Chemical Engineering Science 176 (2018) 157–164

Contents lists available at ScienceDirect

Chemical Engineering Science

journal homepage: www.elsevier.com/locate/ces

Selective protein quantification for preparative chromatography using variable pathlength UV/Vis spectroscopy and partial least squares regression

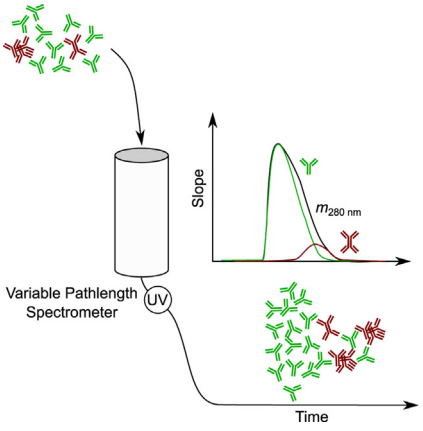
Nina Brestrich [1], Matthias Rüdt [1], Daniel Büchler, Jürgen Hubbuch *

Institute of Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology, Engler-Bunte Ring 3, 76131 Karlsruhe, Germany

HIGHLIGHTS

- PAT tool for monitoring of preparative chromatographic steps.
- In-line UV/Vis spectral measurement at protein concentrations up to 75 g/l.
- Selective protein quantification by partial least squares regression (PLS).
- Control of purification steps by calculating pool purities with PLS models.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 31 March 2017
Received in revised form 8 October 2017
Accepted 22 October 2017
Available online 24 October 2017

Keywords:
Preparative chromatography
Process Analytical Technology
Partial least squares regression
In-line monitoring
Variable pathlength spectroscopy
Selective protein quantification

ABSTRACT

In preparative protein chromatography, broad dynamic ranges of protein concentrations as well as co-elution of product and impurities are common. Despite being the standard in biopharmaceutical production, monitoring of preparative chromatography is generally limited to surrogate signals, e.g. UV absorbance at 280 nm. To address this problem, variable pathlength (VP) spectroscopy in conjunction with Partial Least Squares regression (PLS) was used to monitor preparative chromatography. While VP spectroscopy enabled the acquisition of absorbance data for a broad concentration range, PLS modelling allowed for the differentiation between the protein species. The approach was first implemented for monitoring the separation of lysozyme from cytochrome c at an overall loading density of 92 g/l. The same method was then applied to the polishing step of a monoclonal antibody (mAb) at 40 g/l loading density. For PLS model prediction of the mAb monomer and the high molecular weight variants (HMWs), the root mean square error (RMSE) was 1.07 g/l and 0.42 g/l respectively. To demonstrate the usability of the approach for in-line control, pooling decisions for both separation problems were subsequently taken based on the computed concentrations or thereof derived purities. In summary, VP spectroscopy in conjunction with PLS modelling is a promising option for in-line monitoring and control of future chromatography steps at large scale.

* Corresponding author.
E-mail address: juergen.hubbuch@kit.edu (J. Hubbuch).
[1] Contributed equally.

https://doi.org/10.1016/j.ces.2017.10.030
0009-2509/

1. Introduction

In current purification processes of biopharmaceuticals, preparative liquid chromatography is key for separating the target product from media components, DNA, host cell proteins, and product related impurities (Carta and Jungbauer, 2010). The method is used because of its high separative power while minimizing product loss. Despite being the standard, monitoring of preparative chromatography is generally limited to surrogate signals, e.g. UV absorbance at 280 nm. The Process Analytical Technology (PAT) initiative of the US Food and Drug Adminstration (FDA) however promotes the acquisition of critical quality attributes in (near) real-time (PAT, 2004). Especially for chromatography, PAT is an active field of research (Rathore et al., 2010; Rathore and Kapoor, 2015).

Preparative chromatographic processes are generally run at high loading densities to realize efficient processes (Carta and Jungbauer, 2010). High loading densities subsequently lead to a broad range of protein concentrations eluting from the column. Thus, monitoring techniques must feature a broad dynamic range to capture the peak concentration as well as lower concentrated impurities in the peak flanks. Furthermore, due to the limited resolution of preparative chromatography, baseline separation between product and impurities is rarely achieved. PAT techniques should therefore selectively quantify the product and the main impurities. Due to the rapid changes in chromatography, the typical decision time within chromatographic processes is limited (Rüdt et al., 2017a; Rathore et al., 2010).

In literature, a number of different approaches have been proposed for (near) real-time monitoring of chromatography. Online monitoring and control of preparative chromatography has been realized by automated sampling and subsequent analysis by high performance liquid chromatography (Fahrner et al., 1998; Kaltenbrunner et al., 2012; Rathore et al., 2010). The broad concentration ranges in preparative liquid chromatography can be managed easily by varying injection volumes. Disadvantages comprise time delay due to sample handling and assay times as well as the risk of contamination.

Fourier Transform Infrared (FTIR) spectroscopy has been applied for at-line monitoring of downstream processes (Capito et al., 2015). Samples were taken at various stages of downstream processing, dried and subsequently measured by FTIR. The method allowed to quantify product, high molecular weight variants (HMWs), and host cell proteins over a broad range of concentrations. Despite showing promising results, at-line measurements bear the risk of introducing operator errors and may increase the risk of contamination. Furthermore, the delay due to sample handling and measuring times may be too long for the typical decision times in chromatography.

Previously, UV/Vis spectroscopy has been proposed as a method for in-line monitoring of chromatographic processes (Brestrich et al., 2014; Brestrich et al., 2015; Rüdt et al., 2017b). Its usefulness was shown for multicomponent mixtures of model proteins and for real-life separation problems. Based on UV/Vis spectral data and Partial Least Squares regression (PLS) modelling, process decisions were taken such as the beginning and end of product pooling. Other applications include high throughput process development (Hansen et al., 2011; Baumann et al., 2016) and coupling with chromatography modelling (Brestrich et al., 2016). While being very fast and accurate, previous applications of in-line monitoring using UV/Vis spectroscopy only took process decisions for separation problems in diluted conditions. The problem of broad concentration ranges occurring in preparative liquid chromatography was not addressed.

To increase the dynamic range of spectroscopic acquisitions, measurement cells have been designed which allow to continuously change the optical pathlength to achieve ideal sensitivity for virtually any analyte concentration (Doucen et al., 1985; Flowers and Callender, 1996). This approach in conjunction with PLS modelling was later successfully applied for monitoring the chemical oxygen demand of wastewater (Chen et al., 2014). The methodology has also been further developed by a commercial vendor for protein related applications. With the commercialized product, spectra of proteins were acquired at a broad range of concentrations (Thakkar et al., 2012). Samples were studied for spectral effects of protein-protein interactions in UV/Vis with protein concentrations up to 250 g/l. Recently, an additional product line has been launched for in-line variable pathlength (VP) measurements. The device is able to provide measurement results up to absorbances of approximately 80 AU/mm corresponding to a mAb concentration of almost 600 g/l.

In this publication, we demonstrate in-line monitoring of preparative chromatography runs by UV/Vis VP spectroscopy in conjunction with PLS modelling. While VP spectroscopy allowed to monitor chromatographic processes at almost arbitrary protein concentrations, PLS models selectively quantified multiple co-eluting species from spectral data. The approach was first implemented for the separation of a mixture of the model proteins lysozyme (lys) and cytochrome c (cyt c) at high loading densities. Former was considered the product, latter the contaminant with high respectively low concentration in the feed. Subsequently, a preparative polishing step of mAb monomer and its HMWs was monitored. To demonstrate the usefulness of this approach for process control, the pooling of products in both separation problems was controlled in-line based on either the predicted concentrations of eluting proteins for the model system or on the calculated purity for the mAb purification.

2. Materials and methods

In both studies, cation exchange chromatography runs with different gradient lengths were executed for PLS model calibration and confirmation. Thereby, different mixing ratios and concentrations of proteins were obtained in order to span a calibration space for the PLS models.

2.1. Chromatography instrumentation

All preparative chromatography experiments were performed on an ÄKTA Pure 25 system 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, and an I/O-box E9. The system was controlled with Unicorn 6.4.1. (all GE Healthcare, Chalfont St Giles, UK). The column effluent was monitored using a FlowVPE VP UV/Vis spectrometer (C Technologies, Bridgewater, USA). It was integrated into the flow path of the ÄKTA system between the conductivity monitor and fraction collector.

The reference analytics of collected fractions was performed using a Dionex UltiMate 3000 rapid separation liquid chromatography system (Thermo Fisher Scientific, Waltham, USA). The system was composed of a HPG-3400RS pump, a WPS-3000 analytical autosampler, a TCC-3000RS column thermostat, and a DAD-3000RS detector.

2.2. VP spectroscopy

The FlowVPE VP spectrometer uses a mobile optical fiber to change the pathlength L of the flow cell (cf. Fig. 1). A VP measurement cycle consists of a screening phase and a subsequent acquisition phase. During the screening phase, optimal pathlengths for the acquisition phase are estimated based on 3 measurements at

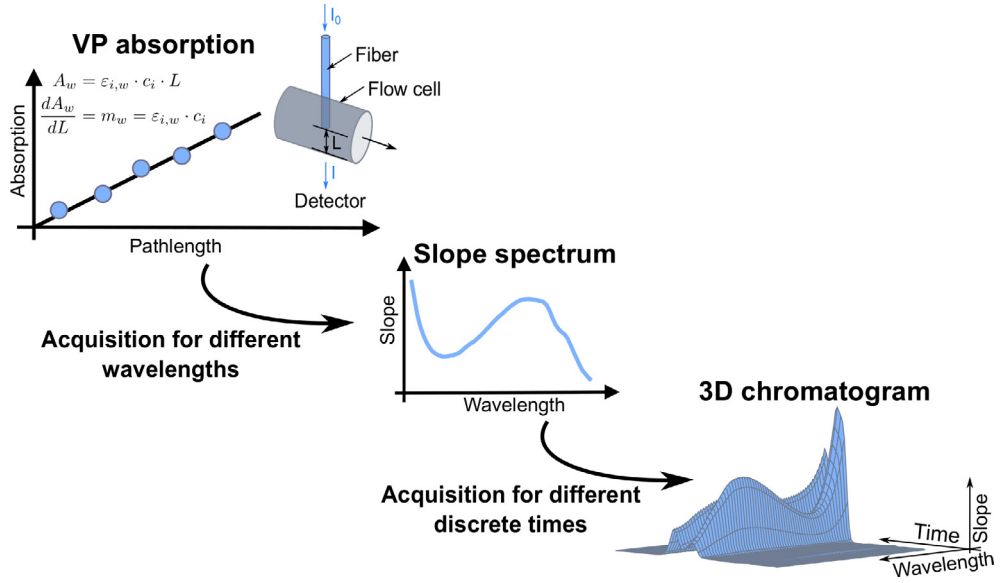


Fig. 1. VP spectroscopy in chromatography: Each measurement cycle, variable pathlengths L are set by actuating the optical fiber of the FlowVPE. After the screening phase for the identification of the optimal measurement pathlengths, UV absorbance spectra at five different pathlengths are acquired. The slope m_w is calculated for each wavelength giving a slope spectrum. For a whole chromatography run, this results in a 3-dimensional chromatogram in (time, wavelength, slope).

$L_s \in \{5, 25, 100\}$ µm at 280 nm. During the acquisition phase, absorbances $A_{w,i}$ are measured at five different pathlengths $i \in \{1, \ldots, 5\}$ for each wavelength of interest w. Based on the Lambert-Beer law, a linear correlation is assumed between the absorbance estimate $\widetilde{A}_w$ and the pathlength L (Eq. (1)). By solving the least squares problem in Eq. (2), slope m_w and intercept b_w of the regression are estimated.

$$\widetilde{A}_w(L; b_w, m_w) = b_w + m_w L \tag{1}$$

$$\min_{b_w, m_w} \sum_i \left(A_{w,i} - \widetilde{A}_w(L_i; b_w, m_w) \right)^2 \tag{2}$$

As slope spectra are obtained during a whole chromatography run, the result of an experiment is a point set in (time, wavelength, slope), thus a 3-dimensional chromatogram. By equating Eq. (1) and the Lambert-Beer law and taking the derivative with respect to L, Eq. (3) is obtained.

$$m_w = \boldsymbol{C} \boldsymbol{\varepsilon}_w^T \tag{3}$$

This shows, that the obtained slope spectra are correlated with the protein concentrations $\boldsymbol{C}$ over the absorbance coefficients $\boldsymbol{\varepsilon}_w$. Since only the slopes are of interest, the fiber offset has no impact on the final result. Generally, for mid-UV spectra measurement cycles take approximately 30 s.

2.3. Case study I: separation of Cyt c from Lys

2.3.1. Proteins and buffers

As model system, a protein mixture consisting of lys from hen egg white and cyt c from equine heart was used (both Sigma-Aldrich, St. Louis, USA). For the preparative runs, the equilibration buffer was 20 mM sodium phosphate (pH 7). Elution was carried out with 20 mM sodium phosphate and 500 mM sodium chloride (pH 7). For the in-line pooling decision, the sodium chloride concentration was increased to 550 mM to simulate a process disturbance. For reference analytics (analytical cation exchange chromatography), equilibration was performed with 20 mM Tris (pH 8) and elution was carried out with 20 mM Tris and 700 mM sodium chloride (pH 8). All buffer components were purchased from VWR, West Chester, USA. 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, Port Washington, USA), and degassed by sonification before usage.

2.3.2. Chromatography runs

A HiTrap 16 × 25 mm column prepacked with SP Sepharose FF (GE Healthcare) was first equilibrated for 5 column volumes (CVs) and then loaded with 418 mg lys and 41.8 mg cyt c. After a wash of 1 CV with equilibration buffer, the proteins were eluted by performing a linear gradient from 0 to 100% elution buffer. The gradient length was 2, 4, 6, and 8 CVs. While the results of the runs with gradient lengths 2, 4, and 8 CVs were applied for the PLS model calibration, the run with a gradient length of 6 CVs was used to confirm the model. At the beginning of each linear gradient, data acquisition of the FlowVPE was started and slope spectra from 240 nm to 300 nm with 2 nm resolution were obtained. After the linear gradient elution, the column was regenerated for 3 CVs with the elution buffer. The flow rate was kept constant at 0.5 ml/min for all steps and experiments. During the elution and regeneration, 1000 µl fractions were collected in deep well plates (VWR).

2.3.3. Reference analytics

The collected fractions were analyzed by analytical cation exchange chromatography. For each injection, the Proswift SCX-1S 4.6 × 50 mm column (Thermo Fisher Scientific) was first equilibrated for 2 min, loaded with sample, washed for 0.5 min with equilibration buffer, and eluted with a linear salt gradient from 10 to 100% elution buffer in 2 min. The column was subsequently regenerated with 100% elution buffer for 1 min. The flow rate was constant at 1.5 mL/min for all steps.

2.4. Case study II: separation of HMWs from mAb monomer

2.4.1. Proteins and buffers

mAb Protein A pool was obtained from Lek Pharmaceuticals d.d. (Mengeš, Slovenia) and stored at −80 °C before experimentation. Because of the low, not detectable HMW content of the mAb Protein A pool, it was partly low pH stressed to reach an overall

HMW level of 10%. For the preparative runs, an equilibration buffer consisting of 20 mM sodium citrate (pH 6) was used. Elution was performed with 20 mM sodium citrate and 500 mM sodium chloride (pH 6). For the in-line pooling decision, the sodium chloride concentration was increased to 550 mM to simulate a process disturbance. For reference analytics (analytical size exclusion chromatography), a buffer with 200 mM potassium phosphate and 250 mM potassium chloride (pH 7) was used. All buffer components were purchased from VWR. All buffers were prepared with Ultrapure Water. Prior to the experiments, the buffers and the feed were filtrated with a cellulose acetate filter with a pore size of 0.22 µm (Pall). The buffers were also degassed by sonification before usage.

2.4.2. Chromatography runs

The HiTrap 16 × 25 mm SP Sepharose FF column was first equilibrated for 5 CVs and then loaded with 200 mg mAb (monomer and HMWs). After a wash of 3 CVs with equilibration buffer, variable linear gradients from 0 to 100% elution buffer were performed. Gradient length was set to 4, 5, 6, and 7 CVs. The results of the runs with gradient lengths 4, 6, and 7 CVs were used to calibrate the PLS model, while the run with a gradient length 5 CVs was applied to confirm the model. As for the model protein study, data acquisition of the FlowVPE was started at the beginning of the gradient and slope spectra from 240 nm to 340 nm with 2 nm resolution were recorded. After the linear gradient elution, the column was regenerated for 3 CVs with the elution buffer. The flow rate was kept constant at 0.5 ml/min for all steps and experiments. During the elution and regeneration, 1000 µl fractions were collected in deep well plates (VWR).

2.4.3. Reference analytics

As reference analytics, size exclusion chromatography was performed with the collected fractions to determine the concentration of mAb monomer and HMWs. Samples were injected into a 4.6 × 150 mm TSKgel SuperSW mAb HTP column (Tosoh, Tokio, Japan). The flow rate was set to 0.3 ml/min.

2.5. Data analysis

Protein slope spectra were extracted from 3-dimensional chromatograms while, according to off-line analytics, pure protein eluted. Each slope spectrum was normalized by dividing by its average slope.

As the total duration of a measurement cycle of the FlowVPE (including the screening for the linear range) varied slightly, the results of the off-line reference analytics were linearly interpolated such that they matched the slope spectra. The slope spectra were then preprocessed by mean centering and correlated with the results of the off-line analytics using PLS1 regression (Eriksson et al., 2006; Höskuldsson, 1988; Martens and Næs, 1989). PLS1 was used as little correlation was observed between the measured concentrations of the different species (Wold et al., 2001). The number of latent variables in the corresponding PLS model was based on the minimization of the root mean square error (RMSE) of the model prediction in a cross validation.

For both case studies, the concentrations of the different species were subsequently smoothed over time by a Savitzky-Golay filter in Matlab® (MathWorks, Inc., Natrick, USA) (Savitzky and Golay, 1964). This filtering approach was chosen as the PLS model does not consider time-wise correlation of the data. The Savitzky-Golay filter allows to smooth data over time. For the calibration and validation runs, the Savitzky-Golay filter was used in a symmetric form smoothing the central data point in a given frame (frame size: 11 points, 2nd order polynomial). The Savitzky-Golay filter was not applied for the in-line pooling decision to avoid a time delay. For calculating the mass balances for purity and yield off-line, the concentrations were again smoothed as described above. Smoothing in spectral dimension was not applied as information crucial to the PLS may be lost in the smoothing process. Furthermore, PLS models already suppress noise in the spectral dimension by mapping spectral data onto latent variables (Wold et al., 2001).

2.6. In-line monitoring and control

The pooling of the products in both case studies was controlled in-line based on the predicted concentrations of the PLS model. A .NET assembly provided by the vendor of the FlowVPE triggered a call-back function in Matlab after each spectral measurement. Matlab subsequently computed the slope spectra and the protein concentrations. The communication between Matlab and Unicorn was implemented analogous to the protocol published earlier (Brestrich et al., 2014). Shortly, Matlab triggers a block in a Unicorn method by sending a digital signal when a predefined pooling criterion is fulfilled.

To demonstrate that the method can handle process disturbances, modified elution buffers with increased salt concentrations (550 mM sodium chloride instead of 500 mM sodium chloride in both case studies) were used. For the second case study, the loading density was additionally decreased to 150 mg as it may occur in production when the remainder of a batch is loaded on a column. In the first case study, lys was declared as product. Lys was automatically collected, as soon as its concentration exceeded 2 g/l and the cyt c concentration fell below 1.8 g/l. Pooling was continued as long as the lys concentration remained above 2 g/l. Pooling in the second case study was triggered when the mAb monomer concentration exceeded 2 g/l. Pooling was stopped as soon as the purity of the pool fell below 95%. For calculating the mass balances, negative concentrations computed by the PLS model were set to 0. The termination of pooling was only allowed after the eluate was collected for three minutes to prevent a fractionation stop due to noise in the prediction.

3. Results and discussion

As described above, linear gradients with variable lengths were performed in both case studies to obtain different mixing ratios and concentrations of the examined proteins. The acquired slope spectra and the results of the fraction analysis of three runs were used to calibrate a PLS model, i.e. to span a calibration space for the model. The fourth run was subsequently applied to test the PLS model in both studies. Before the results of both studies are discussed in detail, typical chromatograms obtained by the combination of chromatography and VP spectroscopy will be presented.

3.1. Application of VP spectroscopy for chromatography

In VP spectroscopy, absorbance is measured at different pathlengths. Fig. 2A illustrates a typical chromatogram obtained by VP spectroscopy at 280 nm. The green lines represent the measured absorbance over time at different pathlengths. The corresponding pathlengths are illustrated by the grey lines. In orange, the slope over time is depicted. As shown by Eq. (3), the slope is linearly dependent on the total protein concentration. It is worth noting that the absorbance values stay almost constant during protein elution, while the pathlengths decrease. During the screening phase, the longest pathlength for the measurement phase is selected such that the absorbance is expected to be 1 AU at 280

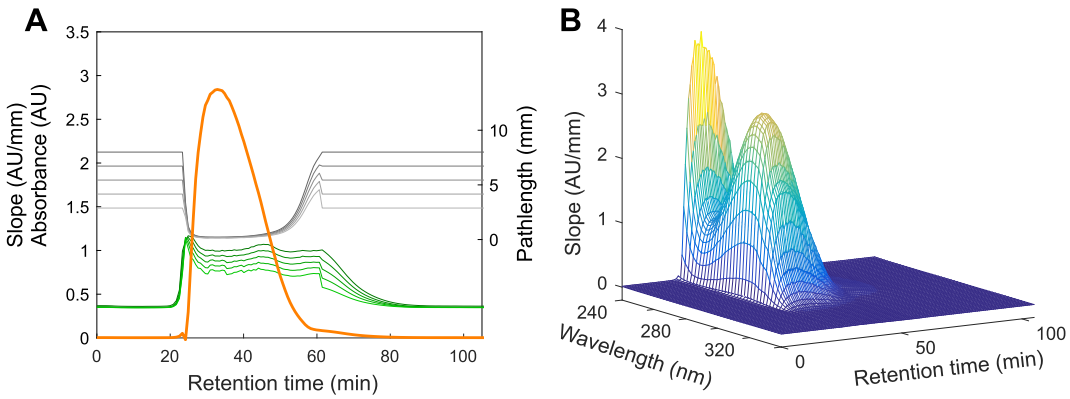


Fig. 2. Typical chromatograms obtained by VP spectroscopy for one wavelength (A) and multiple wavelengths (B). A: The green lines represent the obtained absorbance values at the different pathlengths (grey lines). At each point in time, the slope (orange line) was determined by a linear regression between five absorbance values and pathlengths. B: If slope spectra are recorded during a chromatography run, a 3-dimensional chromatogram is obtained. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

nm. In consequence, the set pathlengths are inversely proportional to the slope and thus to the protein concentration (cf. Fig. 2A).

While Fig. 2A displays a typical chromatogram for one wavelength, a chromatogram for multiple wavelengths is presented in Fig. 2B. Instead of single wavelength slopes, slope spectra were acquired. As the slopes were recorded during a whole chromatography run, this resulted in a 3-dimensional chromatogram in (time, wavelength, slope). The obtained 3-dimensional chromatograms of the chromatography experiments were the starting point for the PLS model calibration and confirmation.

3.2. Case study I: separation of Cyt c from Lys

Cyt c and lys feature significant spectral differences in UV due to the different mass fraction of aromatic amino acids as well as the heme group in cyt c (cf. Fig. 3A). The two proteins thus allow to test the proposed setup for a simple application.

Based on the performed cross validation, PLS1 models with 2, respectively 3 latent variables were selected for lys and cyt c. The resulting PLS model predictions for the three calibration runs (gradient length of 2, 4, and 8 CVs) are displayed in Fig. 4A, B, and D. The plots compare the PLS model prediction for lys (solid blue lines) and cyt c (solid red lines) with the results of the corresponding reference analytics (blue bars for lys and red bars for cyt c). A good agreement was observed between PLS model prediction and reference for all calibration runs.

To confirm the model, predictions for a 6 CV gradient run were made, which closely follow the corresponding reference analytics (cf. Fig. 4C). The different shades of blue and red of the PLS model prediction in Fig. 4 illustrate the unsmoothed and smoothed data (lighter colors for the unsmoothed and darker colors for the smoothed data). The RMSE of this run was 0.53 g/l for cyt c and 1.11 g/l for lys. The RMSE of the smoothed prediction data with a Savitzky-Golay filter was 0.48 g/l for cyt c and 1.05 g/l for lys. Applying the Savitzky-Golay filter clearly improved the RMSE for both the prediction of lys and cyt c (cf. Table 1). The results show that the RMSE of the model predictions is partly due to random noise but also partly systematic. As Savitzky-Golay filtering helps to reduce random noise, the method was selected as a suitable post-run data treatment.

The combination of VP spectroscopy and PLS modelling allowed for a selective quantification of lys and cyt c during chromatography runs with highly loaded columns and lys peak concentrations of up to 80 g/l. This shows that the method is applicable for the typical concentration range of preparative chromatography.

3.3. Case study II: separation of HMWs from mAb monomer

In contrast to cyt c and lys, mAb monomer and HMW presumably contain the same mass fraction of aromatic amino acids and disulfide bridges. Spectral differences are therefore either related to changes in tertiary structure or due to light scattering (Jiskoot and Crommelin, 2005). As a result, the differences in the spectra of mAb monomer and HMW are comparably small (cf. Fig. 3B). This makes the quantification by PLS modelling more challenging.

Based on the performed cross validation, a PLS1 model with 4 latent variables was selected for the mAb monomer, while a PLS1 model with 8 latent variables was used for the HMWs. Fig. 5A, C, and D display the PLS model prediction for the three calibration runs (gradient length of 4, 6, and 7 CVs), while Fig. 5B displays the results of the confirmation run (5 CV gradient). The figures show a comparison between the PLS model prediction for the mAb monomer (solid blue lines) and the HMWs (solid red lines) with the results of the reference analytics (blue bars for mAb monomer and red bars for HMWs). In all four runs, the model prediction matched the reference analytics closely. The results thus demonstrate the applicability of the method for proteins with only slight differences in the absorption spectra. The RMSE of prediction was 1.26 g/l for the mAb monomer and 0.50 g/l for HMWs. The RMSE could be further reduced by smoothing the data with a Savitzky-Golay filter. This led to RMSE values of 1.07 g/l for mAb monomer and 0.42 g/l for HMWs. Fig. 4 compares the unsmoothed and the smoothed data for all runs. Applying the Savitzky-Golay filter improved the RMSE for both the prediction of the mAb monomer and HMWs indicating the presence of random noise (cf. Table 1).

3.4. In-line pooling decisions

To show the usefulness of VP spectroscopy in conjunction with PLS modelling, the implemented methods were used for process decisions. For both case studies, the pooling decision for chromatographic runs with process disturbances, i.e. away from the set point, were performed. For the first case study, lys was purified from cyt c in a 6 CV gradient with an increased salt concentration in the elution buffer (550 mM sodium chloride). The start of the pooling was triggered by the PLS model as soon as the concentration of lys exceeded 2 g/L and the cyt c concentration fell bellow 1.8 g/l (cf. Fig. 6A). For the second case study, a 5 CV gradient was performed also with an increased salt concentration in the elution buffer (550 mM sodium chloride). Furthermore, the loading

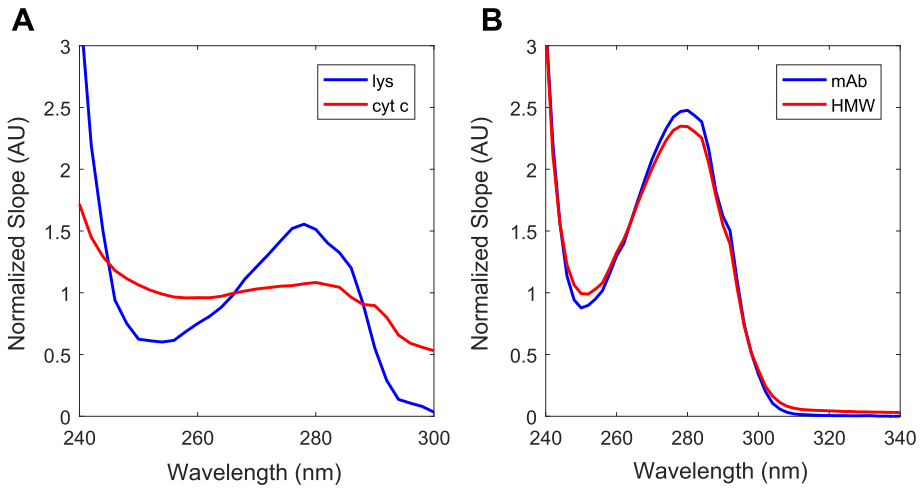


Fig. 3. Comparison of normalized protein slope spectra for case study I (A) and case study II (B).

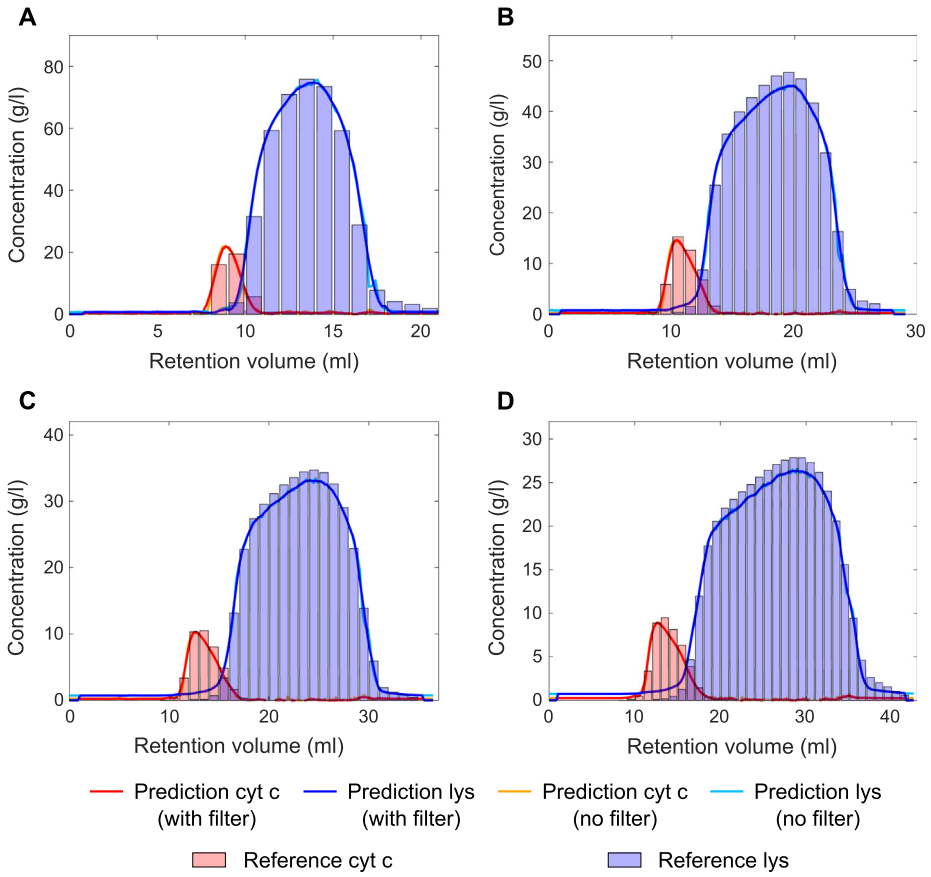


Fig. 4. Comparison of the PLS model prediction for lys and cyt c with the results of the off-line reference analytics for a gradient length of (A) 2 CV, (B) 4 CV, (C) 6 CV (confirmation run), and (D) 8 CV. The different shades of blue and red of the PLS model prediction illustrate the effects of smoothing. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

density of the column was decreased from 40 g/l to 30 g/l. Such a situation may occur in production processes, when the pool from the previous process step cannot be transferred to an integer number of column cycles. Thus, a number of cycles are performed with fully loaded columns. The remainder is then loaded onto a column resulting in a lower loading density, which might change the separation and pool purity. This problem can be solved by applying the described in-line control. The method was applied to trigger pooling as soon as the mAb monomer concentration exceeded 2 g/l. Pooling was stopped, when the purity fell bellow 95% (cf. Fig. 6B).

The collected pools were subsequently analyzed by off-line analytics. In both case studies, accurate predictions could be made with minor deviations. For the separation of cyt c and lys, a purity of 99.0% was predicted while off-line analytics measured a 99.7% purity. For the mAb monomer product pool, the PLS model predicted a purity of 94.4 % compared to 94.2 % measured by off-line analytics. Here, Savitzky-Golay filtering did not lead to an

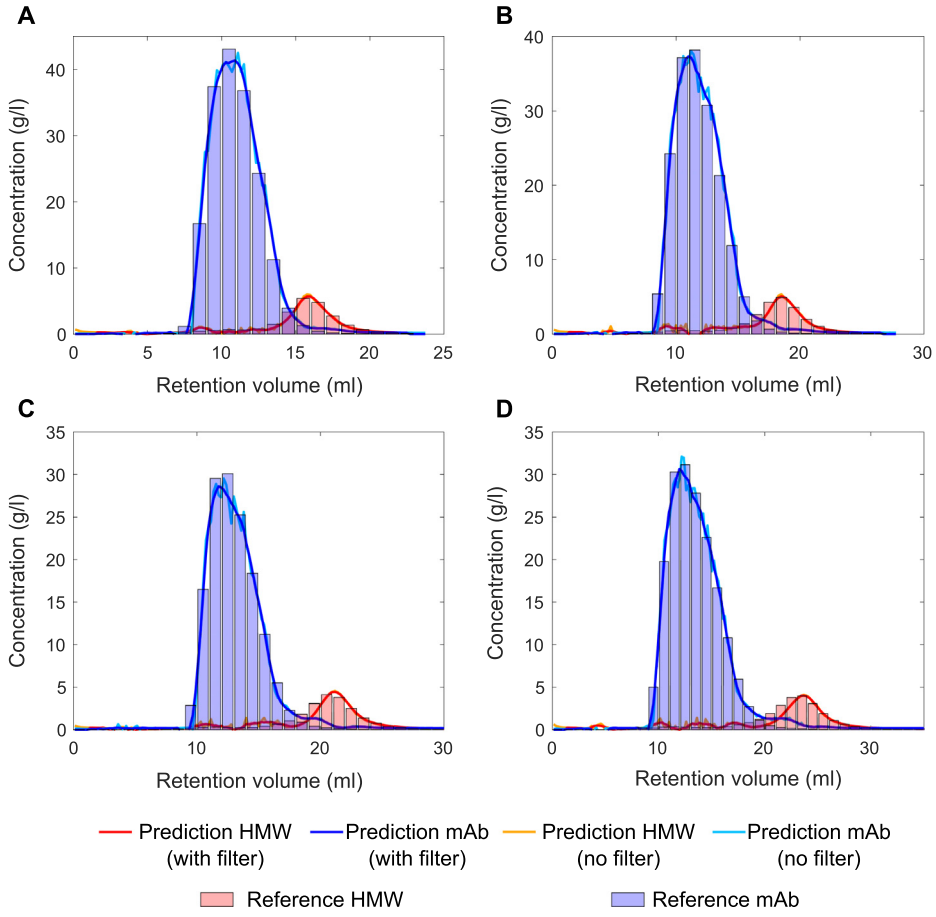


Fig. 5. Comparison of the PLS model prediction for mAb monomer and HMWs with the results of the off-line reference analytics for a gradient length of (A) 4 CV, (B) 5 CV (confirmation run), (C) 6 CV, and (D) 7 CV. The different shades of blue and red of the PLS model prediction illustrate the effects of smoothing. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 1
RMSE for the prediction in the confirmation runs of the two case studies: The RMSE of unfiltered predictions is compared with the RMSE of the filtered data.

	Case study I		Case study II	
	lys [g/l]	cyt c [g/l]	mAb monomer [g/l]	HMW [g/l]
No filter	1.10	0.53	1.26	0.50
Savitzky-Golay filter	1.04	0.48	1.07	0.42

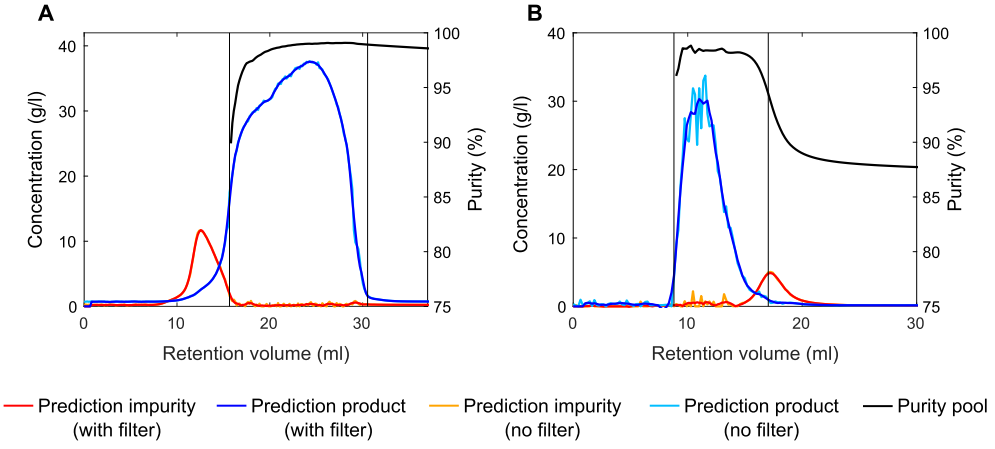


Fig. 6. Prediction of the PLS model during the in-line control run for case study I (A) and case study II (B). The different shades of blue and red of the PLS model prediction illustrate the effect of smoothing. Based on the unsmoothed data, the pool purity was calculated (black). The thin black vertical lines indicate the start and end of pooling.

Table 2
Results of the in-line pooling decisions: For both case studies, the pool purity was calculated based on the PLS model prediction and compared with the corresponding results from off-line analytics as reference.

	Case study I purity [%]	Case study II purity [%]
Reference	99.7	94.2
Prediction (no filter)	99.0	94.4
Prediction (with filter)	99.0	94.4

improved prediction because the measurement noise cancelled itself out over time (cf. Table 2).

The results show that the proposed method may be applied for in-line control of a chromatography system. The used residence time of 10 min is above the industrial standard. However, a reduction of the flow rate during elution may be justifiable as the elution phase is relatively short compared to the complete process. It is worth noting that the discussed approach is significantly faster than on-line HPLC PAT methods proposed in literature. In production scale, on-line HPLC was applied with 2.5 times to 3 times slower response time (1.3 min resp. 96 s) (Rathore et al., 2010; Kaltenbrunner et al., 2012). Thus, the applicability of VP spectroscopy should be further evaluated on a case-to-case basis.

4. Conclusion and outlook

In-line monitoring of preparative chromatography was successfully realized in this study. It was demonstrated, that the combination of VP UV/Vis spectroscopy and PLS modelling allows for a selective in-line protein quantification in a broad dynamic range of concentrations. The method enabled the monitoring of chromatography runs with highly loaded columns. Product peak concentrations varied between 30 g/l to 80 g/l, while contaminant peak concentrations were only 4 g/l to 20 g/l. Consequently, the proposed method has potential for the in-line monitoring and control of preparative chromatography. It might also be applicable for controlling switching times in continuous chromatography. Future challenges are especially related to the scale up and robustness of the method as well as to the optimization of the measurement time of VP spectroscopy. As each VP measurement relies on accurate mechanical positioning of the optical fiber, an increase in measurement speed is challenging. Additionally, further investigations have to be made to improve understanding and sensitivity for HMW detection.

Acknowledgment

This project has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No. 635557. We kindly thank Lek Pharmaceuticals d. d. (Mengeš, Slovenia) for providing the mAb Protein A pool.

References

Carta, G., Jungbauer, A., 2010. Protein Chromatography: Process Development and Scale-up. John Wiley & Sons.
Guidance for Industry. PAT – A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance, 2004.
Rathore, A.S., Bhambure, R., Ghare, V., 2010. Process analytical technology (PAT) for biopharmaceutical products. Anal. Bioanal. Chem. 398 (1), 137–154. https://doi.org/10.1007/s00216-010-3781-x.
Rathore, A.S., Kapoor, G., 2015. Application of process analytical technology for downstream purification of biotherapeutics. J. Chem. Technol. Biotechnol. 90 (2), 228–236.
Rüdt, M., Briskot, T., Hubbuch, J., 2017a. Advances in downstream processing of biologics – spectroscopy: an emerging process analytical technology. J. Chromatogr. A 1490, 2–9. https://doi.org/10.1016/j.chroma.2016.11.010.
Fahrner, R.L., Lester, P.M., Blank, G.S., Reifsnyder, D.H., 1998. Real-time control of purified product collection during chromatography of recombinant human insulin-like growth factor-I using an on-line assay. J. Chromatogr. A 827 (1), 37–43.
Kaltenbrunner, O., Lu, Y., Sharma, A., Lawson, K., Tressel, T., 2012. Risk-benefit evaluation of on-line high-performance liquid chromatography analysis for pooling decisions in large-scale chromatography. J. Chromatogr. A 1241, 37–45.
Rathore, A.S., Parr, L., Dermawan, S., Lawson, K., Lu, Y., 2010. Large scale demonstration of a process analytical technology application in bioprocessing: use of on-line high performance liquid chromatography for making real time pooling decisions for process chromatography. Biotechnol. Prog. 26 (2), 448–457.
Capito, F., Zimmer, A., Skudas, R., 2015. Mid-infrared spectroscopy-based analysis of mammalian cell culture parameters. Biotechnol. Prog. 31 (2), 578–584.
Brestrich, N., Briskot, T., Osberghaus, A., Hubbuch, J., 2014. A tool for selective inline quantification of co-eluting proteins in chromatography using spectral analysis and partial least squares regression. Biotechnol. Bioeng. 111 (7), 1365–1373.
Brestrich, N., Sanden, A., Kraft, A., McCann, K., Bertolini, J., Hubbuch, J., 2015. Advances in inline quantification of co-eluting proteins in chromatography: process-data-based model calibration and application towards real-life separation issues. Biotechnol. Bioeng. 112, 1406–1416.
Rüdt, M., Brestrich, N., Rolinger, L., Hubbuch, J., 2017b. Real-time monitoring and control of the load phase of a protein a capture step. Biotechnol. Bioeng. 114 (2), 368–373. https://doi.org/10.1002/bit.26078.
Hansen, S.K., Skibsted, E., Staby, A., Hubbuch, J., 2011. A label-free methodology for selective protein quantification by means of absorption measurements. Biotechnol. Bioeng. 108 (11), 2661–2669.
Baumann, P., Huuk, T., Hahn, T., Osberghaus, A., Hubbuch, J., 2016. Deconvolution of high-throughput multicomponent isotherms using multivariate data analysis of protein spectra. Eng. Life Sci. 16 (2), 194–201. https://doi.org/10.1002/elsc.201400243.
Brestrich, N., Hahn, T., Hubbuch, J., 2016. Application of spectral deconvolution and inverse mechanistic modelling as a tool for root cause investigation in protein chromatography. J. Chromatogr. A 1437, 158–167. https://doi.org/10.1016/j.chroma.2016.02.011 http://www.sciencedirect.com/science/article/pii/S0021967316300978.
Doucen, R.L., Houdeau, J.P., Cousin, C., Menoux, V., 1985. Variable path-length, low-temperature cells for absorption spectroscopy. J. Phys. E: Scient. Instrum. 18 (3), 199 http://stacks.iop.org/0022-3735/18/i=3/a=007.
Flowers, P., Callender, S., 1996. Variable path length transmittance cell for ultraviolet, visible, and infrared spectroscopy and spectroelectrochemistry. Anal. Chem. 68 (1), 199–202. https://doi.org/10.1021/ac950580w http://www.ncbi.nlm.nih.gov/pubmed/21619236.
Chen, B., Wu, H., Li, S.F.Y., 2014. Development of variable pathlength UV-vis spectroscopy combined with partial-least-squares regression for wastewater chemical oxygen demand (COD) monitoring. Talanta 120, 325–330. https://doi.org/10.1016/j.talanta.2013.12.026.
Thakkar, S.V., Allegre, K.M., Joshi, S.B., Volkin, D.B., Middaugh, C.R., 2012. An application of ultraviolet spectroscopy to study interactions in proteins solutions at high concentrations. J. Pharm. Sci. 101, 3051–3061.
Eriksson, L., Johansson, E., Kettaneh-Wold, N., Trygg, J., Wikström, C., Wold, S., 2006. Multi- and Megavariate Data Analysis. Umetrics Academy.
Höskuldsson, A., 1988. PLS regression methods. J. Chemom. 2, 211–228.
Martens, H., Næs, T., 1989. Multivariate Calibration. Wiley.
Wold, S., Sjöström, M., Eriksson, L., 2001. PLS-regression: a basic tool of chemometrics. Chemometr. Intell. Lab. Syst. 58 (2), 109–130. https://doi.org/10.1016/S0169-7439(01)00155-1.
Savitzky, A., Golay, M.J.E., 1964. Smoothing and differentiation of data by simplified least squares procedures. Anal. Chem. 36 (8), 1627–1639.
Jiskoot, W., Crommelin, D., 2005. Methods for Structural Analysis of Protein Pharmaceuticals. AAPS Press.