pH was maintained at 7.0 ± 0.2 via addition of 12.5% ammonia solution and the dissolved oxygen stabilized at a setpoint of 30% by cascade control of the stirrer speed, aeration rate, and gassing composition. For the induction of protein expression, isopropyl-β-D-thiogalactopyranoside (IPTG; Carl Roth) was added after approximately 0.4 doublings after the feed start. Induction was performed to yield a constant concentration of 2 μmol IPTG/g theoretical CDM.

All details regarding the media composition and fermentation process are shown in Fink et.al. [32].

Offline analysis

Upstream sampling and analysis

From the time of induction onwards, samples were drawn every second hour. For measurement of the OD_{600}, an Ultrospec 500 pro Spectrophotometer (Amersham Biosciences, UK) was used. To reach the linear range of the measurement, the samples were diluted with phosphate-buffered saline (PBS). The CDM was determined using dried and pre-weighed reaction tubes by centrifuging 1 mL of sample, washing the pellet, and drying it to a constant weight at 105°C. For product analysis, a sample volume equivalent to a calculated mass of 1 mg CDM was centrifuged. The pellets and supernatants were stored separately at -20°C. In addition, to determining the expression of the intracellular soluble Fab, the extracellular Fab in the supernatant was determined using a specific sandwich enzyme-linked immunosorbent assay (ELISA). Sample preparation and analysis methods were described in detail by Fink et al. [31].

DSP parameters

Endpoint samples were drawn to determine DSP parameters. Cell harvest and the effectiveness of centrifugation, viscosity of the fermentation broth and supernatant, and the extracellular DNA and endotoxin contents in the fermentation supernatant were specified as steps of early stage DSP. Quality control aspects were evaluated by mass spectrometry (MS) analysis of the full-size Fab.

4

Cell harvest

Cell harvest experiments were performed using a Heraeus Multifuge X3FR (Thermo Fisher Scientific, USA) equipped with a TX-750 swinging bucket rotor (Thermo Fisher Scientific). We transferred 50 mL cell broth to 50 mL conical centrifuge tubes (Greiner Bio-One, Austria) and centrifuged them at 4000 g and 4°C for 30 min. After centrifugation, the pellet height was measured to evaluate the solid fraction and the supernatant used for further analysis of turbidity and viscosity.

Turbidity

Turbidity was measured using a 2100Q nephelometer (HACH Company, USA). For the measurement of high turbidity, the suspensions were diluted with RO-H_2O to reach the measurement range of the nephelometer (i.e., <1000 NTU). The turbidity of the fermentation broth and supernatant was determined to evaluate the efficiency of centrifugation. Approximately 15 mL of homogeneous sample was transferred to the glass cuvette and measured by the device.

Viscosity

A Viskometer DV-II+Pro equipped with a CPA41Z spindle (Brookfield Viscometers Ltd., UK) was used to determine viscosity. According to spindle specifications, 2 mL of sample were transferred to the sample cup. For each fermentation, the cell broth and supernatant were measured. The viscosity was determined at a shear rate of 70 s^{-1}.

Extracellular DNA

The extracellular DNA content was determined at the end of fermentation and over a time course during fermentation to monitor the release of DNA as evidence of cell lysis. Supernatant samples were drawn and analyzed using a Qubit 4 fluorometer (Invitrogen, USA) with a dsDNA BR assay kit (Invitrogen) according to the manufacturer's protocol.

Extracellular endotoxin

The supernatants of the endpoint samples of fermentation were analyzed with regards to the endotoxin content. The samples were processed using the EndoZyme® II kit (bioMérieux SA, France) according to the manufacturer's protocol. Measurements were performed using an Infinite® M200 Pro plate reader (Tecan Trading AG, Switzerland).

MS analysis

For MS analysis, samples were directly injected into a liquid chromatography electrospray ionization MS (LC-ESI-MS) system (Dionex Ultimate 3000 LC, Thermo Fisher Scientific Inc., USA). We injected 5 µL of Source 15S (GE Healthcare Life Sciences, USA) cation exchange resin eluate via µL-pick up using 0.1% formic acid (Acros Organics, Belgium) as the transport liquid to fill up the 20 µL sample loop. In the case of the KanCap G (Kaneka Corp., Japan) fraction, a full loop mode injection was performed. For sample preparation, a 3D chromatography approach was used, as described in detail in [33]. Samples were run on a ProSwift RP-4H column (0.2 x 250 mm; Thermo Fisher Scientific Inc.) operated with a 20-80% acetonitrile (VWR®, USA) gradient in 0.05% trifluoroacetic acid (Thermo Fisher Scientific) at a flow rate of 8 µL/min and a gradient time of 30 min. Detection was performed via the maXis 4G quadrupole time-of-flight (Q-TOF) instrument (Bruker Corp., USA) equipped with the ESI source in positive ion MS mode with a range of 400-3000 Da. For calibration, an ESI calibration mixture (Agilent Technologies, USA) was used. Data processing was performed using Data Analysis 4.0 (Bruker). MaxEnt was used for spectrum deconvolution in the mass range of 10,000-100,000 Da.

Results

The different combinations of host, Fab, and translocation pathway have led to a wide range of impacts on cell growth, Fab expression level and localization, extracellular DNA and endotoxin content, viscosity, and centrifugation efficiency. These parameters are summarized in Table 1. The values were standardized using a

5

z-score calculation to illustrate the interdependencies of the parameters caused by the different combinations and plotted in radar charts (Fig. 2 and 3).

Cultivation process and product formation

Product formation and its leakage into the supernatant were monitored offline from the time of induction onwards to evaluate biomass accumulation during cultivation. The CDM was measured and Fab concentrations quantified for the intracellular and extracellular fractions.

Cell growth

Severe negative effects on cell growth were observed as a function of the expressed Fab leader combination in both investigated hosts (Fig. 1A-D). The strongest effect occurred with the expression of Fabx. The Fabx-expressing clones always achieved the lowest CDM within the host/Fab combinations. In general, HMS174(DE3) was more affected than BL21(DE3) clones, and the Fab-specific dependency was more pronounced. Overall, the lowest CDM was reached with H<oFabx> (12.45 g), whereas the highest was reached with B<dFabx> (50.06 g). On average, all BL21(DE3) clones had 30% higher final CDM than the corresponding HMS174(DE3) clones. The final CDM values are given in Table 1. With regard to growth over time (Fig. 1A-D), we observed that, in almost all clones, the negative impact on CDM formation started between 6 and 9 hours after induction, which is close to one generation in the induced state. For H<oFabx>, the combination with the strongest response, it started 3 hours after induction. Overall, the influence was more pronounced and appeared slightly earlier for HMS174(DE3).

Fab expression

To monitor product formation, samples were taken every 2 hours from the time of induction onwards. Due to the expected product release into the supernatant, soluble intracellular and extracellular Fab concentrations were determined per gram of biomass (Fig. 1E-H and I-L). The results showed a wide range of final yields with different distributions, with the antigen-binding region as the main factor influencing these differences, followed by the expression strain (Table 1). The highest yields were achieved with H<oFTN2>, which had intracellular and extracellular yields of 1.9 and 3.3 mg/g CDM, respectively. This corresponds to a total yield of 5.2 mg/g CDM and a product titer of approximately 156.4 mg/L. However, we did not observe any expression in B<dFabx>. All combinations showed the same ranking in terms of yield, with FTN2 at the top, followed by BIWA4, BIBH1, and Fabx, except for the B<dFab> clones, in which BIBH1 and BIWA4 switched positions. When comparing the two host strains, we observed 69% higher average total specific yields with HMS174(DE3) for all Fab/leader combinations. A significant increase in product release into the supernatant was observed approximately 8 hours after induction (Fig. 1I-L). This temporal correlation is characteristic and can be associated with the negative effects on growth and cell lysis. The final extracellular fractions ranged from ~38% (B<oBIBH1>) to 71% (H<oFabx>). Higher extracellular fractions were observed with DsbA^{ss} in the case of BL21(DE3), whereas this was observed with OmpA^{SS} for HMS174(DE3).

Early stage DSP characteristics

Extracellular DNA and endotoxin contents in the fermentation broth, viscosity, and centrifugation efficiency are relevant parameters for early stage DSP operability. The results are summarized in Table 1. We used 3D chromatography followed by MS analysis to evaluate the expressed Fabs in terms of mass, possible alterations in the supernatant, and free LC content [33].

Extracellular DNA

The DNA content in the fermentation broth was determined in endpoint samples and defined as an indicator of cell lysis associated with a reduced final CDM and product leakage. With regard to DSP operability, DNA in the supernatant leads to increased viscosity and can negatively impact cell harvesting and subsequent steps in the process. DNA analysis showed that the different host/leader/Fab combinations resulted in a wide range of extracellular DNA concentrations, from 175 µg/mL for B<oFTN2> up to 1016 µg/mL

6

for H<dBIBH1>. The HMS174(DE3) clones had 2.5-times the content of their BL21(DE3) counterparts on average (Fig. 2A). Higher DNA concentrations may be associated with a stronger negative impact on growth and higher extracellular product fraction (Fig. 3A and B). In addition, the translocation pathway influenced BL21(DE3), as DsbA^{SS} led to higher DNA concentrations in the supernatant.

Extracellular endotoxin

In addition to DNA, extracellular endotoxins play a major role in the processing of fermentation broths. Comparing the two host strains illustrated in Figure 2B, HMS174(DE3) had significantly higher endotoxin contents for almost all combinations tested, approximately 28% higher on average. Furthermore, the influence of the signal sequence on endotoxin levels varied between the two hosts. In BL21(DE3), OmpA^{SS} led to higher levels, and this was the case for DsbA^{SS} in HMS174(DE3). Endotoxin contents varied between 3.3×10^6 EU/mL for B<dFabx> and 8.1×10^6 EU/mL for H<dBIBH1>.

Viscosity

The viscosity of the fermentation broth directly influences early stage DSP during the cell harvest unit operation. In this study, the viscosities of the cell broth and supernatant were determined at a shear rate of 70 s^{-1}. Comparing the two host strains, the average reduction in viscosity from broth to supernatant was approximately 30% for BL21(DE3) and 23% for HMS174(DE3) (Fig. 2C). For almost all combinations, broths and supernatants originating from HMS174(DE3) had significantly higher viscosities. The viscosity of cell broths varied from 1.61 mPas for B<dFabx> up to 13.81 mPas for H<oBIBH1>. For the supernatants, the minimum and maximum were 1.05 and 11.64 mPas, respectively.

Centrifugation efficiency

Centrifugation efficiency was described as a reduction in turbidity [%] by centrifugation under well-defined conditions for all samples (Fig. 2D). Thus, turbidity was measured before and after centrifugation. Centrifugation efficiency was calculated according to Equation 1, where η_c is the centrifugation efficiency, NTU_f is the turbidity of the feed broth, NTU_s is the turbidity of the supernatant, and NTU_0 is the minimal achieved turbidity (B<dFabx>).

$$\eta_c = \frac{NTU_f - NTU_s}{NTU_f - NTU_0} \bullet 100\% \quad (1)$$

Based on the higher cell densities, BL21(DE3) fermentation broths had greater turbidity. All BL21(DE3) clones had high reduction of turbidity, at least 94%, except B<oFabx>, which was also the one that had the highest impact on growth. Notably, the highest reduction was achieved for B<dFabx>, which did not express any Fab and grew unaffected. In general, the HMS174(DE3) clones had lower centrifugation efficiency, reflecting the higher impact on cell growth.

MS analysis

LC-ESI-MS was used to evaluate the Fab products. We determined the molecular mass of both intracellular and extracellular Fab. All of FTN2 (100%) was the correctly folded full Fab both intracellularly and extracellularly, except for B<dFTN2>, in which a minimal amount (~1%) presented a mass shift of +148 Da.

In the case of BIWA4, all clones produced only the correct Fab, except for H<dBIWA4>, which had 70% with +2 Da in the extracellular fraction, and B<oBIWA4>, which had a mass shift of -178 Da for approximately 5% of the intracellularly retained Fabs. BIBH1 had a more pronounced mass shift, as the ratio of correctly processed Fab varied from 40% to 75% for both fractions. The observed mass shift was -17 Da for all clones. For B<dBIWA4>, the intracellular fraction had a mixture of -16 and +100 Da. Due to the low amounts that were gained for Fabx, there was only one MS result, which indicate a ratio of 30% -17 Da in the extracellular fraction of B<oFabx>. Wide variations were observed in terms of the ratio of free LC and LC fragments of total Fab. The ratio ranged from 0% for B<oBIWA4> up to 74% for H<oFTN2>;

7

nevertheless, all combinations showed similar ratios in intracellular and extracellular fractions (Table 2). In general, the ranking for these impurities was FTN2>BIBH1>BIWA4, independent of the investigated host and leader combination. Due to the very low expression levels of Fabx, this point was only investigated for one exemplary combination.

Discussion

In this study, we investigated how the combination of host strain, expressed Fab, and translocation pathway interact with and influence process efficiency. One of the main results was that the selected variations and their combined application led to scattered product yields and cell growth, as well as significant differences in DSP operability. The greatest influence was from the Fab molecule that was expressed, though all four candidates had the same backbone and only the CDA region varied. The highest product concentrations for all combinations were measured for FTN2, followed by BIWA4, BIBH1, and Fabx, with few exceptions. For all Fab-leader combinations, HMS174(DE3) was superior to its BL21(DE3) counterparts in terms of specific Fab titers. However, the BL21(DE3) host responded less strongly to the expression of the Fabs, and cells were generally more vital with better growth behavior. This was also confirmed by the DSP investigations, in which considerably less DNA was quantified in the supernatant, resulting in lower viscosity and better clearance (Fig. 3A and B). Considering the values from Table 1, linearity between clearance efficiency, viscosity, and DNA content is clearly apparent if the values for B<oFabx> are left aside. The least influence on process and yield was observed for the leader sequence. No general overriding behavior could be identified, only that OmpA^{SS} led to slightly higher expression levels in both strains in most cases. Taken together with DSP operability, the chosen procedure in the USP, with early induction and a long production phase independent of strain and Fab, resulted in very unfavorable distribution ratios of intracellular and extracellular Fabs, in the range of 60:40 to 40:60. One of the fractions would have to be omitted in DSP, which is not acceptable in view of the generally low product titers. Although high extracellular fractions were observed in both strains, there was a significant difference in extracellular DNA content, which was much lower in BL21(DE3), indicating increased release of Fab into the supernatant independent of cell lysis (Fig. 2A and Fig. 3C and D). This finding is clearly a positive characteristic attributable to the strain with high impact on process design and efficiency.

The traditional approach for process development focused on USP yield would favor H<oFab> clones as expression hosts due to the higher specific total Fab titers. However, the high cell lysis rates, high extracellular DNA concentrations, and high viscosities would negatively affect the DSP performance and pose great challenges for process development. DSP would be faced with cell broths containing high levels of extracellular product fraction and increased DNA and endotoxin levels, in combination with low centrifugation efficiencies due to the elevated viscosity. Therefore, a classical purification process for intracellular products would not be economic and would lead to high product loss. This could be circumvented by homogenizing the cell broth before harvesting. However, a drawback would be that DSP would be confronted with a broth containing media residue and extracellular metabolites that are normally removed by harvesting the cells. These impurities would need to be removed during subsequent unit operations.

These results highlight the importance of a holistic approach for process optimization, including integrating the upstream and downstream aspects. The parameters selected for the evaluation of DSP performance were found to be a good choice. These parameters allow for a comprehensive process evaluation and provide relevant information for integrated process optimization. In our specific case, due to the observed growth and product leakage kinetics, shifting the induction and harvest time points could be a starting point for optimization. Later induction would shorten the expression period and, therefore, could reduce cell lysis and product leakage.

The MS analysis showed that, in almost all cases, the correct regular Fab was expressed and the extracellular fraction was not subjected to changes. Only BIBH1 presented a significant mass shift caused by the cyclization of N-terminal glutamine to pyroglutamate, which is a well-described phenomenon in the literature for the production of recombinant antibodies [38, 39]. This represents another issue affecting DSP operability, as separation of product variants is challenging due to their small differences in charge. Furthermore, in

8

all combinations except B<oBIWA4>, remarkable amounts of free LC and LC-fragments were found, which can impede DSP by limiting Protein L column capacities, as these impurities will compete with the product for binding [34].

Thus, we showed that strong impairment of growth by Fab expression leads to a high extracellular product fraction and contamination, resulting in high viscosity and low centrifugation efficiency. The chosen host and expressed Fab product exerted the main influences on growth, titer, and DSP operability. In particular, HMS174(DE3) clones had higher extracellular product fractions, DNA and endotoxin contents, and viscosities, and lower final CDM.

Our results indicate significant interdependencies between the two disciplines, and that decisions made at the USP level are especially important for the performance of the first steps of DSP. Therefore, the classical yield-oriented USP development approach would create major challenges for DSP development, as higher extracellular impurities, product localization, and poor manufacturability would lead to low overall product yields and high manufacturing costs. This underlines the need for a holistic and integrated process development approach to establish reliable and economical production processes, and to shorten the overall time to market. Our results provide the basis for further research into integrated and holistic process development. By monitoring the impact on growth and product localization caused by changes in the USP strategy, we are now able to better predict fluctuations in DSP.

Acknowledgements

Conflict of interest

The authors declare no financial or commercial conflict of interest.

5 References

[1] R. J. Carvalho, J. Cabrera-Crespo, M. M. Tanizaki, V. M. Gonçalves, Development of production and purification processes of recombinant fragment of pneumococcal surface protein A in Escherichia coli using different carbon sources and chromatography sequences, *Applied Microbiology and Biotechnology* 2012, 94, 683.

[2] C. Walther, D. Boras, L. Demmel, M. Berkemeyer, et al., Integrated process development - quality by design compliant evaluation of upstream variations at the microscale level, *Journal of Chemical Technology and Biotechnology* 2018, 93, 2021.

[3] S. Ali, M. A. Perez-Pardo, J. P. Aucamp, A. Craig, et al., Characterization and feasibility of a miniaturized stirred tank bioreactor to perform E. coli high cell density fed-batch fermentations, *Biotechnology Progress* 2012, 28, 66.

[4] Q. Li, G. J. Mannall, S. Ali, M. Hoare, An ultra scale-down approach to study the interaction of fermentation, homogenization, and centrifugation for antibody fragment recovery from rec E. coli, *Biotechnology and Bioengineering* 2013, 110, 2150.

[5] D. J. Roush, Y. Lu, Advances in primary recovery: Centrifugation and membrane technology, *Biotechnology Progress* 2008, 24, 488.

[6] A. Meitz, P. Sagmeister, T. Langemann, C. Herwig, An integrated downstream process development strategy along QbD principles, *Bioengineering* 2014, 1, 213.

[7] J. M. Newton, D. Schofield, J. Vlahopoulou, Y. Zhou, Detecting cell lysis using viscosity monitoring in E. coli fermentation to prevent product loss, *Biotechnology progress* 2016, 32, 1069.

[8] J. P. Aucamp, R. Davies, D. Hallet, A. Weiss, et al., Integration of host strain bioengineering and bioprocess development using ultra-scale down studies to select the optimum combination: An antibody fragment primary recovery case study, *Biotechnology and Bioengineering* 2014, 111, 1971.

[9] A. Shokri, A. M. Sandén, G. Larsson, Growth rate-dependent changes in Escherichia coli membrane structure and protein leakage, *Applied Microbiology and Biotechnology* 2002, 58, 386.

[10] E. Bäcklund, D. Reeks, K. Markland, N. Weir, et al., Fedbatch design for periplasmic product retention in Escherichia coli, *Journal of Biotechnology* 2008, 135, 358.

[11] D. N. Adhyaru, N. S. Bhatt, H. A. Modi, Optimization of upstream and downstream process parameters for cellulase-poor-thermo-solvent-stable xylanase production and extraction by Aspergillus tubingensis FDHN1, *Bioresources and Bioprocessing* 2015, 2.

[12] E. Faulkner, M. Barrett, S. Okor, P. Kieran, et al., Use of fed-batch cultivation for achieving high cell densities for the pilot-scale production of a recombinant protein (phenylalanine dehydrogenase) in Escherichia coli, *Biotechnology Progress* 2006, 22, 889.

[13] M. A. Perez-Pardo, S. Ali, B. Balasundaram, G. J. Mannall, et al., Assessment of the manufacturability of Escherichia coli high cell density fermentations, *Biotechnology Progress* 2011, 27, 1488.

[14] C.-C. Hsu, O. R. T. Thomas, T. W. Overton, Periplasmic expression in and release of Fab fragments from Escherichia coli using stress minimization, *Journal of Chemical Technology & Biotechnology* 2015, n/a.

[15] Y. Okamoto, K. Ohmori, C. E. Glatz, Harvest time effects on membrane cake resistance of Escherichia coli broth, *Journal of Membrane Science* 2001, 190, 93.

[16] Y. Zhou, Z. Lu, X. Wang, J. N. Selvaraj, et al., Genetic engineering modification and fermentation optimization for extracellular production of recombinant proteins using Escherichia coli, *Applied Microbiology and Biotechnology* 2018, 102, 1545.

[17] M. Luo, M. Zhao, C. Cagliero, H. Jiang, et al., A general platform for efficient extracellular expression and purification of Fab from Escherichia coli, *Applied Microbiology and Biotechnology* 2019, 103, 3341.

[18] M. Barazesh, Z. Mostafavipour, S. Kavousipour, S. Mohammadi, et al., Two simple methods for optimizing the production of "difficult-to- Express" GnRH-DFF40 chimeric protein, *Advanced Pharmaceutical Bulletin* 2019, 9, 423.

[19] J. Eichmann, M. Oberpaul, T. Weidner, D. Gerlach, et al., Selection of High Producers From Combinatorial Libraries for the Production of Recombinant Proteins in Escherichia coli and Vibrio natriegens, *Frontiers in Bioengineering and Biotechnology* 2019, 7.

[20] A. Kulmala, T. Huovinen, U. Lamminmäki, Effect of DNA sequence of Fab fragment on yield characteristics and cell growth of E. coli, *Scientific Reports* 2017, 7.

[21] G. Walsh, Biopharmaceutical benchmarks 2014, *Nat Biotech* 2014, 32, 992.

[22] Z. A. Ahmad, S. K. Yeap, A. M. Ali, W. Y. Ho, et al., scFv antibody: principles and clinical application, *Clin Dev Immunol* 2012, 2012, 980250.

[23] M. Ellis, P. Patel, M. Edon, W. Ramage, et al., Development of a high yielding E. coli periplasmic expression system for the production of humanized Fab' fragments, *Biotechnology Progress* 2017, 33, 212.

[24] S. Kumar, J. Kalsi, D. S. Latchman, L. H. Pearl, et al., Expression of the Fabs of human auto-antibodies in Escherichia coli: Optimization and determination of their fine binding characteristics and cross-reactivity, *Journal of molecular biology* 2001, 308, 527.

[25] T. Gundinger, A. Pansy, O. Spadiut, A sensitive and robust HPLC method to quantify recombinant antibody fragments in E. coli crude cell lysate, *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences* 2018, 1083, 242.

[26] Y. Zhou, P. Liu, Y. Gan, W. Sandoval, et al., Enhancing full-length antibody production by signal peptide engineering,*Microbial Cell Factories* 2016, 15.

[27] K. F. J. Metzger, W. Padutsch, A. Pekarsky, J. Kopp, et al., IGF1 inclusion bodies: A QbD based process approach for efficient USP as well as early DSP unit operations, *Journal of Biotechnology* 2020, 312, 23.

[28] H. J. Lee, C. M. Lee, K. Kim, J. M. Yoo, et al., Purification of antibody fragments for the reduction of charge variants using cation exchange chromatography, *Journal of Chromatography B: Analytical Technologies in the Biomedical and Life Sciences* 2018, 1080, 20.

[29] A. L. Cataldo, D. Burgstaller, G. Hribar, A. Jungbauer, et al., Economics and ecology: Modelling of continuous primary recovery and capture scenarios for recombinant antibody production, *Journal of Biotechnology* 2020, 308, 87.

[30] M. Melcher, T. Scharl, M. Luchner, G. Striedner, et al., Boosted structured additive regression for Escherichia coli fed-batch fermentation modeling, *Biotechnology and Bioengineering* 2017, 114, 321.

[31] M. Fink, S. Vazulka, E. Egger, J. Jarmer, et al., Microbioreactor Cultivations of Fab-Producing Escherichia coli Reveal Genome-Integrated Systems as Suitable for Prospective Studies on Direct Fab Expression Effects, *Biotechnology Journal* 2019, 14.

[32] M. Fink, M. Cserjan-Puschmann, D. Reinisch, G. Striedner, High-throughput microbioreactor cultivations greatly contribute to early-stage fermentation process development Under Review.

[33] C. Schimek, M. Kubek, D. Scheich, M. Fink, et al., Three-dimensional chromatography for purification and characterization of antibody fragments and related impurities from *Escherichia coli* crude extracts, UNDER REVIEW.

[34] G. Rodrigo, M. Gruvegard, J. Alstine, Antibody Fragments and Their Purification by Protein L Affinity Chromatography,*Antibodies* 2015, 4, 259.

Table 1. Overview of relevant upstream and downstream parameters.

Hosted file

`image1.emf` available at https://authorea.com/users/374714/articles/492163-combinatorial-analysis-of-host-leader-fab-can-greatly-improve-integrated-process-design

Table 2. Fraction (%) of free LC/LC-fragments in proportion to full Fab.

Hosted file

`image2.emf` available at https://authorea.com/users/374714/articles/492163-combinatorial-analysis-of-host-leader-fab-can-greatly-improve-integrated-process-design
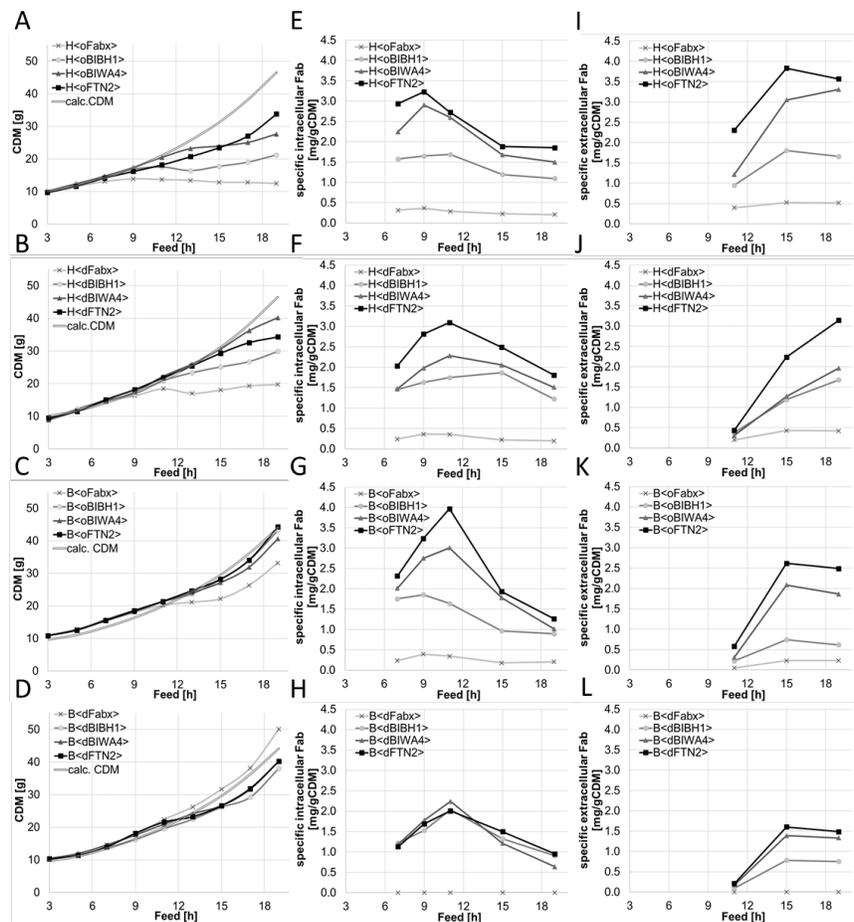
Figures

Figure 1. Fed-batch cultivation of *E. coli* HMS174(DE3) and *E. coli* BL21(DE3) with different Fab/leader combinations. (**A-D**) Kinetics of biomass formation with calculated CDM indicated by double lines. (**E-H**) Specific intracellular and (**I-L**) extracellular product over time. All graphs start from induction (3 hours of feed), and rows from top to bottom show data for H<oFab>, H<dFab>, B<oFab>, and B<dFab>. Data are shown as the mean of three biological triplicates. Due to clarity, no error bars are shown.
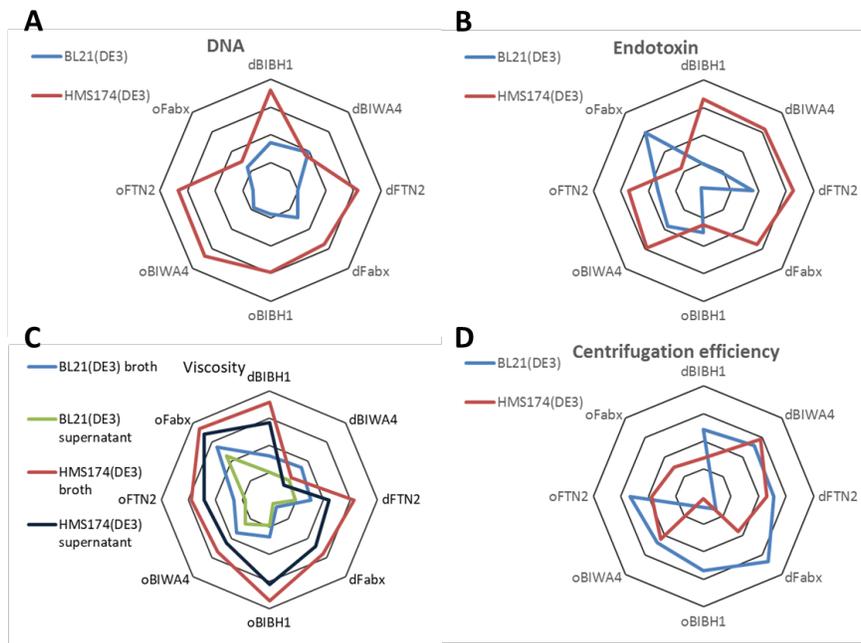
Figure 2. Radar charts for Z-score-normalized DNA (**A**), endotoxin (**B**), viscosity (**C**), and centrifugation efficiency (**D**) comparing BL21(DE3) and HMS174(DE3). n = 3 for DNA, CDM, and extracellular Fab; n = 2 for endotoxin, centrifugation efficiency, and viscosity.
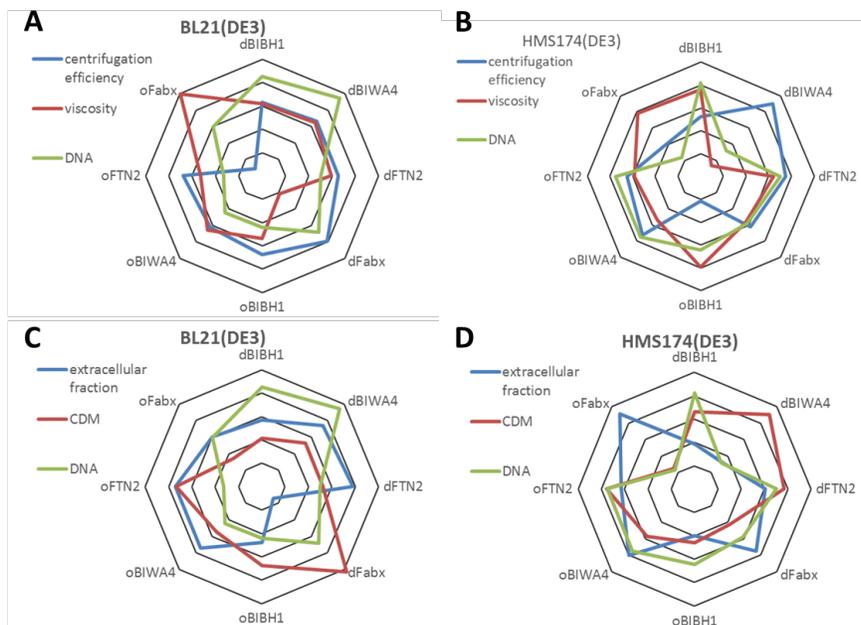
Figure 3. Radar charts of Z-score-normalized DSP parameters for the visualization of observed interdependencies in BL21(DE3) and HMS174(DE3). DNA content, viscosity, centrifugation efficiency (**A** , **B**), extracellular Fab fraction, DNA, CDM (**C** , **D**). n = 3 for DNA, CDM, and extracellular Fab; n = 2 for centrifugation efficiency and viscosity.