ARTICLE

A multisensor approach for improved protein A load phase monitoring by conductivity-based background subtraction of UV spectra

Laura Rolinger [ID] | Matthias Rüdt [ID] | Jürgen Hubbuch

Karlsruhe Institute of Technology, Karlsruhe, Germany

Correspondence
Jürgen Hubbuch, Fritz-Haber-Weg 2, 76131 Karlsruhe, Germany.
Email: juergen.hubbuch@kit.edu

Abstract

Real-time monitoring and control of protein A capture steps by process analytical technologies (PATs) promises significant economic benefits due to the improved usage of the column's binding capacity, by eliminating time-consuming off-line analytics and costly resin lifetime studies, and enabling continuous production. The PAT method proposed in this study relies on ultraviolet (UV) spectroscopy with a dynamic background subtraction based on the leveling out of the conductivity signal. This point in time can be used to collect a reference spectrum for removing the majority of spectral contributions by process-related contaminants. The removal of the background spectrum facilitates chemometric model building and model accuracy. To demonstrate the benefits of this method, five different feedstocks from our industry partner were used to mix the load material for a case study. To our knowledge, such a large design space, which covers possible variations in upstream condition besides the product concentration, has not been disclosed yet. By applying the conductivity-based background subtraction, the root mean square error of prediction (RMSEP) of the partial least squares (PLS) model improved from 0.2080 to 0.0131 g L^{-1}. Finally, the potential of the background subtraction method was further evaluated for single wavelength-based predictions to facilitate implementation in production processes. An RMSEP of 0.0890 g L^{-1} with univariate linear regression was achieved, showing that by subtraction of the background better prediction accuracy is achieved then without subtraction and a PLS model. In summary, the developed background subtraction method is versatile, enables accurate prediction results, and is easily implemented into existing chromatography setups with typically already integrated sensors.

KEYWORDS
antibody quantification, capture step, partial least squares regression, process analytical technology, protein A chromatography

1 | INTRODUCTION

The profitability of biopharmaceutical companies is decreasing (Thakor et al., 2017) due to decreasing research and development (R&D) productivity and increased drug price competition from biosimilars (Kessel, 2011). Therefore, the sector is looking to reduce costs in R&D and production by automation of the production processes (Grilo & Mantalaris, 2019; Rantanen & Khinast, 2015). The implementation of PAT is key for the digital transformation and automation of processes to gain a competitive edge over business rivals. As automation in the downstream process is economically most valuable for protein A capture steps due to the high costs of protein A resin, this area has received a lot of attention (Rüdt et al., 2017), especially in the past year (Feidl, Garbellini, Vogg, et al., et al., 2019; Thakur et al., 2019). Rüdt et al. (2017) published an approach in 2017, where ultraviolet and visible (UV/Vis) spectra were used to monitor the breakthrough of a protein A column and to control the load phase, if a certain concentration in the breakthrough was reached. While the approach itself is interesting, little explanation was given in the article on the used PLS model and what spectral changes it leverages. Additionally, a background subtraction at a constant UV signal was necessary to improve the prediction for low concentrations as the change in host cell protein (HCP) in different feeds influenced the model. This background subtraction at constant absorption is difficult, as a displacement of HCP species or highly concentrated feedstock can lead to insufficient fulfillment of UV criteria and thereby to the failure of the method.

Feidl, Garbellini, Luna et al. (2019) and Feidl, Garbellini, Vogg, et al. (2019) published an approach to monitor the breakthrough with Raman spectroscopy. Due to the low scatter efficiency of proteins, measurement times of 30 s per spectra were necessary (Feidl, Garbellini, Luna, et al., 2019; Feidl, Garbellini, Vogg, et al., 2019) and with an average of two spectra (Feidl, Garbellini, Vogg, et al., 2019), resulting in a measurement time of 1 min. Measurement times of 1 min can be insufficient for process control, especially when looking at protein A membranes with high flow rates and short load times. Even though measurement times per spectra were quite high compared to UV/Vis, additional extensive data analysis was necessary to remove high noise and make accurate predictions possible.

A limitation of current publications is furthermore the comparably small change in harvested cell culture fluid (HCCF) composition due to the usage of only one or two feedstocks in each study. Rüdt et al. (2017) used HCCF and mixed it with mock from a different cultivation. Feidl et al. used HCCF from a perfusion reactor with two different monoclonal antibody (mAb) concentration. Thakur et al. prepared flow-through and purified mAb from one batch of HCCF for a near-infrared (NIR)-based control for continuous chromatography. In all three studies, the calibration space was thus spanned by only one or two HCCF batches. Since inter-batch variations can result in a significant impact on HCP composition and DNA content (Goey, 2016), the obtained models may be limited in their predictive power for an independent HCCF batch.

To tackle sensor complexity and model validity over upstream fluctuations in this study, a product containing HCCF was mixed with three different mock materials and purified bispecific mAb. This accounts for various changes in the cell line, cell culture medium, host cell profile, and also for changes in the bispecific product profile due to the changes in the

concentration of mispaired species relative to the product. Due to the increased and random variability compared to previous studies, a prediction of the mAb concentration in the breakthrough becomes more challenging. To compensate the increased variability in the background, a novel background subtraction method was developed in this study. Specifically, a background spectrum is subtracted when the conductivity reaches a stable point. This allows to determine the breakthrough of the flow-through as the protein concentration contributes very little to the overall conductivity of the HCCF.

Finally, the usage of single wavelength absorption in combination with the conductivity-based background subtraction for product concentration prediction in the effluent is evaluated. The use of only one absorption wavelength and conductivity allows for an easy implementation of load control strategies in current manufacturing processes as those sensors are typically implemented in chromatographic equipment.

2 | MATERIALS AND METHODS

2.1 | Biologic material and buffers

All biologic material was stored at 5°C before experimentation after delivery from our industry partner. To obtain a variable mAb concentration—in this study a bispecific mAb—a variable mispaired species to product ratio, and a variable impurity profile in the load material, the product containing HCCF (Feedtsock 1) with a product concentration of $2\,g\,L^{-1}$ was mixed with purified product (Feedstock 2) and three different mock HCCFs solutions (Feedstock 3–5). One mock solution was cultivated with a nonproducing cell line. The other two mock solutions were prepared as flow-through by preparative protein A chromatography. These two mock solutions were derived from HCCFs of two different cell lines, which produce two different mAbs, respectively. Before this study, it was ensured that the protein A flow-through did not contain antibodies in detectable concentrations (based on analytical protein A chromatography). For product spiking, the used bispecific mAb (Feedstock 2) was purified to the second polishing step by our industry partner and was concentrated up to 20 $g\,L^{-1}$ to reduce dilution effects of the impurities by addition of the concentrated product.

In the product containing HCCF (Feedstock 1), different mispaired species were present, while the purified product (Feedstock 2) only contained the desired mAb. By mixing the product containing HCCF with the purified product, variation in the concentration of the different mAb species was introduced into the design space as well.

The product containing HCCF, purified mAb, and the three mock HCCFs were filtered with a cellulose acetate filter with a pore size of 0.22 μm (Pall Corporation, Port Washington, NY, USA) before mixing. In Table 1, the used volume of the different stock materials for each run are shown. The composition of the mixtures between the three mock materials was determined by Latin hypercube sampling to provide a random multidimensional distribution.

For all preparative runs, the following buffers were applied: Equilibration with 25 mM TRIS and 0.1 M sodium chloride at pH 7.4, wash with 1 M TRIS and 0.5 M potassium chloride at pH 7.4, elution

TABLE 1 Sample composition for the calibration runs 1–4 and the validation run 5 with volumes of the product containing HCCF (Feedstock 1), purified mAb (Feedstock 2), mock HCCF (Feedstock 3), and flow-through 1 and 2 (Feedstock 4 and 5)

Run number	Data usage	HCCF (ml)	mAb (ml)	Flow-through 1 (ml)	Flow-through 2 (ml)	Mock HCCF (ml)
Run 1	Calibration	52.50	0.00	9.85	21.91	20.74
Run 2	Calibration	35.00	1.75	14.82	1.36	17.08
Run 3	Calibration	21.00	3.15	6.48	3.93	7.44
Run 4	Calibration	17.50	3.50	6.57	6.13	1.30
Run 5	Validation	26.25	2.63	2.01	12.65	8.96

Abbreviations: HCCF, harvested cell culture fluid; mAb, monoclonal antibody.

with 20 mM citric acid at pH 3.6, sanitization with 50 mM sodium hydroxide and 1 M sodium chloride, and storage with 10 mM sodium phosphate, 130 mM sodium chloride, 20% ethanol.

For analytical protein A chromatography, column equilibration was carried out using a buffer with 10 mM phosphate (from sodium phosphate and potassium phosphate) with 0.65 M chloride ions (from sodium chloride and potassium chloride) at pH 7.1. Elution was performed with the same buffer, but titrated to pH 2.6 with hydrochloric acid. All buffer components were purchased from VWR. The buffers were prepared with Ultrapure Water (PURELAB Ultra, ELGA LabWater, Viola Water Technologies, Saint Maurice, France), filtrated with a cellulose acetate filter with a pore size of 0.22 μm (Pall), and degassed by sonification.

2.2 | Chromatographic instrumentation

All preparative runs were realized with an Äkta Pure 25 purification system controlled with Unicorn 6.4.1 (GE Healthcare). The system was equipped with a sample pump S9, a fraction collector F9-C, a column valve kit (V9-C, for up to 5 columns), a UV-monitor U9-M (2 mm pathlength), a conductivity monitor C9, a pH valve kit (V9-pH) and an I/O-box E9. Additionally, an UltiMate 3000 diode array detector (DAD) equipped with a semipreparative flow cell (0.4 mm optical pathlength) and operated with Chromeleon 6.8 (Thermo Fisher Scientific) was connected to the Äkta Pure. The DAD was positioned between the conductivity monitor and the V9-pH valve. Additionally, a second sensor and flow cell were positioned before the DAD. The data was not used for this study.

Reference analysis of collected fractions was performed using a Vanquish Flex Binary High-Performance Liquid Chromatography (HPLC) system (Thermo Fisher Scientific) by analytical protein A chromatography. The system consisted of a Binary Pump F, Split Sampler FT, Column Compartment H and a Diode Array Detector HL. Chromeleon Version 7.2 SR4 (Thermo Fisher Scientific) was used to control the HPLC.

2.3 | Chromatography runs

To generate variable mixtures between the product bispecific mAb, mispaired species and, other impurities for the PLS model calibration and validation, breakthrough experiments with variable mAb titers in the feed were performed. The mAb titers in the different load materials were 1, 1.5, 2, 2.5, and 3 g L^{-1}. For each experiment, a prepacked

5 × 50 mm, MabSelect SuRe column (0.982 ml; Repligen) was first equilibrated for 5 column volumes (CV) and then loaded with 100 mg of mAb. At the beginning of the load phase, the DAD equipped with a semipreparative flow cell (optical pathlength 0.4 mm) was triggered to record absorption spectra from 200 to 800 nm and the column flow-through was collected in 200 μL fractions.

2.4 | Analytical chromatography

The collected fractions of all runs were examined by analytical protein A chromatography to obtain the mAb concentrations. For each sample, a 2.1 × 30 mm POROS prepacked protein A column (Applied Biosystems) was equilibrated with 2 CV of equilibration buffer, flowed by an injection of 20 μL of sample. The column was then equilibrated with 0.8 CV of equilibration buffer and eluted with 1.4 CV of elution buffer. The flow rate was 2 mL min^{-1} for all phases and experiments.

2.5 | Data analysis

The data analysis workflow is depicted in Figure 1. The recorded 3D field, results from the analytical chromatography, and run data from the Äkta system were read in and pre-processed with MATLAB 2019R (The MathWorks, Inc.). From the conductivity data, the stable point of the conductivity was determined by smoothing the data with a moving mean filter with a window size of 5 s. If the conductivity did not change in the third decimal point for 10 s after the first CV, the conductivity was seen as stable. This point was used to subtract the background spectrum from the UV spectra, as depicted in Figure 2. The goal of this background subtraction is to remove signal originating from contaminants from the spectrum to improve product concentration predictions.

The background subtraction was performed by subtracting the measured UV spectrum closest to the stable point of the conductivity. The spectra were averaged according to the fraction size data from the Äkta. For the correlation of the averaged absorption spectra with the mAb concentrations, PLS models were calibrated using SIMCA 13.0.3 (Sartorius). SIMCA applies the Nonlinear Iterative Partial Least Squares algorithm for PLS model building (Eriksson et al., 2006). Before the PLS model calibration, all spectra and the mAb concentration were pretreated by mean-centering using SIMCA. For the calibration of the PLS model, Runs 1–4 were used as calibration data set. SIMCA applies a
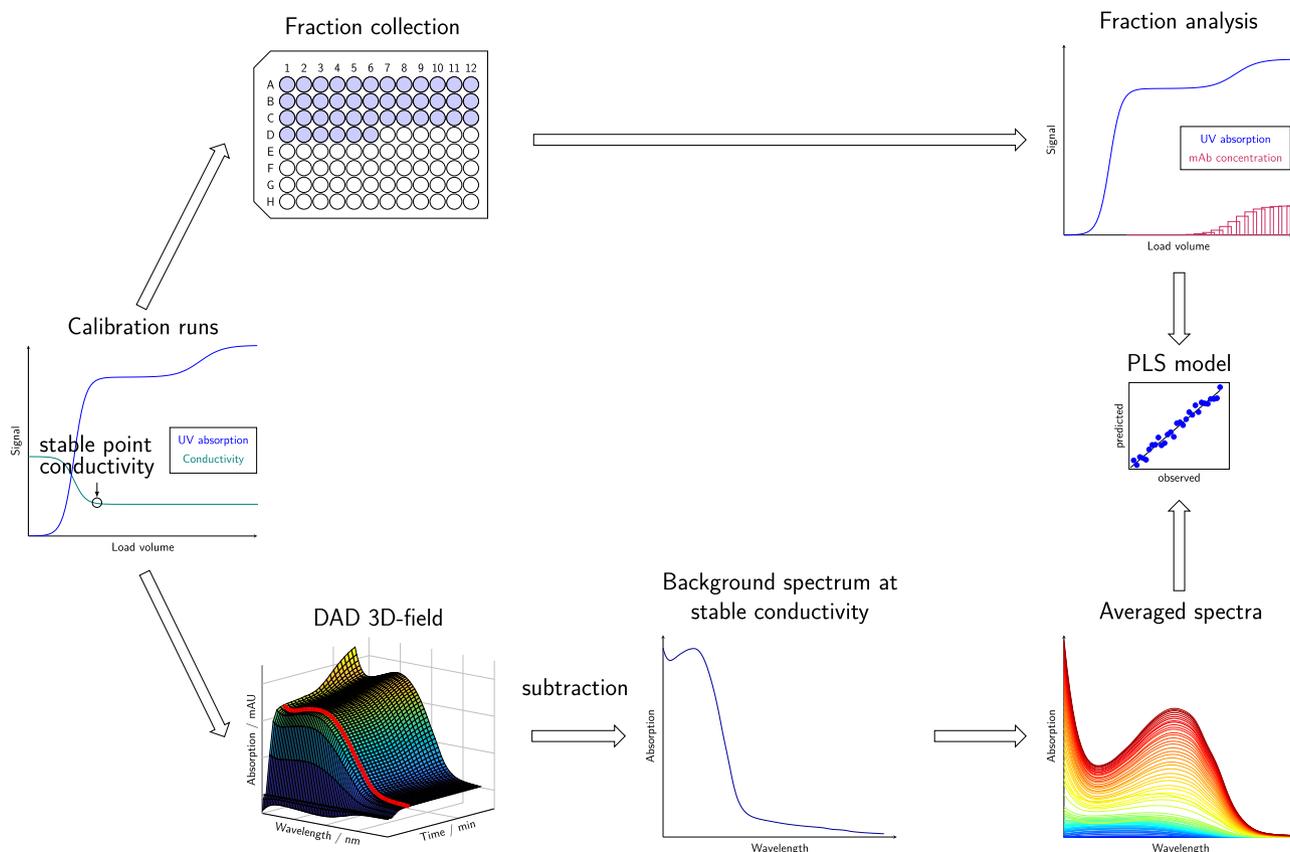
FIGURE 1 Experimental procedure for the PLS model calibration with background correction: For each calibration run, 200 µl fractions were collected and analyzed by analytical protein A chromatography to obtain the mAb breakthrough curves. During the breakthrough, 3D chromatograms and the conductivity were recorded. When the initial breakthrough of impurities was completed, determined by the stability of the conductivity signal, this background spectrum (highlighted red in 3D field) was subtracted from the 3D-field. Then the averaged spectra corresponding to the fraction size were calculated from the background-corrected absorption 3D-field. Averaged spectra and mAb concentrations were correlated using PLS modeling. 3D, three-dimensional; mAb, monoclonal antibody; PLS, partial-least square [Color figure can be viewed at wileyonlinelibrary.com]
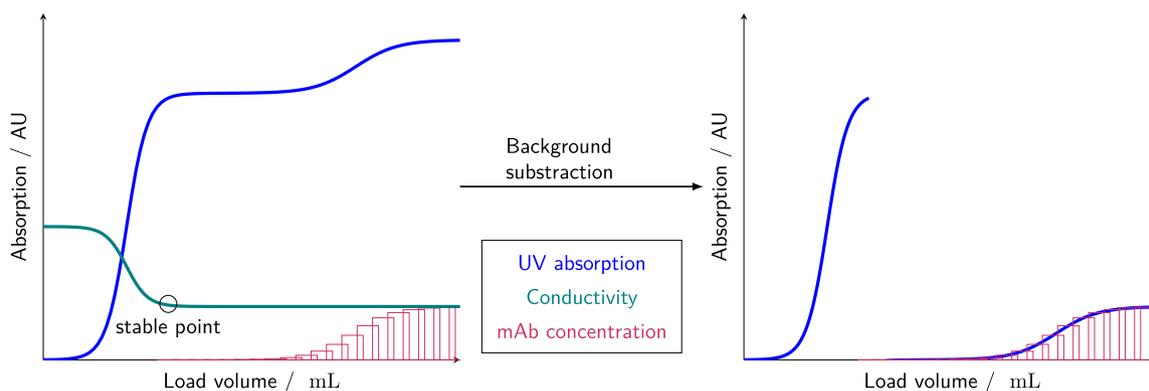

FIGURE 2 The goal of the background subtraction is to determine the complete breakthrough of the HCCF background by conductivity and to subtract the spectrum at complete background breakthrough. Through this most effects of the background are removed from the spectrum and estimation of the mAb concentration can be improved. Additionally, background effects in the HCCF due to changing conditions in the medium, HCP profile or DNA amount are excluded. HCCF, harvested cell culture fluid; HCP, host cell protein; mAb, monoclonal antibody [Color figure can be viewed at wileyonlinelibrary.com]

seven-fold cross validation as internal validation. The number of latent variable (LV) was determined by the autofit function of SIMCA. Run 5 was chosen as external validation.

The model complexity, in this case the number of LV, is important for the robustness of the model (Eriksson et al., 2006). It is important to find the right compromise between fit and predictive ability of the model. While an increase in LVs increases the fit of the model, also noise in the data can be fitted, which reduces the prediction ability of the model for new data with unknown noise or other non-idealities (Kessler, 2007).

3 | RESULTS AND DISCUSSION

In this study, the breakthrough of mAb during the protein A load phase was monitored by UV spectroscopy in combination with a PLS model. To calibrate the PLS model, four chromatographic runs (Runs 1–4) at mAb concentrations of 1, 1.5, 2.5, and $3\,g\,L^{-1}$ in the feed were performed and analyzed by off-line analytics. The actual concentration in the load material were slightly higher due to inaccuracies in the initial titer measurement of the HCCF and purified product. A validation run (Run 5) was performed at a mAb concentration of $2\,g\,L^{-1}$ in the feed. Not only was the mAb concentration varied, but also the composition of mock mixture to dilute the HCCF. This was done to imitate possible variability in upstream processing, like changes in cell culture medium, different amounts of DNA through different harvest time points, and changes in the HCP profile. This variation generates a large design space for model application.

Figure 3 compares the absorption at 280 nm A_{280} recorded by the DAD to the conductivity recorded by the Äkta. The stability criterium of the conductivity is reached between 6.6 to 10.6 ml, depending on the remaining buffer volume in the sample pump due to incomplete purging. It can be seen, that while the conductivity is stable after this point, the absorption at 280 nm is still increasing due to the displacement of impurities. It has been shown, that DNA and certain HCP species interact with the mAb bound to the Protein A resin (Aboulaich et al., 2014; Nogal et al., 2012; Sisodiya et al., 2012; Van de Velde et al., 2020). This interaction can lead to a retention effect of the interacting impurities in comparison to noninteracting impurities, which could lead to a delayed breakthrough of the interacting impurities. The difference in interaction strength between the impurities and the bound mAb could also lead to a displacement of weakly interacting contaminants by stronger interacting HCP species with progression of the load. The increase in absorption due to the displacement, while no mAb breakthrough occurs, varies between runs as the impurity profile varies. Therefore, the conductivity-based criterium is more robust for the background subtraction than a UV-based criterium.

3.1 | PLS model calibration and validation

The results of the model calibration without background subtraction are depicted in Figure 4. It compares the absorption at 280 nm A_{280} to the concentrations measured by off-line analytics

and the prediction calculated by the calibrated PLS model. It can be seen, that from the A_{280} alone, it is not possible to determine the breakthrough of mAb, because no clear plateau is visible. Likely, HCPs are displaced during loading which is overlaying with the breakthrough of the mAb (Aboulaich et al., 2014,2006; Shukla & Hinckley, 2008). The data show, that with decreasing mAb concentration and increased background variation, the off-set between the model prediction and actual concentration at low concentrations is increasing. In Table 2, the coefficient of determination R^2, the cross-validated coefficient of determination Q^2, the root mean square error of cross-validation (RMSECV) and number of LVs are compared for the model with background subtraction and without. In general, the model with background subtraction has a higher R^2 and Q^2 with 0.999 compared to 0.980, respectively, for the model without background subtraction.

The results of the model calibration with background subtraction are depicted in Figure 6. The PLS model fits the actual concentration profile over the complete concentration range better than the model without subtraction. The corrected absorption at 280 nm $A_{280corr}$ does not plateau during the load, showing that conductivity-based background subtraction is better suitable than a UV-based criterium. To further visualize the accuracy of both methods, an observed versus predicted mAb concentration plot is discussed in the Appendix A.

Additionally, the PLS model with background subtraction has two LVs compared to three LVs of the PLS model without background subtraction. In general, models with less LVs are preferred as the chances of overfitting are smaller and therefore the robustness of the model can be better.

Figure 5 compares the spectra of the uncorrected and corrected data. Apparently, the background contributes most to the UV spectrum of the load material. Typically, the background contributes between 710 mAU to 768 mAU at 280 nm to the overall absorption, while the mAb and the displaced proteins contribute between 0 mAU to 162 mAU. This indicates, that other UV active components in the load material besides the product are the main contributors to the spectrum. The local spectral maxima shift for the not background-subtracted spectra from 271 nm towards 272 or 273 nm, depending on the background spectrum. Probably different DNA concentrations in the load material cause difference in the local maximum at same concentration in different runs. DNA has a local absorption maximum around 260 nm, which could cause the spectrum of the load material to lay between 260 and 280 nm, depending on the DNA concentration. The varied concentration of UV-active components in the different load materials is probably the reason for the shift and the different total absorption values of the spectra without background subtraction by mAb concentration in Figure 5. As the mAb concentration increases the local maximum shifts in the direction of 280 nm, which is typically considered as the local maximum of proteins (Langel et al., 2009).

The local maxima in the background-corrected spectra (Figure 5b) remain constant at 279 nm as soon as the mAb concentration starts to increase. While the mAb concentration is still $0\,g\,L^{-1}$, the overall absorption still increases over time, may be due to a baseline drift by the DAD. Additionally a small contribution of the background is visible.
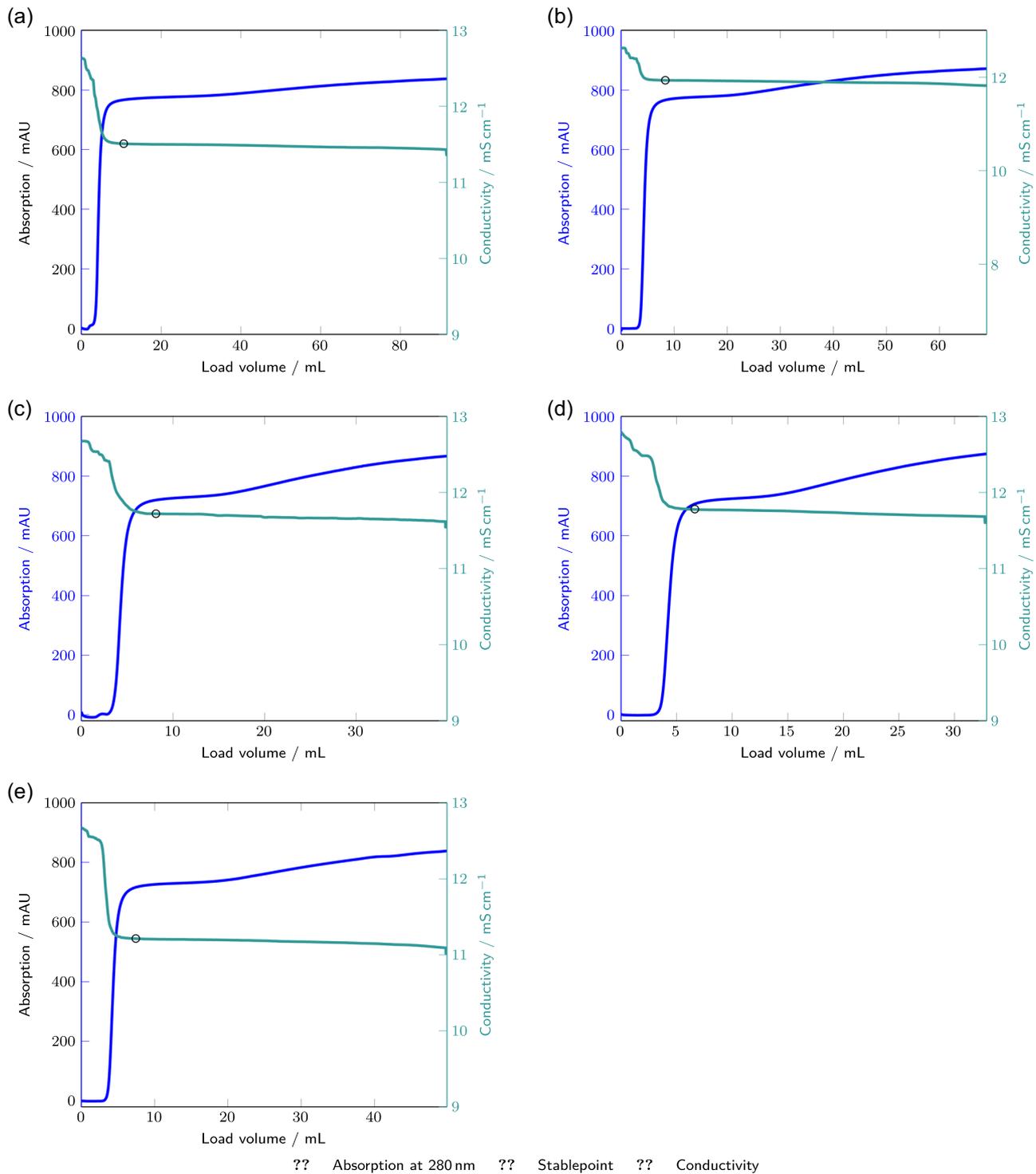
FIGURE 3 The absorption at 280 nm A_{280} recorded by the DAD (displayed as blue line) is compared with the conductivity recorded by the Äkta (teal line). The calculated stable point of the conductivity is indicated as black circle. All five runs exhibited variable mAb titers in the feed (a) 1 g L^{-1}, (b) 1.5 g L^{-1}, (c) 2 g L^{-1}, (d) 2.5 g L^{-1}, and (e) 3 g L^{-1}. DAD, diode array detector; mAb, monoclonal antibody [Color figure can be viewed at wileyonlinelibrary.com]

This may be caused by the displaced of impurities from the column due to the binding of the mAb. Both phenomena could explain, why the absorption does not stay at 0 mAU for all wavelengths, while no mAb is breaking through the column. To compare the background-corrected spectra with the absoprtion spectrum of the procduct, the absoprtion spectrum during the elution of Run 1 is plotted in black in Figure 5b. The absorption spectrum during the elution of Run 1 was normalized to maximum absorption in Figure 5 and shifted up by 5 mAU to enhance
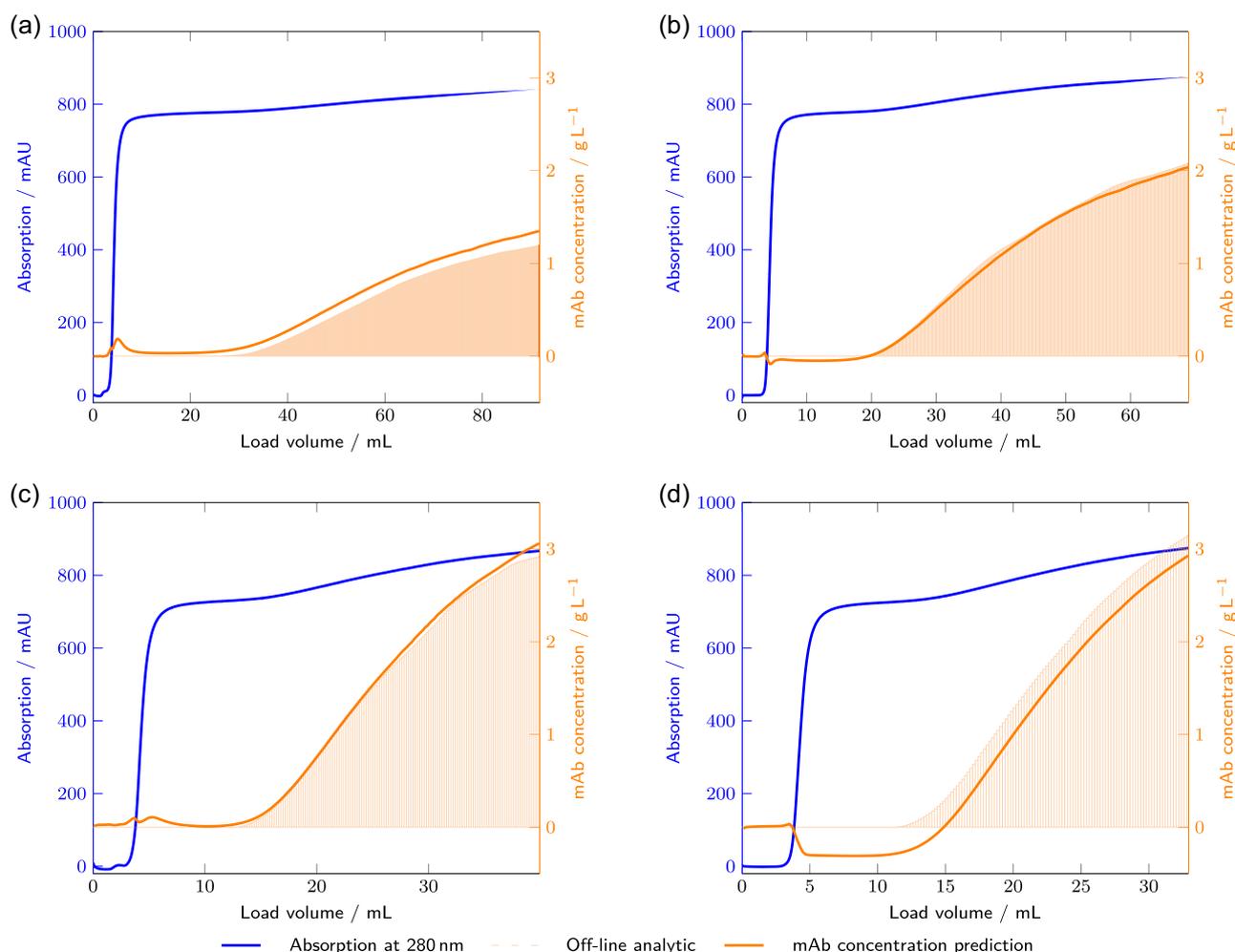
FIGURE 4 Results of the PLS model calibration without background subtraction. The absorption at 280 nm A_{280} recorded by the DAD (displayed as blue line) is compared with the results of the off-line analytics for mAb quantification (orange bars). The PLS model prediction is illustrated as orange lines. The four runs (Runs 1–4) exhibited variable mAb titers in the feed (a) $1\,g\,L^{-1}$, (b) $1.5\,g\,L^{-1}$, (c) $2.5\,g\,L^{-1}$, and (d) $3\,g\,L^{-1}$. DAD, diode array detector; PLS, partial-least square [Color figure can be viewed at wileyonlinelibrary.com]

readability. The elution spectrum has its local maximum at 279 nm like the background-corrected spectra. The absorption in the elution spectrum around the local minimum at 252 nm is lower compared to the background-corrected spectra. This could be caused by impurities contributing to the background-corrected spectra. It seems more challenging for the PLS model to extract the mAb concentration from the spectra

TABLE 2 R^2, Q^2, RMSECV, RMSEP and number of LVs for both PLS models

Background subtraction	R^2	Q^2	RMSECV in $(g\,L^{-1})$	RMSEP $(g\,L^{-1})$	Number of LVs
No	0.980	0.980	0.1170	0.2080	3
Yes	0.999	0.999	0.0246	0.0131	2

Abbreviations: LV, latent variable; PLS, partial-least square; RMSECV, root mean square error of cross-validation; RMSEP, root mean square error of prediction.
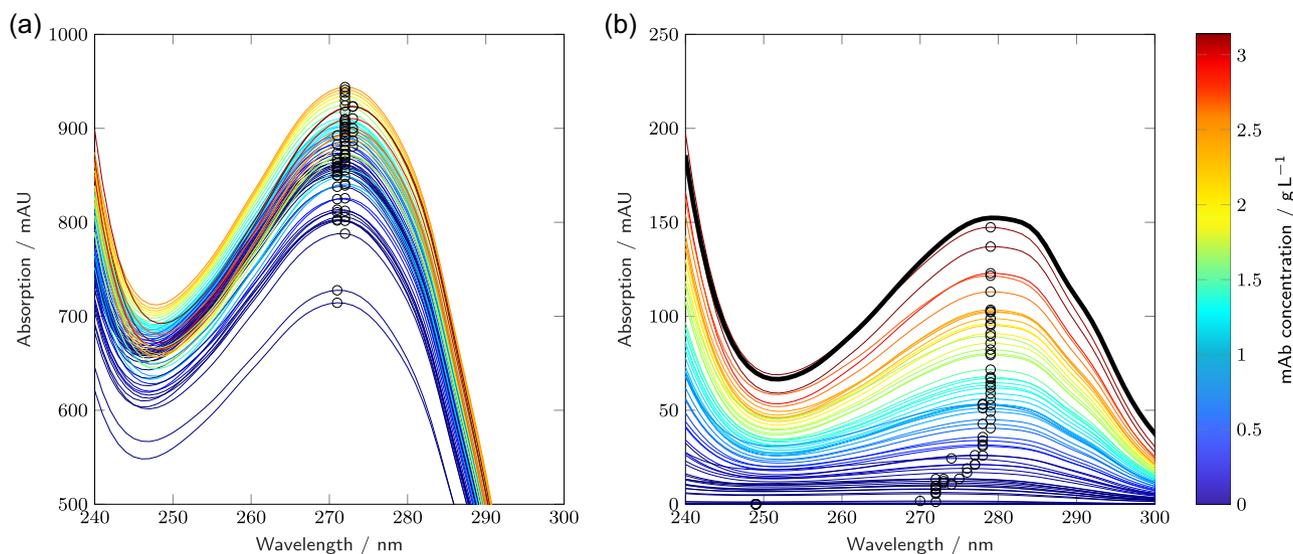
with the random variation in the background, because the PLS model without background subtraction needs more LVs to fitted the data.

The spectra with background subtraction are ordered according to mAb concentration and the local maximum stays at 279 nm, indicating that the spectrum originates from a proteinous source.

Additionally, from Figure 6 it seems, that the product concentration does not follow the absorption at 280 nm entirely, because the difference between the absorption and product concentration grows bigger with increase in product concentration. The higher the mAb concentration in the breakthrough the more HCPs seem to be displaced from the column as the column saturates. The results of the model validation without background subtraction and with background subtraction are depicted in Figure 7. For the prediction without background subtraction, an offset between the actual mAb concentration at low protein concentration persists as in the calibration data. Again, the model with background subtraction fits the actual breakthrough at low concentration better. This is also

FIGURE 5 Comparison of the averaged spectra before (a) and after background subtraction (b). Every 20th spectrum is plotted. The spectra are colored accordingly to the mAb concentration in the spectra from low concentration (blue) to high concentration (red). The local maxima of the spectra are highlighted with black circles. It is shown, that without the background subtraction, the positions of the local maximum shift with higher mAb concentration closer from 264 nm towards 270 nm. For the background-subtracted spectra, the position of the local maxima stay consistent at 279 nm after the initial breakthrough. Highlighted in black is a normalized absorption spectrum during the elution of Run 1. mAb, monoclonal antibody [Color figure can be viewed at wileyonlinelibrary.com]

represented in the RMSEP of 0.0131 g L^{-1} of the model with background subtraction compared with the RMSEP of 0.2080 g L^{-1} of the model without background subtraction (see Table 2).

Additionally, we provide the limit of detection (LOD) and the limit of quantification (LOQ) of both models with and without background subtraction in the Appendix C.

3.2 | Comparison to other publications

To set the results of this study into perspective to recent publications, the results are compared to the obtained results by (Thakur et al., 2019) for the usage of NIR spectroscopy to monitor the breakthrough and to the results by (Feidl, Garbellini, Vogg, et al., 2019) for the usage of Raman spectroscopy. As these studies were carried out on different data set and different steps for model optimization were undertaken, a final conclusion cannot be drawn by solely comparison of the results. However, a comparison can give a general steer on which method might be the most suitable for the monitoring of the Protein A load phase.

Thakur et al. (2019) published an RMSEP for the breakthrough experiments in their publication of 0.1540 g L^{-1} for NIR spectroscopy in combination with PLS models. This error is almost 10-fold higher than the RMSEP of 0.0186 g L^{-1} for the model with background subtraction from this article. As it is sometimes misleading to compare RMSEPs due to difference in involved sample concentration and sample distribution in the design space, it would be better to as well compare the goodness of fit R^2 and the goodness of prediction during cross-validation Q^2 values. However, the coefficient of determination R^2 mentioned in the paper must not be mistaken for the goodness of fit R^2, as the coefficient of

determination describes only the goodness of fit of the regression line on the observed versus predicted plot and not on the actual prediction. Another point to consider is, that no orthogonal off-line analytic was performed by Thakur et al. It remains unclear, how the actual mAb concentration was calculated, if it was not measured. Also, the chemometric side of the data analysis is not explained. This makes it difficult to evaluate, whether an effect caused by the actual difference in mAb concentration was measured or an effect due to the mixture of the feedstocks is correlated to the mAb concentration. Additionally, a more challenging design space presented in this study, as five different feedstocks were used compared to one in the case study by Thakur et al.

Therefore, it is difficult to compare the NIR-based model with the UV-based models. With the presented evidence, however, we would conclude that UV-based models seem to have a lower prediction error compared to NIR-based methods. This is also in good agreement with literature, which generally concludes, that UV absorption spectroscopy has higher accuracy due to the low impact of temperature and water background on the spectra (Kessler, 2012; Rolinger et al., 2020; Swartz, 2010).

The same can be said about Raman spectroscopy, which is also reported to have a lower accuracy and higher limit of detection in comparison to UV absorption spectroscopy for proteins (Kessler, 2012; Rolinger et al., 2020; Swartz, 2010). An average RMSEP of 0.12 g L^{-1} was published by (Feidl, Garbellini, Vogg, et al., 2019) for the breakthrough monitoring with Raman spectroscopy and PLS modeling in a concentration range, which is comparable to this study. This is again an almost 10-fold higher RMSEP as for the model with background subtraction presented in this study. Also extensive chemometric model optimization was used to achieve this RMSEP,
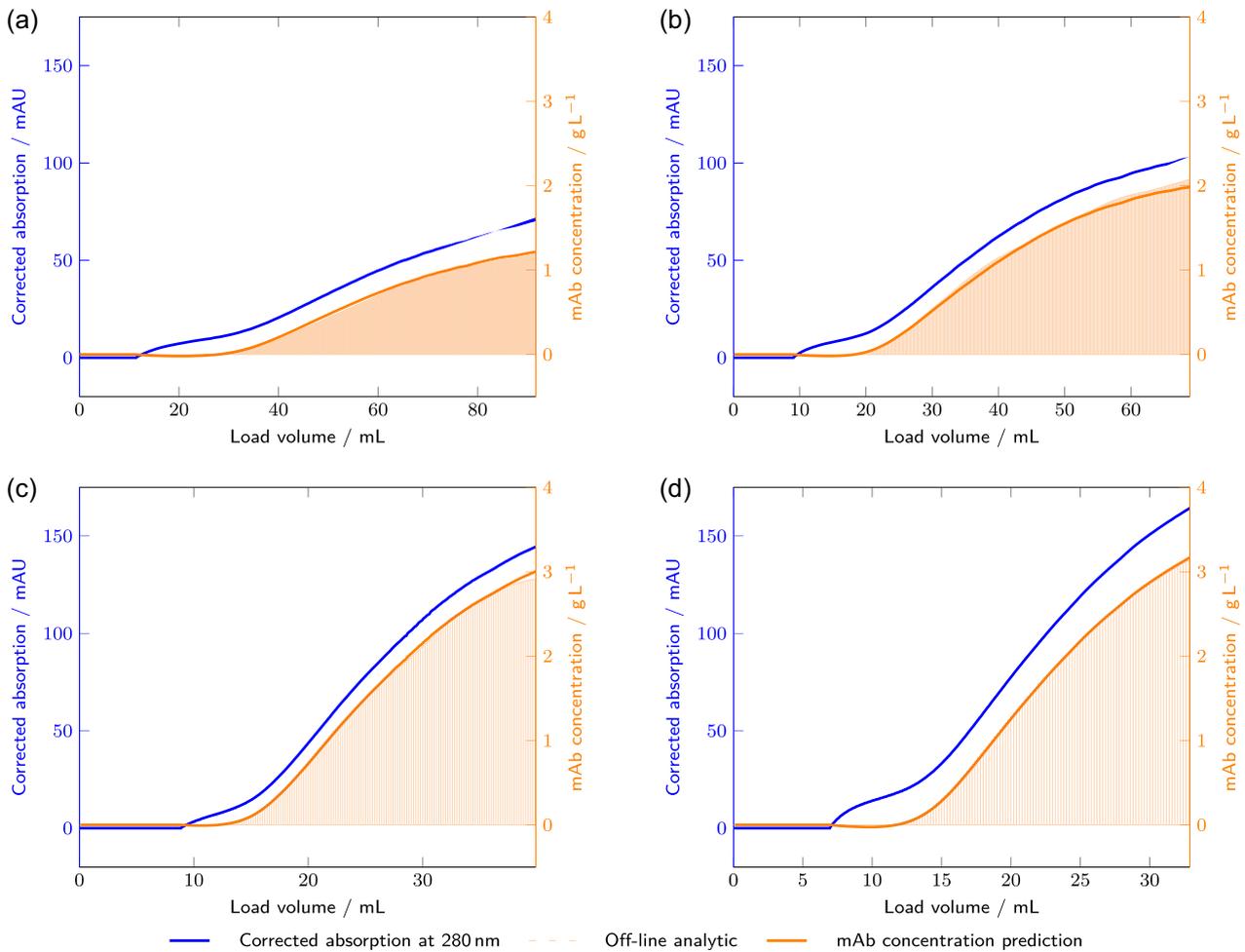
FIGURE 6 Results of the PLS model calibration with background subtraction. The absorption at 280 nm A$_{280}$ recorded by the DAD (displayed as blue line) is compared with the results of the off-line analytics for mAb quantification (orange bars). The PLS model prediction is illustrated as orange lines. The four runs (Runs 1–4) exhibited variable mAb titers in the feed (a) 1 g L^{-1}, (b) 1.5 g L^{-1}, (c) 2.5 g L^{-1}, (d) 3 g L^{-1}. DAD, diode array detector; PLS, partial-least square [Color figure can be viewed at wileyonlinelibrary.com]
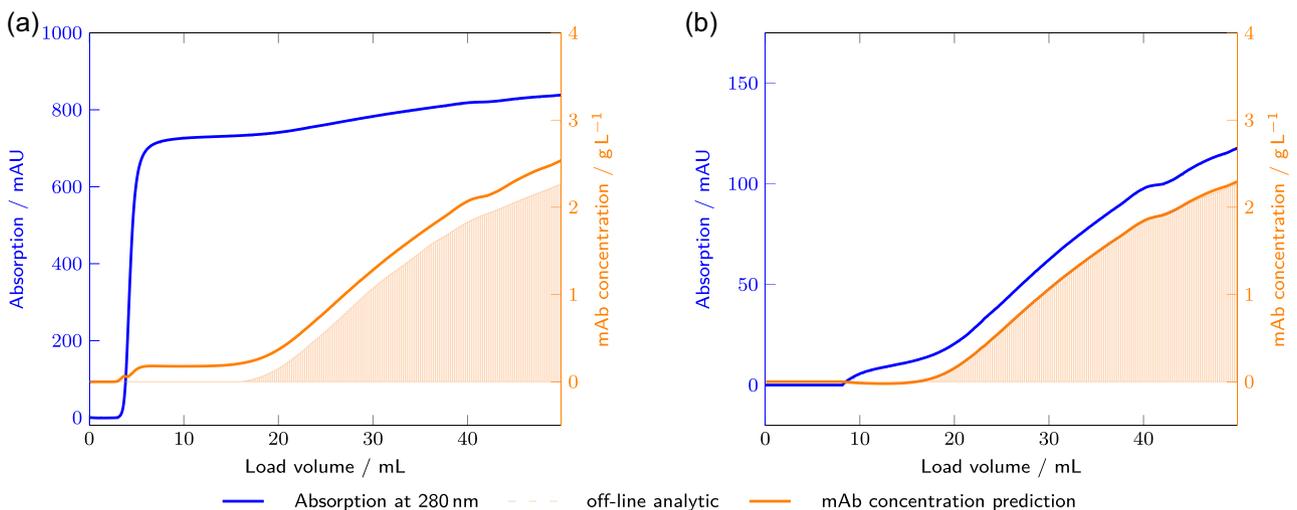

FIGURE 7 Results of the PLS model validation (Run 5). The absorption at 280 nm A$_{280}$ recorded by the DAD and displayed as blue line) is compared with the results of the off-line analytics for mAb quantification (orange bars). The PLS model prediction is illustrated as orange lines. (a) Model prediction without background subtraction and (b) the model prediction with background subtraction at a feed concentration of 2 g L^{-1}. DAD, diode array detector; PLS, partial-least square [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 R^2, Q^2, RMSECV, and RMSEP for both PLS models with only 279 nm as input

Background subtraction	R^2	Q^2	RMSECV (g L^{-1})	RMSEP (g L^{-1})
No	0.172	0.172	0.7574	0.7348
Yes	0.985	0.985	0.101	0.0890

Abbreviations: PLS, partial-least square; RMSECV, root mean square error of cross-validation; RMSEP, root mean square error of prediction.

whereas in this study no optimization was undertaken, as it was not deemed necessary. In the next study, Feidl, Garbellini, Luna, et al. (2019) investigated the usage of a lumped kinetic model and an extended Kalman filter to improve the PLS model prediction for low mAb concentrations. In the lower concentration range between 0 and 0.42 g L^{-1}, an RMSEP of 0.055 g L^{-1} was achieved. The implementation of an extended Kalman filter improved the RMSEP to 0.026 L^{-1}. Even an RMSEP of 0.026 g L^{-1} is still almost double as high as the best RMSEP of this study. Although the use of an extended Kalman leads to a prediction improvement at first, the underlying model can change during the lifetime of protein A column due to column fouling, which could make the predictions worse in the long run. Additionally, the Raman measurements were quite slow with a total measurement time of 1 min (Feidl, Garbellini, Vogg, et al., 2019) in comparison to NIR or UV measurements, which can be carried out in less than a second. An RMSEP of 0.12 g L^{-1} (Feidl, Garbellini, Vogg, et al., 2019) obtained with Raman spectroscopy and an RMSEP of 0.026 g L^{-1} (Feidl, Garbellini, Luna, et al., 2019) of Raman spectroscopy with an extended Kalman filter and extensive chemometric processing. The RMSEP in this study is still lower even though the concentration range was 10 times larger. It seems, that in general the prediction obtained by Raman spectroscopy is more corrupted by measurement noise and the use of a signal filter is obligatory to derive a more reliable prediction compared to the raw prediction.

3.3 | Application of single wavelength UV-measurements

The implementation of a DAD is not standard in most production processes. Therefore, the use of the absorption only at 280 nm was tested, with and without background subtraction. In Table 3 the R^2, the Q^2, the RMSECV and RMSEP are compared for the model with background subtraction and without. Without background subtraction, the model cannot fit the breakthrough of mAb. The R^2 and Q^2 are with 0.172 too low for spectroscopic models (Eriksson et al., 2006) and the RMSEP is with 0.7348 g L^{-1} too high for an effective control of a protein A load phase. With background subtraction, an R^2 and Q^2 of 0.985 and an RMSEP of 0.0890 g L^{-1} are achieved. This shows, that the background subtraction eliminates most effects not caused by the increase in mAb concentration. In Appendix D, a further visualization of the prediction capability of the single wavelength approach is given. The simple linear regression on a single wavelength with background subtraction allows the implementation in production processes with already available process sensor, that is, conductivity and absorption at 280 nm. No advanced chemometric methods are necessary. Instead, the approach works almost out-of-the-box. As the accuracy of the sensors are crucial for the application, low-noise sensors are required in the process.

4 | CONCLUSION AND OUTLOOK

In this study, a multisensor approach for real-time monitoring of the load phase in a protein A capture step was presented and compared to other published approaches. The proposed method relies on a dynamic UV background subtraction based on the leveling out of the conductivity signal. The background-corrected spectra can be used for product breakthrough predictions in combination with a PLS model or by single wavelength regression. In this study, a large design space with possible variations arising during fermentation was created by using five different feedstocks to mix the load material for the protein A step. The mixtures account for possible changes in contaminant profile and concentration, like buffer components, DNA, HCP, and mispaired species of the bispecific mAb. It was demonstrated that by subtracting the background spectrum during the breakthrough, the prediction of the mAb concentration is facilitated and improved compared models using the raw spectra. The proposed method offers a robust quantification of the product breakthrough regardless of large variability in the cell culture fluid.

We conclude that UV-based methods, especially with background subtraction, yield better prediction accuracies than NIR- or Raman-based methods judged by the RMSEPs published in other publications (Feidl, Garbellini, Luna, et al., 2019; Feidl, Garbellini, Vogg, et al., 2019; Thakur et al., 2019). The application of the background subtraction to product concentration determination with only one absorption wavelength shows great potential for the application to production processes as the required sensors are already implemented in most processes.

CONFLICT OF INTERESTS
The authors declare that there are no conflict of interests.

AUTHOR CONTRIBUTIONS
L. Rolinger designed the study, carried out the experiments, analyzed the data and wrote the manuscript. M. Rüdt contributed to the study design and supported the data analysis and writing of the manuscript. J. Hubbuch supervised the project and reviewed the manuscript.

ORCID
Laura Rolinger http://orcid.org/0000-0002-6061-654X
Matthias Rüdt http://orcid.org/0000-0002-8583-6645

REFERENCES

Aboulaich, N., Chung, W. K., Thompson, J. H., Larkin, C., Robbins, D., & Zhu, M. (2014). A novel approach to monitor clearance of host cell proteins associated with monoclonal antibodies. *Biotechnology Progress*, 30(5), 1114–1124. https://doi.org/10.1002/btpr.1948

Allegrini, F., & Olivieri, A. C. (2014). IUPAC-consistent approach to the limit of detection in partial least-squares calibration. *Analytical Chemistry*, 86(15), 7858–7866.

Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., & Wold, S. (2006). *Multi- and megavariate data analysis* (Vol. 1, p. 95). Umetrics Academy.

Feidl, F., Garbellini, S., Luna, M. F., Vogg, S., Souquet, J., Broly, H., Morbidelli, M., & Butté, A. (2019). Combining mechanistic modeling and Raman spectroscopy for monitoring antibody chromatographic purification. *Processes*, 7(10), 683.

Feidl, F., Garbellini, S., Vogg, S., Sokolov, M., Souquet, J., Broly, H., Butté, A., & Morbidelli, M. (2019). A new flow cell and chemometric protocol for implementing in-line Raman spectroscopy in chromatography. *Biotechnology Progress*, 35(5), e2847.

Goey, C. H. (2016). Cascading effects in bioprocessing: The impact of cell culture environment on CHO cell behaviour and host cell protein species. Thesis.

Grilo, A. L., & Mantalaris, A. (2019). The increasingly human and profitable monoclonal antibody market. *Trends in Biotechnology*, 37(1), 9–16.

Kessel, M. (2011). The problems with today's pharmaceutical business-Ťan outsider's view. *Nature Biotechnology*, 29(1), 27–33.

Kessler, R. W. (2012). *Prozessanalytik: Strategien und Fallbeispiele aus der industriellen Praxis*. John Wiley & Sons.

Kessler, W. (2007). *Multivariate datenanalyse: Für die pharma, bio-und prozessanalytik*. John Wiley & Sons.

Langel, U., Cravatt, B. F., Graslund, A., VonHeijne, N., Zorko, M., Land, T., & Niessen, S. (2009). *Introduction to peptides and proteins*. CRC Press.

Nogal, B., Chhiba, K., & Emery, J. C. (2012). Select host cell proteins coelute with monoclonal antibodies in protein A chromatography. *Biotechnology Progress*, 28(2), 454–458.

Rantanen, J., & Khinast, J. (2015). The future of pharmaceutical manufacturing sciences. *Journal of Pharmaceutical Sciences*, 104(11), 3612–3638.

Rolinger, L., Rüdt, M., & Hubbuch, J. (2020). A critical review of recent trends, and a future perspective of optical spectroscopy as PAT in biopharmaceutical downstream processing. *Analytical and Bioanalytical Chemistry*, 412, 2047–2064.

Rüdt, M., Brestrich, N., Rolinger, L., & Hubbuch, J. (2017). Real-time monitoring and control of the load phase of a protein A capture step. *Biotechnology and Bioengineering*, 114(2), 368– 373.

Shukla, A. A., & Hinckley, P. (2008). Host cell protein clearance during protein A chromatography: Development of an improved column wash step. *Biotechnology Progress*, 24(5), 1115–1121. https://doi.org/10.1002/btpr.50

Sisodiya, V. N., Lequieu, J., Rodriguez, M., McDonald, P., & Lazzareschi, K. P. (2012). Studying host cell protein interactions with monoclonal antibodies using high throughput protein A chromatography. *Biotechnology Journal*, 7(10), 1233–1241.

Swartz, M. (2010). HPLC detectors: A brief review. *Journal of Liquid Chromatography & Related Technologies*, 33(9–12), 1130–1150.

Thakor, R. T., Anaya, N., Zhang, Y., Vilanilam, C., Siah, K. W., Wong, C. H., & Lo, A. W. (2017). Just how good an investment is the biopharmaceutical sector? *Nature Biotechnology*, 35 (12), 1149.

Thakur, G., Hebbi, V., & Rathore, A. S. (2019). An NIR-based PAT approach for real-time control of loading in protein A chromatography in continuous manufacturing of monoclonal antibodies. *Biotechnology and Bioengineering*, 117(3), 673–686.

Van de Velde, J., Saller, M. J., Eyer, K., & Voloshin, A. (2020). Chromatographic clarification overcomes chromatin-mediated hitch-hiking interactions on Protein A capture column. *Biotechnology and Bioengineering*, 117(11), 3413–3421.

APPENDIX A: LINEARITY OVER CONCENTRATION RANGE FOR THE PLS MODELS

Figure A1 shows the predicted mAb concentration over the observed/measured mAb concentration for the PLS models with and without background subtraction. Predicted and observed mAb concentrations
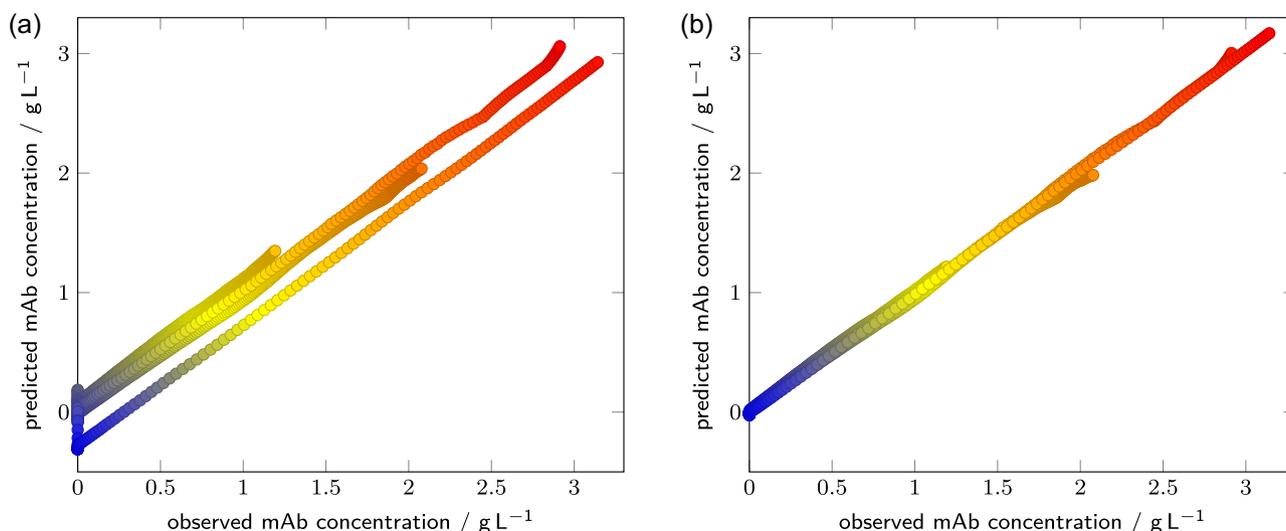
FIGURE A1 Predicted mAb concentration by the PLS model over the measured (observed) mAb concentration for (a) the not background-corrected PLS model and (b) the background-corrected PLS model. mAb, monoclonal antibody; PLS, partial-least square

show a linear relationship for both models. Deviations from the linear relationship could be possibly caused by the off-line analytic due to carry-over between samples. However the not background-corrected models show different offsets depending on the individual run. The largest offset can be observed for Run 1. These offset could originate from the differences in load material. The PLS model with background subtraction shows no significant offsets, which seems to be a result of the removal of different spectral contributions from the different feedstock material.

APPENDIX B: BACKGROUND COMPOSITION

All feedstocks used in this study could be differentiated by the color of the HCCF. Figure B1 shows the different background
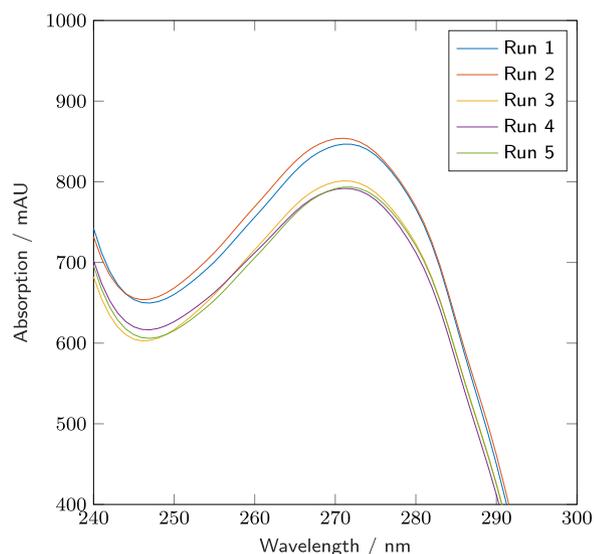
spectra, which were subtracted. As contaminants, like DNA, HCP, some buffer componants and scattering molecules contribute to this background spectrum, the diversity in the feedstock can be spectrally assessed. Interestingly, the background spectra cluster into 2 groups. Even though Runs 1 and 2 have a very different composition, the background spectra look similar to Run 2 possibly having a higher DNA concentration due to the increased absorption at 260 nm, but not at 280 nm. Also Runs 3–5 show similar background spectra with regard to the total amount of absorption, but also differ in the composition possibly due to different DNA, HCP, and amount of large molecules, which cause light scattering.

APPENDIX C: LIMIT OF DETECTION

The LOD and LOQ interval for the data set was calculated based on the MATLAB code provided by Allegrini and Olivieri (2014). The results are displayed in Table C1. If the background subtraction is done, both the LOD and LOQ intervals are lower in comparison to without background subtraction. Additionally are the intervals itself smaller with background subtraction. The reduced spectral contribution of interfering components due to the background subtraction could explain these findings, allowing for better detection and quantification.

FIGURE B1 Comparison of the background spectra for the calibration runs (Runs 1–4) and the validation run (Run 5)

TABLE C1 LOD interval and LOQ interval for multivariate models with and without background subtraction

Background subtraction	LOD interval ($g L^{-1}$)	LOQ interval ($g L^{-1}$)
Yes	0.0130–0.0144	0.039–0.043
No	0.0752–0.0940	0.226–0.282

Abbreviations: LOD, limit of detection; LOQ, limit of quantification.

APPENDIX D: SINGLE WAVELENGTH PREDICTION

Figure D1 shows the predicted mAb concentration over the observed/measured mAb concentration for the linear regression models with and without background subtraction at 279 nm. The regression without background subtraction shows large offsets for the different runs, which seem to be driven by the contribution of the background spectra (see Figure B1). The regression model with background subtraction shows little offsets. Only Run 4 seems to have a larger offset compared to the other runs, which could be explained by a comparable little earlier subtraction of the background as with the other runs. Interestingly, even though the offsets are minimized by the background subtraction, the predicted mAb concentration over the observed/measured mAb concentration show different slopes for the different runs. This could be caused by the different interacting species present in the load material, which are displaced at a different rate from the column between the different runs.
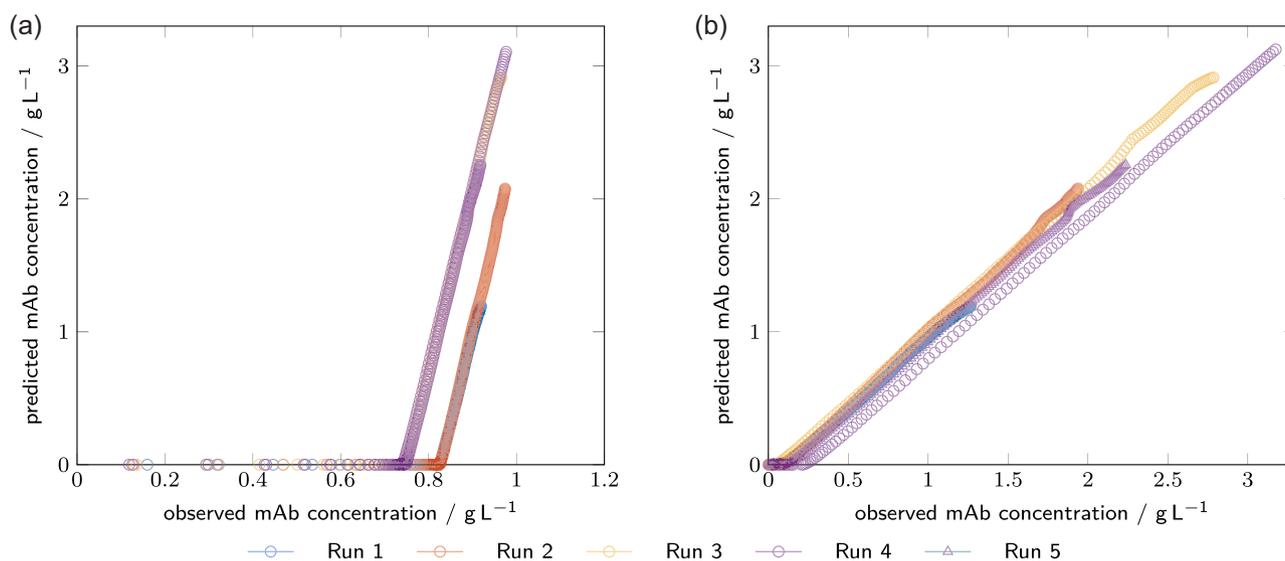

FIGURE D1 Predicted mAb concentration by the PLS model over the measured (observed) mAb concentration for (a) the not background-corrected PLS model and (b) the background-corrected linear regression model. mAb, monoclonal antibody; PLS, partial-least square