processes

MDPI

Article

A Chemometric Tool to Monitor and Predict Cell Viability in Filamentous Fungi Bioprocesses Using UV Chromatogram Fingerprints

Philipp Doppler [1,†], Lukas Veiter [1,2,†], Oliver Spadiut [1], Christoph Herwig [1,2,3] and Vignesh Rajamanickam [1,*]

1 Envirionmental and Bioscience Engineering, Research Area Biochemical Engineering, Technische Universität Wien, Institute of Chemical, Gumpendorfer Strasse 1a, 1060 Vienna, Austria; philipp.doppler@tuwien.ac.at (P.D.); lukas.veiter@tuwien.ac.at (L.V.); oliver.spadiut@tuwien.ac.at (O.S.); christoph.herwig@tuwien.ac.at (C.H.)

2 Competence Center CHASE GmbH, Altenbergerstraße 69, 4040 Linz, Austria

3 Christian Doppler Laboratory for Mechanistic and Physiological Methods for Improved Bioprocesses, TU Wien, Gumpendorfer Straße 1a, 1060 Vienna, Austria

* Correspondence: vignesh.rajamanickam@tuwien.ac.at; Tel.: +43-1-58801-166-496

† Philipp Doppler & Lukas Veiter contributed equally to this work.

Received: 27 March 2020; Accepted: 10 April 2020; Published: 14 April 2020

Abstract: Monitoring process variables in bioprocesses with complex expression systems, such as filamentous fungi, requires a vast number of offline methods or sophisticated inline sensors. In this respect, cell viability is a crucial process variable determining the overall process performance. Thus, fast and precise tools for identification of key process deviations or transitions are needed. However, such reliable monitoring tools are still scarce to date or require sophisticated equipment. In this study, we used the commonly available size exclusion chromatography (SEC) HPLC technique to capture impurity release information in *Penicillium chrysogenum* bioprocesses. We exploited the impurity release information contained in UV chromatograms as fingerprints for development of principal component analysis (PCA) models to descriptively analyze the process trends. Prediction models using well established approaches, such as partial least squares (PLS), orthogonal PLS (OPLS) and principal component regression (PCR), were made to predict the viability with model accuracies of 90% or higher. Furthermore, we demonstrated the platform applicability of our method by monitoring viability in a *Trichoderma reesei* process for cellulase production. We are convinced that this method will not only facilitate monitoring viability of complex bioprocesses but could also be used for enhanced process control with hybrid models in the future.

Keywords: cell viability; prediction; chromatogram fingerprinting; filamentous fungi; *Penicillium chrysogenum*; *Trichoderma reesei* Rut-C30; HPLC-SEC

1. Introduction

Bioprocesses are dynamic in nature with varying process conditions rendering inconsistent product quality. Process variability arises from changes in critical process parameters (CPPs) and critical material attributes (CMAs) affecting key performance indicators (KPIs) and critical quality attributes (CQAs) [1]. Therefore, process monitoring is of utmost importance to monitoring and controlling changes in KPIs and CQAs to deliver consistent product quality. Furthermore, complex expression systems, such as filamentous fungi, require cumbersome offline methods (e.g., staining) to monitor process variables (e.g., cell viability [2,3]).

Viable biomass is one of the most important process variables in bioprocesses. Its reliable estimation allows the determination of other essential variables for process understanding, such as growth rates, substrate uptake rates and biomass yield [4]. For filamentous bioprocesses, performance, control strategies and productivity highly depend on cellular aspects, which calls for a segregated view of biomass. The determination of viable biomass concentration via chemical methods such as fluorescence staining using propidium iodide (PI) or fluorescein diacetate (FDA) is accurate but time consuming. PI cannot cross the membrane of living cells and FDA is only hydrolyzed by metabolically active cells, making both stains useful for determining viability. In an industrial setting, chemical analytical methods are not preferred in comparison to physical techniques, which are capable of real time measurement [5]. Thoroughly reviewed methods for measuring viable biomass include dielectric spectroscopy, infrared spectroscopy and fluorescence [6,7]. However, inline sensors are prone to high measurement noise and require chemometric knowledge to establish meaningful measurement techniques. As filamentous organisms tend to develop special morphological forms consisting of compact hyphal aggregation [8,9], process monitoring strategies are further complicated. For *Penicillium chrysogenum* and *Trichoderma reesei* bioprocesses, this special morphology (known as "pellets") featuring dense biomass clumps rather than loose mycelia results in low mixing times and improved gas–liquid mass transfer. But pellet morphology also leads to limitations in the transport of substrates and oxygen [10], which negatively affects biomass growth and productivity. As a result, pellets need to be compact enough to ensure a compact and productive biomass density and small enough to avoid diffusional limitations in the pellet's core. This balance is commonly controlled by optimized agitation conditions, medium composition or spore inoculum levels. If mass transport into the pellet cannot be maintained, the biomass will exhibit hyphal degradation beginning in the pellet's core and a decline in overall viability [11]. Consequently, most contributions dealing with the assessment of viable biomass in filamentous cultivations identify a growth phase and a decline phase. The measurable onset of a viability drop initiates the decline phase. Employing capacitance-based probes growth and decline phases can be differentiated by an increase of conductivity, with error prone results, however [4,5]. Reliable monitoring strategies capable of identifying the onset of a cultivation's decline phase are essential in order to avoid over-feeding and further decline of biomass viability. In this respect, the relationship between substrate availability and oxygen consumption is also a most relevant factor in process control: limiting substrate feeding regimes can positively affect productivity in secondary metabolite production while ensuring high viability due to less substrate oxidation and less oxygen consumption within the pellet [12,13].

Spectroscopic and chromatographic data have been used in combination with statistical models for process monitoring strategies and quantifying process variables. Optical sensors have been widely used to measure and monitor different process variables, such as analyte concentration (e.g., product), product quality attributes (e.g., glycosylation) and cell level responses (e.g., cell sub-populations) [14–17]. FDA promotes the implementation of process analytical technology (PAT) and quality by design (QbD) in each unit operation of a bioprocess to monitor and control critical quality attributes (CQAs) [18,19]. In the biotech industry, such multivariate data analysis (MVDA) techniques are gaining acceptance and are implemented in various leading pharmaceutical companies [20,21]. UV chromatography is one of the most commonly used techniques in various bioanalytical assays. Recently, we employed UV chromatography coupled with chemometric approaches to monitor cell lysis in *Escherichia coli* bioprocesses [22] and for process development of downstream unit operations [23,24]. In a similar approach, we used principal component analysis (PCA) with UV chromatographic data of samples from twelve *P. chrysogenum* bioprocesses to monitor cell viability. In contrast to the previous study, where we used a strong anion exchange monolithic column (CIMac QA, BIA separations, Sloveina), in this study we used a size exclusion chromatography analytical column for better resolution of the protein and nucleic acid profiles. Furthermore, the predictive power of the model was tested using partial least squares (PLS), orthogonal partial least squares (OPLS) and principal component regression (PCR). Based on the model results, the drop in cell viability was identified, and thereby used to define

the optimal time point of harvest or measures to maintain high viability through process control; for instance, feeding profiles, power input and dissolved oxygen content [13].

To summarize, in this study we exploit the use of data-driven models (DDM) for bioprocess monitoring in complex expression systems using HPLC fingerprints as a versatile PAT tool. We analyzed cultivation samples using a simple HPLC setup equipped with a SEC column. UV chromatogram fingerprints at 260 nm with statistical models were used to predict and identify the decline of cell viability in *P. chrysogenum* processes. The model results are compared and verified with state-of-the-art viability assessment methods. To show the versatility of the methodology, we implemented the developed workflow for another filamentous fungus strain, *T. reesei* Rut-C30, an industrial workhorse for the production of cellulolytic enzymes [25].

2. Materials and Methods

2.1. Bioreactor Cultivations

2.1.1. Bioreactor Set-Up

P. chrysogenum cultivations were performed either in a Techfors S bioreactor (Infors HT, Bottmingen, Switzerland, with 10 L maximal working volume) or in a DASGIP Mini parallel reactor system (working volume 4 * 2.0 L, Eppendorf, Germany). All *T. reesei* cultivations were performed in the aforementioned Techfors S bioreactor. The stirrer was equipped with three six bladed Rushton turbine impellers, of which two were submersed and one was installed above the maximum liquid level for foam destruction. For supplying pressurized air and oxygen (O_2) Four aeration mass flow controllers (Vögtlin, Aesch, Switzerland) were used. Dissolved oxygen concentration (DO_2) was measured using a dissolved oxygen probe (Hamilton, Bonaduz, Switzerland). pH was measured using a pH probe (Hamilton, Bonaduz, Switzerland). CO_2 and O_2 concentrations in the off-gas were analyzed with an off-gas analyser (M. Müller AG, Egg, Switzerland).

2.1.2. *P. Chrysogenum*

Samples from *P. chrysogenum* cultivations from both small scale (SS; working volume 2 L) and laboratory scale (LS; working volume 10 L) setups were used for UV chromatographic data acquisition. A total of nine small scale cultivations were tested in a DASGIP Mini parallel reactor system (Eppendorf, Germany), and three lab scale cultivations were tested in a Techfors S bioreactor (Infors HT, Bottmingen, Switzerland). The process profiles for the different scales were similar. In general, the batch was inoculated with approximately $2{\cdot}10^8$ spores$\cdot L^{-1}$. During batch phase, pH was not controlled. The end of the batch was defined as an increase in pH of 0.5 by convention. After the batch, the broth was diluted with fed-batch medium (15% broth, 85% medium) and fed-batch was started. Details on batch and fed-batch media can be found in [26]. The fed-batch process lasted for approximately 150–170 h. Temperature was maintained at 25 °C and pH was controlled at 6.5 ± 0.1 by addition of 20% (*w*/*v*) KOH or 15% (*v*/*v*) H_2SO_4, respectively. pH was measured using a pH probe (Hamilton, Bonaduz, Switzerland). After an additional 12 h, nitrogen and phenoxyacetate feeds were started at constant rates (6.5 mL$\cdot h^{-1}$ for nitrogen and 2 mL$\cdot h^{-1}$ for phenoxyacetate).

A feed-forward controller was implemented to maintain a constant specific glucose uptake rate of biomass q_S. Aeration was controlled at 1 vvm in batch and initial fed-batch. Dissolved oxygen concentration was controlled between 40% and 90% during the batch phase and at the set-points 5.0, 22.5% or 40.0% during fed-batch, via adjustment of the gas mix using pressurized air, nitrogen and oxygen. The different q_S and dissolved oxygen values are listed in Appendix A Table A1. The agitation conditions were maintained at 325–700 rpm stirring speed in all process phases.

2.1.3. *T. reesei* Rut-C30

For testing the versatility of our tool, a cultivation process with another industrially relevant strain *T. reesei* Rut-C30 was done. The optimized media recipe for cultivation has been published elsewhere [27]. In short, the pre-culture medium was supplemented with 10 $g \cdot L^{-1}$ glucose and 1 $g \cdot L^{-1}$ peptone from casein. Batch-medium initially contained 10 $g \cdot L^{-1}$ lactose, 0.5 $g \cdot L^{-1}$ urea, 2 $g \cdot L^{-1}$ $(NH_4)_3SO_4$, 2 $g \cdot L^{-1}$ KH_2PO_4, 0.5 $g \cdot L^{-1}$ $MgSO_4 \cdot 7H_2O$ and 0.5 $g \cdot L^{-1}$ $CaCl_2 \cdot 2H_2O$ mixed with 0.5 $L \cdot L^{-1}$ 0.2 M Na_2-HPO_4-citric acid buffer (pH 5.0).

Next, a 500 mL pre-culture was inoculated with $5 \cdot 10^8$ spores$\cdot L^{-1}$ equally split in two 1000 mL Erlenmeyer shake flasks. After 24 h at 28 °C and 180 rpm on a rotary shaker (Infors HT, Bottmingen, Switzerland) the pre-culture was transferred to inoculate 4.5 L batch-medium in the reactor. During the whole cultivation, the pH was constantly controlled at 5.0 ± 0.05 by automatic addition of 20% (*w*/*v*) KOH or 20% (*v*/*v*) H_2SO_4. Following a drop in the CO_2 off-gas signal indicating the end of batch, a fed-batch with $(NH_4)_3SO_4$ and 200 $g \cdot L^{-1}$ lactose feed was started. The initial specific lactose uptake rate of biomass q_S was set to 0.18 $g_{Lac} \cdot g_X^{-1} \cdot h^{-1}$ and fed isocratic until end of process with an average q_S of 0.05 $g_{Lac} \cdot g_X^{-1} \cdot h^{-1}$. The broth was held at 28 °C, pressurized at 1 bar and constantly aerated by 1.0 vvm. The stirrer was set to 600 rpm during the batch phase and 900 rpm during fed-batch phase. The dissolved oxygen level was always controlled above 40% by the addition of pure O_2 in the gas flow.

2.2. Viability Assays

The following published methods ([11,28]) for viability assessment were used in method development for verification purposes.

2.2.1. *P. Chrysogenum*

PI Staining

The membrane impermeable dye PI binds to DNA. If subsequently excited at wavelengths of 488 nm, PI will emit in the red spectral section. This characteristic is used for viability assessment according to the following method: Viability is estimated as a ratio between the fluorescence intensity of an untreated sample and a microwaved, and hence non-viable negative control.

To investigate viability via propidium iodide (PI) staining according to [11], 200 µL of sample was diluted 1:5 with phosphate buffered saline (PBS, 8 $g \cdot L^{-1}$ NaCl, 0.2 $g \cdot L^{-1}$ KCl, 1.44 $g \cdot L^{-1}$ and 0.24 $g \cdot L^{-1}$ KH_2PO_4, see [29]). In addition, 1 mL of sample was diluted 1:5 with PBS and microwave treated by leaving it for 30 s at 940 W in a M510 microwave oven (Philips, Amsterdam, The Netherlands). One milliliter of the microwave-treated sample was used for further investigation. In a next step, duplicates of all samples (including microwave-treated and untreated samples) were centrifuged for 15 min at 500 rpm. 800 µL of supernatant was removed and 800 µL of PBS buffer was added. The pellet was resuspended, and the washing step repeated; 100 µL of the resuspended sample was pipetted into a microtiter well, and 1 µL of 200 µM PI solution (Sigma Aldrich, St. Louis, MO, USA) was added. The PI was prepared by diluting a 20 mM PI stock solution in DMSO, 1:100 in PBS. After an incubation time of 20 min at room temperature in darkness, the measurement was performed in a Tecan well-plate reader (Tecan, Männedorf, Switzerland; ex./em. 535/600 nm). Each sample was measured six times simultaneously using 96 well plates. Viability assessment was subsequently performed according to the following equation:

$$\text{Via} = 1 - (\text{FL\_red}_{\text{native}} / \text{FL\_red}_{\text{microwaved}}) \quad (1)$$

where Via is the viability, $\text{FL\_red}_{\text{native}}$ is the red fluorescence signal of the untreated sample and $\text{FL\_red}_{\text{microwaved}}$ is the red fluorescence of the microwaved negative control. Viability was measured in six replicates, leading to a maximum error of 5% for each sample.

2.2.2. *T. reesei* Rut-C30

FDA Staining

Assessment of viability was performed via fluorescein diacetate staining. FDA is a non-fluorescent molecule. Esterase activity in live cells leads to hydrolyzation of FDA, resulting in fluorescent fluorescein [30,31].

For viability staining, 500 µL sample was diluted 1:10 in phosphate-buffered saline (PBS, 8 $g \cdot L^{-1}$ NaCl, 0.2 $g \cdot L^{-1}$ KCl, 1.44 $g \cdot L^{-1}$ and 0.24 $g \cdot L^{-1}$ KH_2PO_4). In total, 490 µL of this solution was incubated with 10 µL of 12 mM FDA in an acetone solution for 5 min in the dark at room temperature prior to flow cytometry analysis. The calculated viability is the ratio of metabolically active cells to the total number of cells, similar to that described by [28]. Detailed information about the used CytoSense flow cytometer (CytoBuoy, Woerden, Netherlands) is described elsewhere [30,31].

2.3. Data Analysis

2.3.1. Data Acquisition

P. Chrysogenum

Samples from twelve (9 SS and 3 LS) *P. chrysogenum* cultivations were used for acquiring UV chromatographic data through size exclusion chromatography (SEC). UV chromatographic data have been shown to contain process information with respect to nucleic acids (having maximum absorbance at 260 nm) and protein impurities (having maximum absorbance at 280 nm) [23,24]. Significant changes in the impurity release profiles, especially during metabolic stress and viability decline phases, can be used to monitor process performance. The total number of samples from the SS and LS runs were 189 in total. UV chromatographic data at 260 nm were recorded using a modular HPLC device (PATfinder^{TM}) purchased from BIAseparations (Ajdovscina, Slovenia). The setup comprised an autosampler (Knauer Optimas), a pump (Azura P 6.1 L) and a UV detector (Azura MWD 2.1 L). The samples were loaded onto a Tosho TSKgel G3000SWxl size exclusion chromatography (SEC) column purchased from Tosho Bioscience LLC (Tokyo, Japan). A loading buffer with 20 mM potassium phosphate, 150 mM sodium chloride, pH 7.0, was used. The flow velocity was kept constant at 0.75 $mL \cdot min^{-1}$. All samples were centrifuged and filtered with a 0.22 µm PVDF filter. Random samples (one in every 10 samples) were injected twice to ensure reproducibility and quality of the UV chromatographic data.

T. reesei Rut-C30

The collected samples from *T. reesei* Rut-C30 fed-batch were analyzed using the HPLC Dionex UltiMate 3000 system (Thermo Fisher Scientific, MA, USA) equipped with autosampler, pump and UV detector. The SEC column BioBasic SEC-300 x 4.6 mm (Thermo Fisher Scientific, MA, USA) heated to 30 °C was loaded with 5 µL of centrifuged and 0.22 µm PTFE filtered supernatant and run in isocratic operation mode by 0.3 $mL \cdot min^{-1}$ 20 mM K_3PO_4, 150 mM NaCl pH 7.0 buffer. Data were acquired at 260 nm by UV detection.

2.3.2. Data Pre-Processing

UV chromatographic data are prone to data misalignments and shifts along the retention time. Therefore, several pre-processing steps are necessary prior to PCA modelling. As described in previous studies [32], we used the optimal correction algorithm to correct misalignments in the raw UV chromatographic data. Three alignment techniques, namely, icoshift [33], peak alignment using fast Fourier transform (PAFFT) and recursive alignment using fast Fourier transform (RAFFT) [34], were screened, and the optimal correction algorithm was chosen as described in [35]. The filamentous fungi cultivations had variations in the estimated viability from the offline analytics; therefore, a smoothing

spline method was used to correct offline data prior to predictive models. All UV chromatographic data were scaled and centered prior to establishing predictive models.

2.3.3. Descriptive Analysis (PCA)

PCA is one of the most commonly used chemometric techniques for compressing high volumes of process data (e.g., spectroscopic sensors [36–38]) into few meaningful process features. We used PCA models to identify process trends using UV chromatographic dataset at 260 nm. We chose to use UV chromatographic data at 260 nm, since the nucleic acids have maximal absorbance at 260 nm and can be used for detecting and predicting viability decline. For clarity, we want to capture the differences in the nucleic acid release pattern along the process using UV chromatographic data and chemometric models. In short, PCA is an exploratory technique which decomposes the entire chromatographic dataset to a few latent principal components. In a PCA model with UV chromatographic data, each sample is represented as a score and is projected across different principal components (PCs) based on its similarities or differences. The resulting score plots from the PCA model can be used to identify possible groupings or trends between samples in the UV chromatographic data. The loadings explain the retention time at which variance in the chromatographic data was significant. In general, the first PCs explain most of the variance in the chromatographic dataset. PCA has been widely reviewed for applications in process development and production [37,39,40].

2.3.4. Predictive Analysis

Three different modelling techniques were used for the prediction of viability using UV chromatogram fingerprints at 260 nm; namely, partial least squares (PLS), orthogonal PLS (OPLS) and principle component regression (PCR). The modelling techniques have been well defined and explained in many publications [41–46]. PLS is the most commonly used multivariate method to assess the relationship between a descriptor matrix X and the response matrix Y. PLS is usually used for prediction of quantitative Y data; however, qualitative Y data can be used for discriminant analysis (PLS-DA). OPLS is an extension of the supervised PLS regression. In simple words, OPLS uses information from the Y matrix to decompose the X matrix into blocks of variation correlated and orthogonal to the Y matrix. In PCR, as a first step the UV chromatographic dataset is rendered as a PCA model and the scores from the model are used to predict the viability. The model results were evaluated based on the root mean squared error of estimation (RMSEE) and 7-fold cross validation (RMSEcv). The workflow for data acquisition, pre-processing, descriptive and predictive analysis was applied to the *T. reesei* data to present the versatility of the tool.

2.3.5. Software

Pre-processing of chromatographic data, namely, peak alignment using correction techniques and offline data correction, were done in MATLAB R2019a version 9.6 (Mathworks, MA, USA). PCA, PLS and OPLS models were established in SIMCA v15.0.2 (Umetrics, Umea, Sweden). PCR models were established in Python (using SpyDer version 3.3.6; distributed under the terms of the MIT License).

3. Results

3.1. Data Acquisiton

A total of 189 samples were drawn from small-scale and laboratory oratory scale runs from the *P. chrysogenum* bioprocesses for offline and at-line analyses. Cell viability, biomass, product and substrate concentrations with their respective rates and yields were calculated using standard analytical techniques. An example time course of cell viability measured via PI treatment using a plate reader as explained in Section 2.2.1 from one small-scale and laboratory scale run, is shown in Figure 1. Both runs were conducted at a maximum q_s setpoint of over 0.05 $g_X \cdot g_S^{-1} \cdot h^{-1}$. While the small-scale cultivation's q_s setpoint could not be sustained due to a continuous loss in viability, the laboratory

scale run conducted at a high average q_s value was stopped before a massive drop in viability occurred. For comparison, the small-scale run SS4 was conducted at a low q_s at consistently high viability. This emphasizes that lower q_s values help to sustain culture viability, as explained in our previous work: using a design of experiments (DoE) approach, we demonstrated the positive effect of lower qs setpoints in a reproducible manner [13].

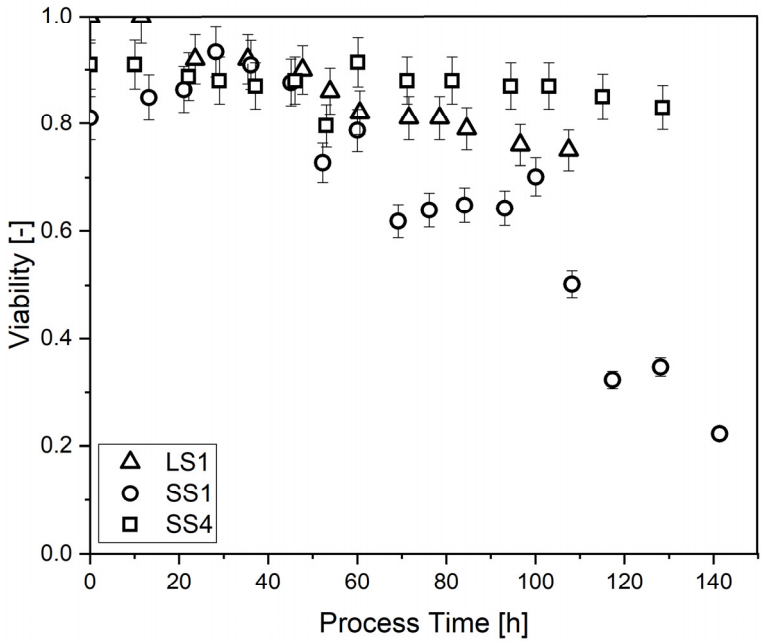


Figure 1. Exemplary drop in viability in *Penicillium chrysogenum* bioprocesses measured via PI staining at prolonged high q_s values in small-scale run 1; laboratory scale run 1 was stopped at the onset of viability decline. For comparison, low q_s values enabled consistently high viability in small-scale run 4.

3.2. Data Pre-Processing

Raw UV chromatograms at 260 nm were acquired using either the modular HPLC setup or a Thermo system with a size exclusion chromatography (SEC) column. Shifts along the retention time in the UV chromatograms as fingerprints were corrected using the PAFFT algorithm [47]. The comparison of the raw data and the pre-processed data is shown in Figure 2.

Peak artefacts and shoulder peaks can be seen in the icoshift and RAFFT algorithms; this is mainly due to aligning the tallest peak from all samples, where the entire data is shifted to give maximum correlation with respect to alignment of the maximum absorbance. It can be inferred from the heatmaps (Figure 2E–H) that PAFFT has removed misalignments, and therefore was chosen for establishing descriptive and predictive models. The offline viability data were corrected to remove noisy measurements using a smoothing spline function. The raw and smoothed viability data are shown in Figure A1.

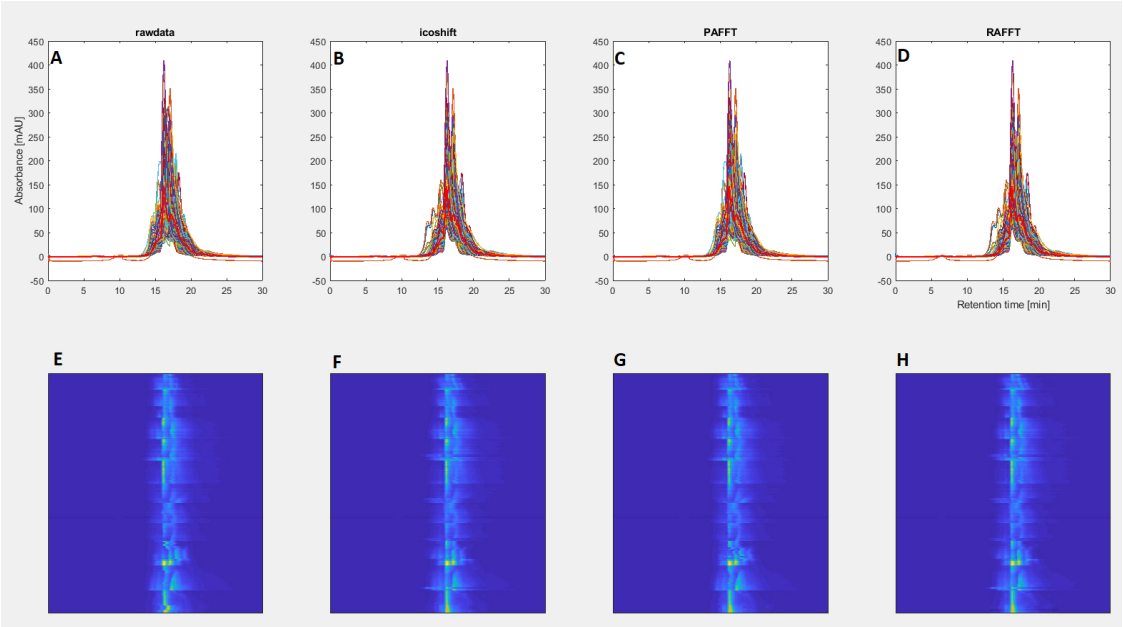


Figure 2. Data alignment correction of UV chromatographic datasets from twelve *P. chrysogenum* cultivations. A, Raw UV data; B, icoshift correction; C, PAFFT correction; D, RAFFT correction; E–H heatmaps of the raw data and correction methods respectively.

3.3. Descriptive Analysis

Descriptive analysis was done as a first step on the UV chromatographic dataset from *P. chrysogenum* processes using PCA models. Three PCA models were developed on (1) the small-scale runs, (2) the laboratory scale runs and (3) the entire dataset to analyze the intrinsic variability between the samples and cultivations. The score plot from the entire dataset is shown in Figure 3.

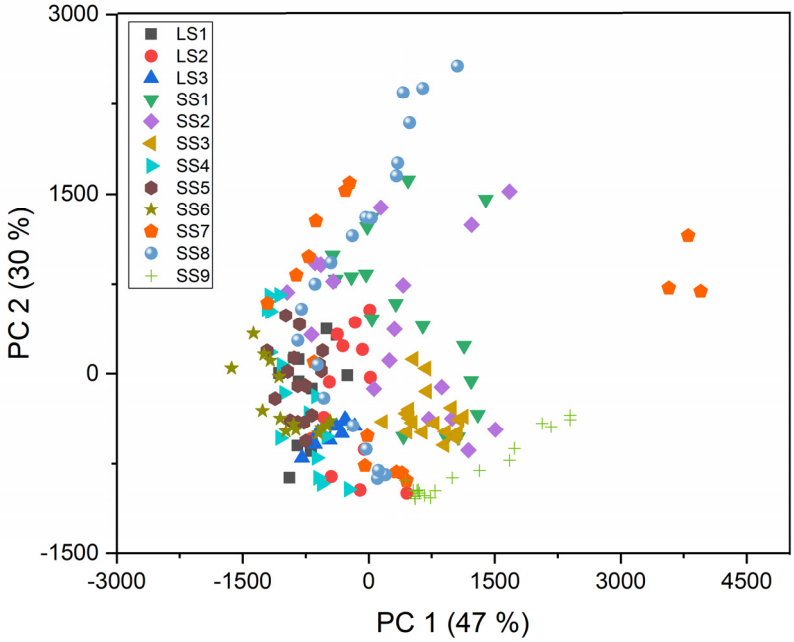


Figure 3. Score plot from the principal component analysis (PCA) model for all 189 samples from the *P. chrysogenum* bioprocesses.

Explained variance from the first six PCs for the entire dataset were 47%, 30%, 10%, 6%, 3% and 2% (adding up to 96%); the remaining PCs were discarded. The PCA model with the entire dataset shows an aggregated cluster at the center of the score plot. The processes, namely, SS 7, 8 and 9 have score spreading upwards and away from the aggregated cluster. It is interesting to note that the aforementioned processes have the lowest viability values (as shown in Figure A1). The scores from the PC2 of the PCA model were plotted across process time, as shown in Figure A2. The scores from the LS runs have a similar trend, since the processes were run under similar conditions; however, huge variability can be seen among the SS runs. We speculate the differences in the feeding regimes, which in turn have dilution effects, caused said high variability in the SS runs. Nevertheless, using descriptive analysis results, a golden batch approach (e.g., using exponentially weighted moving average (EWMA)) can be used to set the standard deviation ranges from run-of-the-mill processes, and significant process deviations can be analyzed. Furthermore, the PCA models were used to detect the outliers from the UV chromatographic datasets based on the distance to model (DmodX) values. All samples which had a DmodX values twice that of Dcrit were removed for further predictive analysis.

3.4. Predictive Analysis

Three predictive modelling techniques, namely, PLS, OPLS and PCR, were used to predict the raw offline viability and the smoothed viability measurements based on the UV chromatographic datasets. The prediction results for both raw and smoothed viability values from the aforementioned modelling techniques for all samples from *P. chrysogenum* cultivations are shown in Figure 4.

Overall, the OPLS models showed best predictive results with an normalized root mean squared error of cross validation (NRMSEcv) of 0.10 and 0.07 for the raw and smoothed viability measurements respectively. It is important to note that irrespective of the scales the model was able to predict cell viability with an accuracy of 90%. The PLS and PCR models showed close prediction accuracy to the OPLS models. The NRMSEcv of the PLS models were 0.11 and 0.08 for the raw and smoothed viability values, and for PCR models they were 0.12 and 0.10 respectively. The NRMSEcv of the two response variables for all models are shown in Appendix A Table A2. The obtained information can be used to detect the onset of a drop in viability and subsequently avoid further decline via adjustment of fermentation parameters, such as the feed rate.

3.5. Tool Versatility

The developed tool and methodology were implemented in *T. reesei* Rut-C30 bioprocesses to test its versatility. The PCA model showed a clear trend with respect to process time, as shown in Appendix A Figure A2. The predictions for the cell viability based on PLS, OPLS and PCR models rendered accurate results, with PLS having an RMSEcv of 0.05, OPLS—0.07 and PCR—0.07. The offline measurements and the predictions from the PLS model over process time are shown in Figure 5.

However, we envision that a robust prediction model could be developed for *T. reesei* processes with higher sample numbers, earlier decline in viability and changing process conditions. The implementation of the workflow showed promising results for this additional organism, highlighting the use of UV chromatographic data from HPLC-SEC for a broader application in filamentous fungi processes.

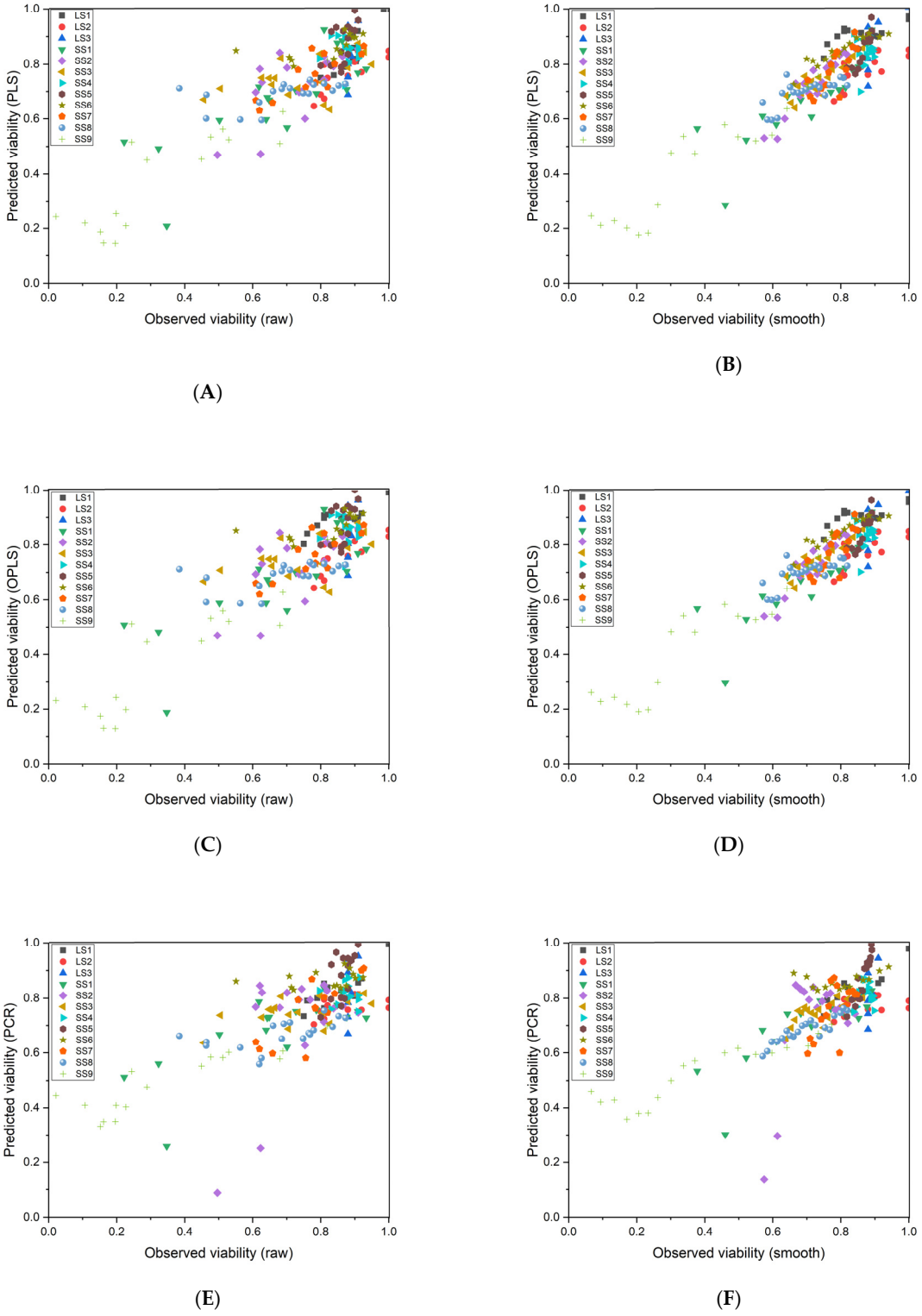
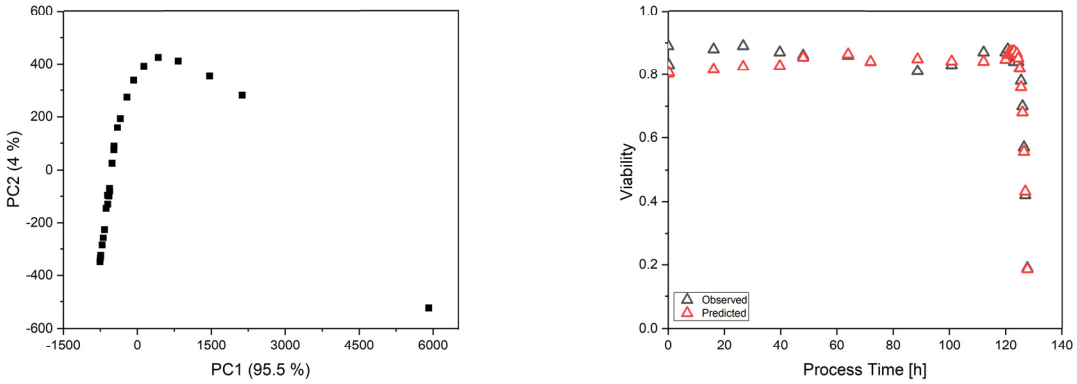


Figure 4. Results from the predictive results of the three chosen modelling techniques for the raw and smoother viability measurements. (**A**,**C**,**E**) Observed vs. predicted for raw viability measurements from partial least squares (PLS), orthogonal PLS (OPLS) and principal component regression (PCR) models respectively. (**B**,**D**,**F**) observed vs predicted for smoothed viability measurements from PLS, OPLS and PCR models respectively.



(**A**) (**B**)

Figure 5. The results from the descriptive and predictive models for the *Trichoderma reesei* cultivations. (**A**) Score plot of the PCA model, and (**B**) observed and predicted viability values.

4. Discussion and Conclusions

UV chromatographic data have been widely used for process monitoring in upstream cultivations and process development in downstream unit operations. With the rising advances in online liquid handling systems and supervisory data analytical methodologies, UV chromatographic data as fingerprints can be exploited for process monitoring, online state estimation and eventually, process control. In this study, we used UV chromatographic samples from different scales and organisms to descriptively analyze process trends, and using supervised prediction models, predicted the cell viability. Although numerous sophisticated techniques are available based on conductivity, dielectric spectroscopy and RAMAN spectroscopy for prediction of cell viability and monitoring bioprocesses, these techniques require expensive hardware. The HPLC-SEC UV chromatographic data at 260 nm and 280 nm contain information regarding the nucleic acids and protein release profiles from the process. In process optimization, the descriptive analysis can be used to follow process trends and identify potential deviations, especially in pilot scale or large-scale production runs. Numerous statistical methods are available to establish boundaries (usually ± 3 SD with a EWMA), and potential deviations can be monitored and acted upon in a timely fashion.

Diffusional limitations within a fungal pellet primarily involve oxygen and occur in dense biomass structures. However, it was shown that lower substrate availability decreases the consumption of oxygen and can enhance pellet viability [12] as well. Consequently, a decrease in viability of *P. chrysogenum* pellets could be detected and moderated via adjustments of the feeding profile, as previously shown [13]. For this purpose, the chromatographic UV datasets show high predictive power, as reported in the results section.

The descriptive score trends from the LS runs are shown in Figure A2. The OPLS models showed high precision for predicting cell viability, and the methodology has been shown to work for another filamentous fungi process; namely, *T. reesei*. Results from the prediction models from *T. reesei* further highlighted the platform applicability of the presented methodology. Prediction models coupled with online HPLC devices can pave way for predicting the cell viability in real time. Product concentration and potential impurity information can be captured using the HPLC data, and feed-rates can be controlled to boost productivity.

We further envision that supervised classification models could be used to distinguish different phases of the process, and with the use of mechanistic descriptors, hybrid models could be used to simulate the rate of decline in viability and thereby enable process control. All forms of analytical data can be combined to holistically analyze the information gaps in the process, and promising modelling techniques can be used to extract maximal information from such processes. Potential deviations can

be encountered early on, and using structured risk-assessment and mitigation tools, the causes for such deviations can be analyzed.

Author Contributions: Conceptualization, C.H., O.S. and V.R.; cultivation, P.D. and L.V.; data analysis, P.D., L.V. and V.R.; writing—original draft preparation, P.D., L.V. and V.R.; writing—review and editing, O.S., C.H., P.D., L.V. and V.R.; visualization, V.R.; supervision, C.H., O.S. and V.R.; funding acquisition, O.S. and C.H. All authors have read and agreed to the published version of the manuscript.

Funding: We thank the Austrian Ministry of Education, Science and Research and the Christian Doppler Laboratory for Mechanistic and Physiological Methods for Improved Bioprocesses for financial support. Further, we thank the TU Wien for funding the doctoral college bioactive. The authors acknowledge TU Wien Bibliothek for financial support through its Open Access Funding by TU Wien. The research article was partly funded in terms of employment (for L.V.) within the framework of Competence Center CHASE GmbH, funded by the Austrian Research Promotion Agency (grant number 868615) as part of the COMET program—Competence Centers for Excellent Technologies by BMVIT, BMDW, the Federal Provinces of Upper Austria and Vienna.

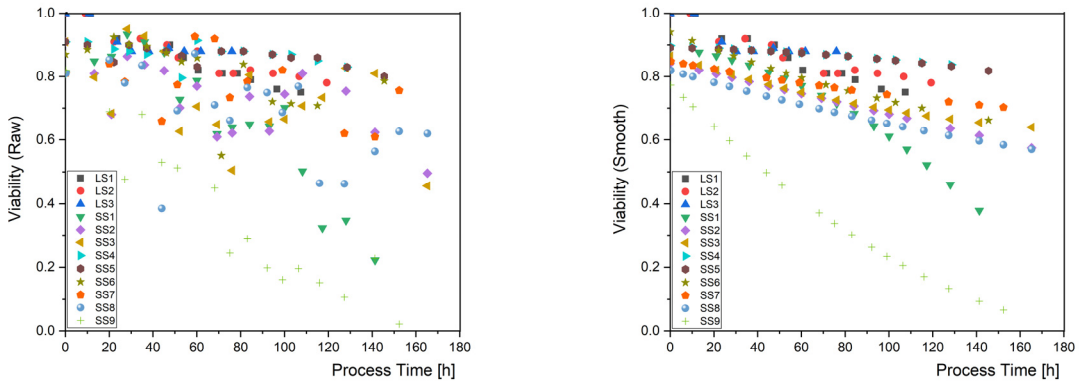
Acknowledgments: Strains for the *P. chrysogenum* experiments were kindly provided by Sandoz GmbH (Kundl, Austria). *T. reesei* Rut-C30 strains were kindly provided by Christian Derntl, TU Wien.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. The average substrate uptake rates q_s, average dissolved oxygen content and average viability values of the *P. chrysogenum* cultivations from small scale (SS) and large scale (LS) cultivations. Please note that some cultivations were stopped at the onset of a viability decline; therefore, the average viability is relatively high at a comparatively low standard deviation.

Name	Average q_s [$g_s/g_x/h$]	Average Dissolved Oxygen Content [%]	Average Viability [-]
LS1	0.054 ± 0.005	40.0 ± 5.4	0.79 ± 0.05
LS2	0.045 ± 0.004	40.0 ± 5.2	0.82 ± 0.09
LS3	0.017 ± 0.003	40.0 ± 5.1	0.88 ± 0.05
SS1	0.042 ± 0.004	40.0 ± 6.5	0.63 ± 0.24
SS2	0.038 ± 0.003	22.5 ± 4.1	0.68 ± 0.18
SS3	0.015 ± 0.004	05.0 ± 0.5	0.70 ± 0.17
SS4	0.026 ± 0.003	22.5 ± 3.9	0.88 ± 0.06
SS5	0.035 ± 0.003	22.5 ± 6.9	0.87 ± 0.03
SS6	0.018 ± 0.001	5.0 ± 0.5	0.78 ± 0.12
SS7	0.034 ± 0.005	22.5 ± 6.6	0.79 ± 0.06
SS8	0.033 ± 0.005	22.5 ± 5.0	0.78 ± 0.10
SS9	0.040 ± 0.012	5.0 ± 0.5	0.35 ± 0.18



(**A**) (**B**)

Figure A1. (**A**) The raw viability and (**B**) the smoothed viability data from *P. chrysogenum* bioprocesses.

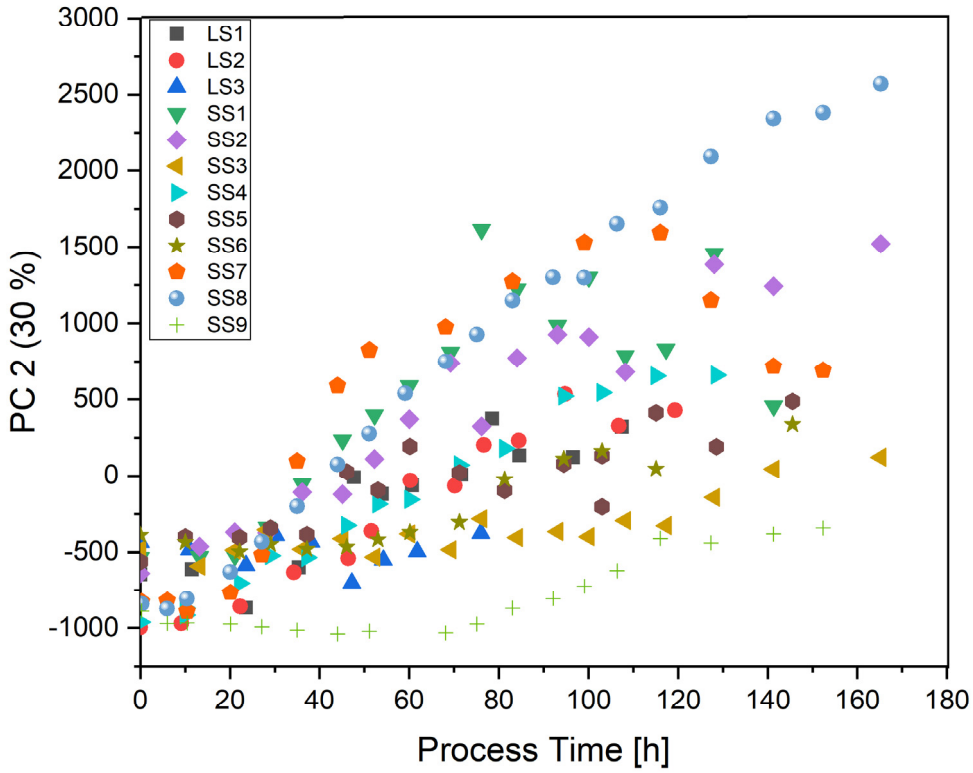


Figure A2. The score trends of the second principal component from the PCA model across process time.

Table A2. The Normalized root mean squared error of cross validation (NRMSEcv) for raw and smoothed viability measurements from the PLS, OPLS and PCR models.

	Raw Viability	Smoothed Viability
PLS	0.11	0.08
OPLS	0.10	0.07
PCR	0.12	0.10

References

1. ICH. *International Conference on Harmonisation Draft Guidance: Q8(R2) Pharmaceutical Development Revision 2*; D.o.H.a.H. Services: Rockville, MD, USA, 2009.
2. Ehgartner, D.; Fricke, J.; Schröder, A.; Herwig, C. At-line determining spore germination of Penicillium chrysogenum bioprocesses in complex media. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 8923–8930. [CrossRef] [PubMed]
3. Veiter, L.; Rajamanickam, V.; Herwig, C. The filamentous fungal pellet-relationship between morphology and productivity. *Appl. Microbiol. Biotechnol.* **2018**, *102*, 2997–3006. [CrossRef] [PubMed]
4. Neves, A.A.; Pereira, D.A.; Vieira, L.M.; Menezes, J.C. Real time monitoring biomass concentration in Streptomyces clavuligerus cultivations with industrial media using a capacitance probe. *J. Biotechnol.* **2000**, *84*, 45–52. [CrossRef]
5. Rønnest, N.P.; Stocks, S.M.; Lantz, A.E.; Gernaey, K.V. Introducing process analytical technology (PAT) in filamentous cultivation process development: Comparison of advanced online sensors for biomass measurement. *J. Ind. Microbiol. Biotechnol.* **2011**, *38*, 1679–1690. [CrossRef]
6. Kiviharju, K.; Salonen, K.; Moilanen, U.; Eerikäinen, T. Biomass measurement online: The performance of in situ measurements and software sensors. *J. Ind. Microbiol. Biotechnol.* **2008**, *35*, 657–665. [CrossRef]
7. Kiviharju, K.; Salonen, K.; Moilanen, U.; Meskanen, E.; Leisola, M.; Eerikäinen, T. On-line biomass measurements in bioreactor cultivations: Comparison study of two on-line probes. *J. Ind. Microbiol. Biotechnol.* **2007**, *34*, 561–566. [CrossRef]

8. Ehgartner, D.; Herwig, C.; Fricke, J. Morphological analysis of the filamentous fungus Penicillium chrysogenum using flow cytometry-the fast alternative to microscopic image analysis. *Appl. Microbiol. Biotechnol.* **2017**, *101*, 7675–7688. [CrossRef]
9. Ehgartner, D.; Sagmeister, P.; Herwig, C.; Wechselberger, P. A novel real-time method to estimate volumetric mass biodensity based on the combination of dielectric spectroscopy and soft-sensors. *J. Chem. Technol. Biotechnol.* **2014**, *90*, 262–272. [CrossRef]
10. Dynesen, J.; Nielsen, J. Surface Hydrophobicity of Aspergillus nidulans Conidiospores and Its Role in Pellet Formation. *Biotechnol. Prog.* **2003**, *19*, 1049–1052. [CrossRef]
11. Ehgartner, D.; Hartmann, T.; Heinzl, S.; Frank, M.; Veiter, L.; Kager, J.; Herwig, C.; Fricke, J. Controlling the specific growth rate via biomass trend regulation in filamentous fungi bioprocesses. *Chem. Eng. Sci.* **2017**, *172*, 32–41. [CrossRef]
12. Bodizs, L.; Titica, M.; Faria, N.; Srinivasan, B.; Dochain, D.; Bonvin, D. Oxygen control for an industrial pilot-scale fed-batch filamentous fungal fermentation. *J. Process. Control.* **2007**, *17*, 595–606. [CrossRef]
13. Veiter, L.; Kager, J.; Herwig, C. Optimal process design space to ensure maximum viability and productivity in Penicillium chrysogenum pellets during fed-batch cultivations through morphological and physiological control. *Microb. Cell Factories* **2020**, *19*, 1–14. [CrossRef] [PubMed]
14. Buckley, K.; Ryder, A.G. Applications of Raman Spectroscopy in Biopharmaceutical Manufacturing: A Short Review. *Appl. Spectrosc.* **2017**, *71*, 1085–1116. [CrossRef] [PubMed]
15. Golabgir, A.; Gutierrez, J.M.; Hefzi, H.; Li, S.; Palsson, B.O.; Herwig, C.; Lewis, N.E. Quantitative feature extraction from the Chinese hamster ovary bioprocess bibliome using a novel meta-analysis workflow. *Biotechnol. Adv.* **2016**, *34*, 621–633. [CrossRef] [PubMed]
16. He, Y.; Friese, O.; Schlittler, M.R.; Wang, Q.; Yang, X.; Bass, L.A.; Jones, M.T. On-line coupling of size exclusion chromatography with mixed-mode liquid chromatography for comprehensive profiling of biopharmaceutical drug product. *J. Chromatogr. A* **2012**, *1262*, 122–129. [CrossRef] [PubMed]
17. Luoma, P.; Golabgir, A.; Brandstetter, M.; Kasberger, J.; Herwig, C. Workflow for multi-analyte bioprocess monitoring demonstrated on inline NIR spectroscopy of P. chrysogenum fermentation. *Anal. Bioanal. Chem.* **2016**, *409*, 797–805. [CrossRef]
18. Rathore, A.S.; Bhambure, R.; Ghare, V. Process analytical technology (PAT) for biopharmaceutical products. *Anal. Bioanal. Chem.* **2010**, *398*, 137–154. [CrossRef]
19. FDA. *Guidance for Industry PAT: A Framework for Innovative Pharmaceutical Development, Manufacuring, and Quality Assurance*; FDA Official Document: Rocksville, MD, USA, 2004; p. 16.
20. Rathore, A.S.; Agarwal, H.; Sharma, A.K.; Pathak, M.; Muthukumar, S. Continuous Processing for Production of Biopharmaceuticals. *Prep. Biochem. Biotechnol.* **2015**, *45*, 836–849. [CrossRef]
21. Rathore, A.S.; Wood, R.; Sharma, A.; Dermawan, S. Case study and application of process analytical technology (PAT) towards bioprocessing: II. Use of ultra-performance liquid chromatography (UPLC) for making real-time pooling decisions for process chromatography. *Biotechnol. Bioeng.* **2008**, *101*, 1366–1374. [CrossRef]
22. Rajamanickam, V.; Wurm, D.; Slouka, C.; Herwig, C.; Spadiut, O. A novel toolbox for E. coli lysis monitoring. *Anal. Bioanal. Chem.* **2016**, *409*, 667–671. [CrossRef]
23. Eggenreich, B.; Rajamanickam, V.; Wurm, D.J.; Fricke, J.; Herwig, C.; Spadiut, O. A combination of HPLC and automated data analysis for monitoring the efficiency of high-pressure homogenization. *Microb. Cell Factories* **2017**, *16*, 134. [CrossRef] [PubMed]
24. Rajamanickam, V.; Krippl, M.; Herwig, C.; Spadiut, O.; Josic, D. An automated data-driven DSP development approach for glycoproteins from yeast. *Electrophor.* **2017**, *38*, 2886–2891. [CrossRef] [PubMed]
25. Peterson, R.; Nevalainen, H. Trichoderma reesei RUT-C30—Thirty years of strain improvement. *Microbiology* **2012**, *158*, 58–68. [CrossRef] [PubMed]
26. Posch, A.E.; Herwig, C. Physiological description of multivariate interdependencies between process parameters, morphology and physiology during fed-batch penicillin production. *Biotechnol. Prog.* **2014**, *30*, 689–699. [CrossRef]
27. Mandels, M.; Andreotti, R. Problems and challenges in the cellulose to cellulase fermentation. *Process Biochem.* **1978**, *13*, 6–13.
28. Lecault, V.; Patel, N.; Thibault, J. Morphological Characterization and Viability Assessment of Trichoderma reesei by Image Analysis. *Biotechnol. Prog.* **2008**, *23*, 734–740. [CrossRef]

29. Ehgartner, D.; Herwig, C.; Neutsch, L. At-line determination of spore inoculum quality in Penicillium chrysogenum bioprocesses. *Appl. Microbiol. Biotechnol.* **2016**, *100*, 5363–5373. [CrossRef]
30. Pekarsky, A.; Veiter, L.; Rajamanickam, V.; Herwig, C.; Grünwald-Gruber, C.; Altmann, F.; Spadiut, O. Production of a recombinant peroxidase in different glyco-engineered Pichia pastoris strains: A morphological and physiological comparison. *Microb. Cell Factories* **2018**, *17*, 183. [CrossRef]
31. Veiter, L.; Herwig, C. The filamentous fungus Penicillium chrysogenum analysed via flow cytometry-a fast and statistically sound insight into morphology and viability. *Appl. Microbiol. Biotechnol.* **2019**, *103*, 6725–6735. [CrossRef]
32. Rajamanickam, V.; Sagmeister, P.; Spadiut, O.; Herwig, C. *Impurity Monitoring as Novel PAT Tool for Continuous Biopharmaceutical Processes, in Repligen Yearly Reports*; Repligen: Waltham, MA, USA, 2018.
33. Tomasi, G.; Savorani, F.; Engelsen, S.B. icoshift: An effective tool for the alignment of chromatographic data. *J. Chromatogr. A* **2011**, *1218*, 7832–7840. [CrossRef]
34. Wong, J.W.H.; Durante, C.; Cartwright, H. Application of Fast Fourier Transform Cross-Correlation for the Alignment of Large Chromatographic and Spectral Datasets. *Anal. Chem.* **2005**, *77*, 5655–5661. [CrossRef] [PubMed]
35. Rajamanickam, V.; Herwig, C.; Spadiut, O. A Generic Workflow for Bioprocess Analytical Data: Screening Alignment Techniques and Analyzing their Effects on Multivariate Modeling. *Biochem. Anal. Biochem.* **2019**, *8*, 1–11.
36. Sales, K.C.; Rosa, F.; Da Cunha, B.R.; Sampaio, P.; Lopes, M.B.; Calado, C. Metabolic profiling of recombinant Escherichia coli cultivations based on high-throughput FT-MIR spectroscopic analysis. *Biotechnol. Prog.* **2016**, *33*, 285–298. [CrossRef] [PubMed]
37. Sampaio, P.; Sales, K.C.; Rosa, F.O.; Lopes, M.B.; Calado, C. High-throughput FTIR-based bioprocess analysis of recombinant cyprosin production. *J. Ind. Microbiol. Biotechnol.* **2016**, *44*, 49–61. [CrossRef]
38. Zavatti, V.; Budman, H.; Legge, R.; Tamer, M. Monitoring of an antigen manufacturing process. *Bioprocess Biosyst. Eng.* **2016**, *39*, 855–869. [CrossRef]
39. Kornecki, M.; Strube, J. Process Analytical Technology for Advanced Process Control in Biologics Manufacturing with the Aid of Macroscopic Kinetic Modeling. *Bioengineering.* **2018**, *5*, 25. [CrossRef]
40. Glassey, J. *Multivariate Data Analysis for Advancing the Interpretation of Bioprocess Measurement and Monitoring Data*; Springer Science and Business Media LLC: Berlin, Germany, 2012; Volume 132, pp. 167–191.
41. Rafferty, C.; Johnson, K.; O'Mahony, J.; Burgoyne, B.; Rea, R.; Balss, K.M. Analysis of chemometric models applied to Raman spectroscopy for monitoring key metabolites of cell culture. *Biotechnol. Prog.* **2020**, e2977. [CrossRef]
42. Chiappini, F.A.; Teglia, C.M.; Forno, Á.G.; Goicoechea, H.C. Modelling of bioprocess non-linear fluorescence data for at-line prediction of etanercept based on artificial neural networks optimized by response surface methodology. *Talanta* **2020**, *210*, 120664. [CrossRef]
43. Zimmerleiter, R.; Kager, J.; Nikzad-Langerodi, R.; Berezhinskiy, V.; Westad, F.; Herwig, C.; Brandstetter, M. Probeless non-invasive near-infrared spectroscopic bioprocess monitoring using microspectrometer technology. *Anal. Bioanal. Chem.* **2019**, *412*, 2103–2109. [CrossRef]
44. Stenlund, H.; Gorzsás, A.; Persson, P.; Sundberg, B.; Trygg, J. Orthogonal Projections to Latent Structures Discriminant Analysis Modeling on in Situ FT-IR Spectral Imaging of Liver Tissue for Identifying Sources of Variability. *Anal. Chem.* **2008**, *80*, 6898–6906. [CrossRef]
45. Narayanan, H.; Luna, M.F.; Von Stosch, M.; Bournazou, M.N.C.; Polotti, G.; Morbidelli, M.; Butté, A.; Sokolov, M. Bioprocessing in the Digital Age: The Role of Process Models. *Biotechnol. J.* **2019**, *15*, e1900172. [CrossRef] [PubMed]
46. Hemmateenejad, B.; Akhond, M.; Samari, F. A comparative study between PCR and PLS in simultaneous spectrophotometric determination of diphenylamine, aniline, and phenol: Effect of wavelength selection. *Spectrochim. Acta Part A Mol. Biomol. Spectrosc.* **2007**, *67*, 958–965. [CrossRef] [PubMed]
47. Vu, T.N.; Laukens, K. Getting Your Peaks in Line: A Review of Alignment Methods for NMR Spectral Data. *Metabolites* **2013**, *3*, 259–276. [CrossRef] [PubMed]