

# BioCapt® Single-Use

Aktiver Luftkeimsammler

*Testfertiger, aktiver Ersatz für Sedimentationsplatten*



- Kontinuierliche Luftprobenahme und Überwachung
- Testfertiger Ersatz für Sedimentationsplatten
- Weniger Handhabungsschritte und weniger falsch-positive Untersuchungen
- Für den mobilen oder remote Einsatz
- Auswahl verschiedener Agarformulierungen und Flussraten
- Erfüllt ISO 14698-1 und den erwarteten EU-GMP-Annex 1 Einweg-Anforderungen
- Verwendung mit aktiven Luftkeimsammler-Instrumenten für eine vollständig validierte Lösung

Mehr erfahren



**PARTICLE  
MEASURING  
SYSTEMS®**  
a spectris company


**Kontaktieren Sie uns für mehr Informationen:**

[www.pmeasuring.de](http://www.pmeasuring.de)

T: +49 351 8896 3850

E: [pmsgermany@pmeasuring.com](mailto:pmsgermany@pmeasuring.com)

# Automated calibration and in-line measurement of product quality during therapeutic monoclonal antibody purification using Raman spectroscopy

Jiarui Wang<sup>1</sup>  | Jingyi Chen<sup>1,2</sup> | Joey Studts<sup>1</sup> | Gang Wang<sup>1</sup>

<sup>1</sup>Late Stage Downstream Process Development, Boehringer Ingelheim Pharma GmbH/Co. KG, Biberach an der Riss, Germany

<sup>2</sup>Karlsruhe Institute of Technology, Karlsruhe, Germany

## Correspondence

Gang Wang  
Email: [gang\\_3.wang@boehringer-ingelheim.com](mailto:gang_3.wang@boehringer-ingelheim.com)

## Funding information

Boehringer Ingelheim Pharma GmbH & Co. KG

## Abstract

Current manufacturing and development processes for therapeutic monoclonal antibodies demand increasing volumes of analytical testing for both real-time process controls and high-throughput process development. The feasibility of using Raman spectroscopy as an in-line product quality measuring tool has been recently demonstrated and promises to relieve this analytical bottleneck. Here, we resolve time-consuming calibration process that requires fractionation and preparative experiments covering variations of product quality attributes (PQAs) by engineering an automation system capable of collecting Raman spectra on the order of hundreds of calibration points from two to three stock seed solutions differing in protein concentration and aggregate level using controlled mixing. We used this automated system to calibrate multi-PQA models that accurately measured product concentration and aggregation every 9.3 s using an in-line flow-cell. We demonstrate the application of a nonlinear calibration model for monitoring product quality in real-time during a biopharmaceutical purification process intended for clinical and commercial manufacturing. These results demonstrate potential feasibility to implement quality monitoring during GGMP manufacturing as well as to increase chemistry, manufacturing, and controls understanding during process development, ultimately leading to more robust and controlled manufacturing processes.

## KEYWORDS

automation, clinical manufacturing, machine learning, process analytical technology, Raman spectroscopy, therapeutic antibody

## 1 | INTRODUCTION

The development and manufacturing of biopharmaceutical products including monoclonal antibodies include inherent risks that need to be understood and controlled. Regulatory agencies and pharmaceutical companies rely on sophisticated control strategies and

technologies for controlling product quality attributes (PQAs) that could impact safety or efficacy (Berry et al., 2016; Rathore, 2014). One emerging technology that promises significant process control capabilities is process analytical technologies (PAT), which aims to monitor critical PQAs and subsequently control processes in real time (Tulsyan et al., 2019). In addition to applications in traditional batch

This is an open access article under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2023 Boehringer Ingelheim Pharma GmbH & Co KG. *Biotechnology and Bioengineering* published by Wiley Periodicals LLC.

bioprocessing, the implementation of PAT has been identified as a key technological driver enabling intensified continuous manufacturing (Khanal & Lenhoff, 2021; Pedro et al., 2022). Thus, there is a cross-industry consensus and interest in bringing PAT tools to implementation (Gillespie et al., 2022). Regulatory agencies have responded to industrial needs with updated guidelines that address best practices for the implementation of this evolving technology (ICH, 2021, 2022), including advising on potential pathways toward real-time release testing (RTRT). While the implementation of PAT in monitoring and controlling molecular targets, biopharmaceutical PQAs remain challenging to monitor using a rapid and nondestructive in-line sensor (Gillespie et al., 2022). Of particular interest is the ability to monitor product breakthrough and aggregation content in affinity chromatography, which would allow continuous downstream purification systems to operate with more efficiency and consistency (Feidl et al., 2019a, 2019b; Rolinger et al., 2020). Other unit operations also benefit from in-line monitoring, for example, product aggregation measurement using multiangle light scattering (MALS) during hydrophobic interaction chromatography (Patel et al., 2018). Ultimately, modern PAT implementation seeks to minimize manufacturing variability by monitoring product quality and controlling process parameters in real-time to guarantee narrower, more controlled quality attribute variations in the final pharmaceutical product.

Another promising approach to reducing manufacturing variability and risk is through increased understanding of the process by automated experimentation using high-throughput process development (HTPD) (Hubbuch, 2012). In addition to applications in early stage process development screening, HTPD and advanced automation robotics have emerged as a promising companion to monitoring and controlling processes by PAT, especially during the development of continuous processes (Silva et al., 2022). Current automation and miniaturization efforts have resulted in everincreasing numbers of experiments being performed automatically. This build-up of samples has created an analytical bottleneck that limits the ultimate generation of large quantities of representative, nonredundant data necessary for machine learning and advanced statistical analysis, since the corresponding analytical methods for measuring high volumes of HTPD samples are more difficult to automate and intensify. Computational analysis scripts are then necessary to render useful measurements from raw spectral data and readily available software for analyzing large quantities of data are similarly difficult to intensify. One potential approach for resolving manual data collection is to mechanistically model and computationally simulate PQAs for each unit operation, which maximizes understanding for a given amount of experimental effort (Saleh et al., 2021). Convolutional neural networks have also been applied to model the relationship between Raman spectral fingerprints and measures of product quality (Rolinger et al., 2021). Thus, there remain potential unexplored possibilities combining aspects of PAT, HTPD, and computational methods that will ultimately lead to exciting advanced process control technologies.

Of the common noninvasive rapid analytical techniques, Raman spectrometry is increasingly being evaluated for in-line pharmaceutical product quality measurement (Pedro et al., 2022). As a light-based technique, Raman spectrometry promises to continuously measure molecular attributes from a nonstationary sample, making it ideally suited as both a PAT tool for product quality monitoring, and as a HTPD tool for intensifying analytical throughput. Raman spectrometry has been used in clinical manufacturing for upstream operations (Esmonde-White et al., 2022), and work is on-going to develop the capability to monitor product quality. Currently, Raman spectrometry can feasibly measure therapeutic protein concentration (Goldrick et al., 2020), product aggregation (Zhou et al., 2015), fragmentation, charge variants, and oxidation (Wei et al., 2021). Although calibration models perform well when considering off-line data, several theoretical and experimental challenges arise when generalizing off-line models to in-line process monitoring. At present, it is possible to measure protein concentration using an in-line Raman sensor (Feidl et al., 2019a, 2019b; Rolinger et al., 2021; Yilmaz et al., 2020). Other in-line sensors have also been investigated for monitoring aggregates under flow, for example, using NMR (Taraban et al., 2019), MALS (Patel et al., 2018), and UV-Vis (Brestrich et al., 2018). Thus, we sought to leverage automation tools to intensify Raman model calibration data collection and evaluate commonly used machine learning models that scale with larger datasets.

## 2 | MATERIALS AND METHODS

### 2.1 | Preparation of critical quality attribute (CQA) calibration material

Stock solutions of therapeutical monoclonal antibody product, force-degraded protein product, and buffer were used to generate mixtures of concentrations and aggregation levels used to calibrate Raman models. One set of pure and degraded product samples was generated for each buffer of interest, and thus it was necessary to generate a unique Raman model for each buffer. To generate aggregates, product samples were frozen to  $-20^{\circ}\text{C}$  and stored for up to 72 h. Freeze-thaw was preferred over other degradation methods due to its specificity in generating aggregates, and due to its relevance in real-world processes (Nowak et al., 2017). Both original and degraded material were exchanged into the buffer of interest by ultra- and diafiltration (UF/DF). Samples were stored at  $-70^{\circ}\text{C}$  until needed, and thawed samples were analyzed by analytical size exclusion ultra-performance liquid chromatography (UPLC) for high and low molecular weight species. Analytical chromatography was performed on an Acquity UPLC system (Waters Corporation) using a BEH200SEC.  $4.6 \times 300$  mm column, and a mobile phase consisting of 200 mM L-Arginine, 120 mM ammonium sulfate, 10% isopropanol, and pH 7.3 adjusted by phosphoric acid. The mobile phase was run under a flow rate of 0.2 mL/min and the targeted sample load was 30  $\mu\text{g}$ /injection.

## 2.1.1 | In-line process analytical measurements

The FlowVPE was paired with a Cary 60 UV-Vis spectrophotometer which allowed measuring light absorbance at a range of wavelengths as well as a range of pathlengths. We used a fixed pathlength of 1 mm on the FlowVPE and a wavenumber range of 190–350 nm at 0.5 nm increments, resulting in 321 absorbance measurement. Four scans were taken over 10 min and the absorbance measurements were averaged. Recording times were stored in the filenames of both the UV-Vis and Raman result files for data alignment offline. Product concentration measurements were validated by comparing results with those obtained on a NanoDrop (Thermo Fisher Scientific Inc.). Raman acquisition settings were optimized for maximizing the signal-to-noise ratio for a given dynamic range of the experimental system (Rolinger et al., 2021). A HyperFlux Pro Plus Raman spectrometer (Tornado Spectral Systems) was equipped with a standard 785 nm laser and operated at the maximum power 495 mW. A flow cell with a dead volume of 200  $\mu$ L (Marqmetrix) was equipped with the spectrometer and was connected in-line to the AKTA Avant 25 system directly downstream of the built-in UV flow cell. To determine the optimal exposure time, we performed multiple autoexposure scans on several buffers common to downstream purification and product formulation, and we determined that 623 ms exposure time and an averaging of 15 exposures per scan was optimal for maximizing the signal-to-noise ratio of product protein concentration, aggregation, and fragmentation. Raman spectra were saved to comma-separated value (CSV) files with the time-stamp of the beginning of recording as the filename, where the time stamp contained the date-time down to 1 ms. In contrast to the UV-Vis, the Raman collected spectra continuously during the experiment, with only the correctly time-aligned files being used at analysis time.

All machine learning model training, testing, and visualization were carried out in Python. The PLS regression model was created using the `PLSRegression` function from the module `sklearn.cross_decomposition` using the default parameters, the principal component regression (PCR) model was created using the `PCA` function from the module `sklearn.decomposition`, and the `LinearRegression` function from the module `sklearn.linear_model`. The  $k$ -nearest neighbors (KNN) model was made using the `KNeighborsRegressor` function from the module `sklearn.neighbors`, using the “distance” as weights. Raman spectra saved as CSV were parsed for the text content and organized according to the timestamps to be paired with the corresponding experiment number and thus mixed concentration and aggregation levels. Odd-numbered spectra files representing 9.3 s of Raman recording were gathered and averaged to make up the training data set. The remaining even-numbered files were averaged to make up the testing data set. After initializing the corresponding regression models, the training data were fed into the model using the `fit` function, and prediction of concentration and aggregation were obtained using the `predict` function. To prevent arbitrary predictions from very small changes to the input Raman spectra, a low amount of Gaussian noise of 1% of the standard deviation was added to both

the training and testing datasets. The `normal` function from the module `numpy.random` was used to randomly sample this noise.

## 2.2 | Mixing calibration on AKTA bench-scale chromatography systems

In mixing experiments where the FlowVPE (C Technologies, Inc.) was utilized, the flow-cell was attached directly downstream of the integrated in-line UV detector. A flow-cell with a dead volume of 0.9 mL was installed onto the FlowVPE, and the FlowVPE was installed after the in-line UV detector of the AKTA Avant 25 chromatography system (Cytiva). The flow path was set to bypass the column and flow restrictor. Each of the pure product, degraded product, and buffer stocks were assigned to inlets Q1-4 for the quaternary valve. The quaternary gradient variables were set by scouting runs. A flow rate of 10 L/min was used to flush using the initial quaternary gradient for 15 mL. A flow rate of 0.1 mL/min was set as the calibration rate, and a digital signal was sent through the input/output (I/O) box to mark the start of calibration. The calibration lasted 1 mL, or 10 min, resulting in a total material usage of 16 mL, plus minor additional volumes wasted between pilot run resets. The order of each mixing run was further randomized to eliminate carry-over effects in between pilot runs. The randomization was based on a unique seed which was saved to recreate the pseudorandom ordering during analysis.

Single and multivariate mixtures were used to generate quaternary gradient variables using the `itertools` Python library, and the resulting gradient percentage values were algorithmically added to an AKTA Unicorn base method by modifying the XML-based method instructions and updating the 64-bit cyclic redundancy check (CRC) hash codes. To sign each new XML file for protection against file corruption, the 32-bit CRC hash was first computed using the `zlib` Python library. The resulting CRC32 hash in hexadecimal notation was extended with leading zeros until an eight-character string was achieved to avoid the rare cases when the resulting hash had fewer than eight characters. The 32-bit hexadecimal hash string was then unpacked as a big-endian signed `int32` data type. This resulting signed integer value was then packed as a signed big-endian `int64` hexadecimal string in all-capital letters and with leading zero characters removed. After obtaining the verification hash strings, they were entered in the new XML files where appropriate to ensure that no data had been corrupted. Following this procedure, we created one template method with the mixing protocol and added a new scouting run for each unique mixture.

A method for communicating between the AKTA system and the computer that controlled the Raman and FlowVPE analytical instruments was necessary to align acquisition times to sample flows. An I/O box for AKTA systems was installed to send a digital output signal each time the quaternary gradient as set, marking the beginning of a calibration instance. The digital output signals from the Akta were recorded using a USB-6009 multifunction data acquisition (DAQ) device and the `NIDAQmx` Python library (National Instruments). The digital out pins from the AKTA I/O box were connected directly to the digital in ports by

jumper wires. A light-emitting diode light and resistor were connected to the digital out ports on the DAQ device to indicate the status of the script. A series of scripts was used to trigger the collection of FlowVPE data using Python, the Windows batch scripting system, and the Agilent ADL API (Agilent Technologies, Inc.). The collection time on the Cary UV-Vis spectrophotometer connected to the FlowVPE was shorter than the total flow duration on the AKTA system, thus communication in the reverse direction from the FlowVPE to the Unicorn software was not necessary and not established. Start and stop times of these digital events were also used to align the FlowVPE spectra to the Raman spectra and the corresponding quaternary gradient variables from the AKTA. Using a measuring caliper, the detector head was manually positioned by incrementing the stepper motor rotation step-by-step from 1747 steps (0 mm pathlength) to 1547 steps (1 mm pathlength), for a step rate of 200/mm. At the start of each scan, the stepper motor position was set by the *SetAcc* function and the number of steps as the distance parameter. Next, UV-Vis scanning settings were set by changing variable values with the *SetVal* function, setting the value "UVVIS averaging" to 28.0 s and "UVVIS interval" to 0.5 nm. Next, the scan was initiated using the *CollectAndGraph* function with "Scan Data" and "Scan Graph" as the first two parameters, the start and stop wavelengths in nanometers, for example, 190–350 nm, as the third and fourth parameters, and the keyword "Wavelength\_XMode" as the fifth and last parameter. At the end, the scanned data is exported to a CSV file using the function *FConvert* with "Scan Data" as the first parameter, the numeric value 1 as the second parameter, and the output filename as a string as the third and last parameters.

The execution of the Cary ADL script for triggering the UV-Vis scan was conducted through a separate Python script. The *nidaqmx* library was used to receive digital output data coming from the AKTA system continuously, filling up a constant array of 30 samples, which required approximately 3 s to fill. At every data query point, the new voltage value was added to the array while removing the oldest sample. Next, if the absolute maximum of the array voltages was above or below 12 standard deviations of the typical voltage fluctuation (2.83 mV), then the appropriate state was switched between the high and the low digital states. Additionally, another state change within 2 s of any previous change is ignored to filter out noise. At the trigger of a state change in either direction, the Cary ADL script was executed using the *check\_output* (shell = True) function from the *subprocess* library. Since the original source text of the ADL script was not modified, it deposited data with one filename for every scan, thus another function was created to rename the UV-Vis output CSV file by including the system time in the filename. A live plot of the current-voltage array was displayed to confirm to the operator that the scans were being consistently triggered.

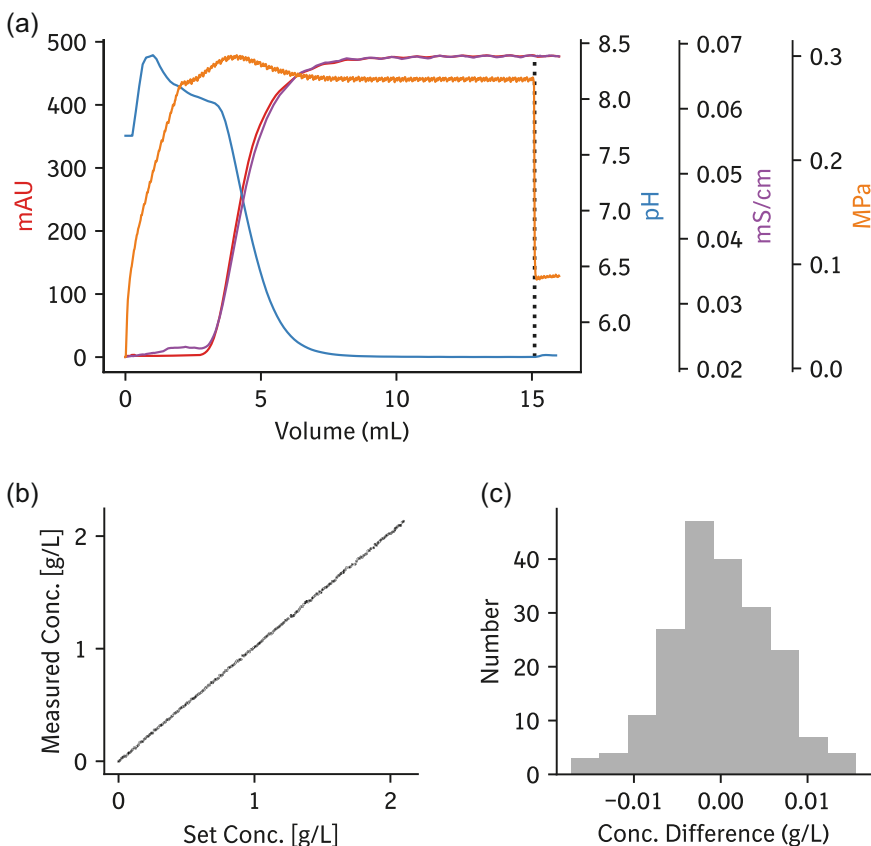
### 3 | RESULTS

A mixing script was assembled in Python for systematically programming multiple compositions of stock solutions on AKTA chromatography systems to generate samples of known mixed ratios

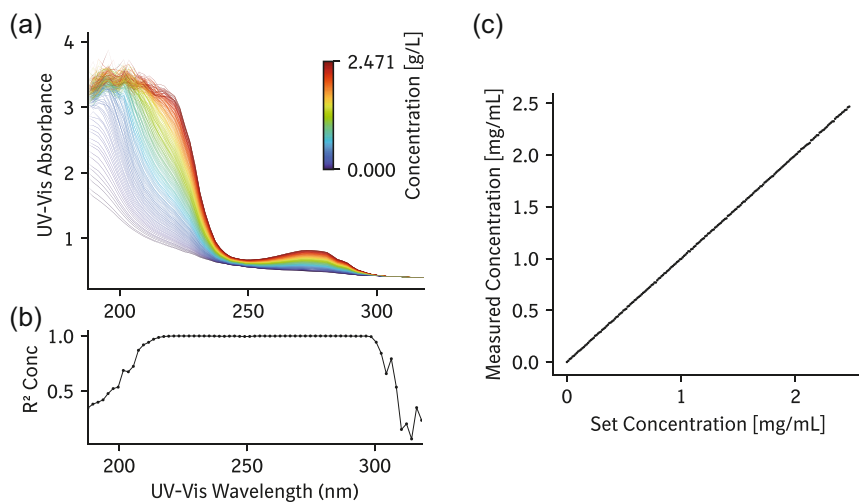
which are then channeled through in-line analytical flow-cells (see Section 2). Early attempts to achieve well-mixed fractions used 9 mL of wash volume at a flow rate of 10 mL/min, however, this resulted in incomplete mixing ( $p = 0.002$ ,  $n = 8$ , two-sided  $t$  test). Thus, a larger volume of 15 mL was decided for mixing (Figure 1a). First, the automated method was validated by mixing water with formulated drug substance as the pure product protein stock (0–2.097 g/L,  $n = 197$ ). The product concentrations of each fraction were measured by Nanodrop in duplicate after each mixing run. The averages of these resulting empirical concentration measurements were then fitted against the algorithmically set concentrations, resulting in an excellent linear fit ( $R^2 = 0.99991$ , intercept = 0.0005, slope = 1.01,  $n = 197$ ) (Figure 1b). The error between the set and empirical concentrations appeared to be normally distributed (Figure 1c), showing a Kullback–Leibler (KL) divergence of 0.015 bits as compared to a baseline divergence of 0.43 bits from the uniform distribution. To quantify variations in concentration useful for engineering, we performed repeat measurements on the Nanodrop and compared errors with concentration values measured by UV-Vis. Repeat measurements ( $n = 16$ ) were performed with a separate stock solution (2.471 g/L) to quantify the Nanodrop error, resulting in a standard deviation of 0.008 g/L (0.3%). The standard deviation of the differences between the set concentrations and measured concentrations of the 197 mixtures was 0.004 g/L, which was comparable to the nanodrop measurement standard deviation. Due to the long experiment duration of several hours, six measurements of product concentration by nanodrop before and after the experiment were collected to assess stability. No statistically significant differences in product concentration were found after conducting a two one-sided test (TOST) at an equivalence range of  $\pm 3\sigma$  between 2.08 and 2.12 g/L (is equivalent,  $p = 0.001$ ). Thus, we were able to demonstrate an automated AKTA-based mixing system that only requires one stock solution and could be controlled by automatic scripting.

Once the AKTA system was validated for mixing, we installed in-line sensors to test whether the system could produce mixtures robust enough for calibration. First, we installed a FlowVPE in-line UV-Vis spectrophotometer manufactured by CTech (Bridgewater) in front of the integrated in-line UV detector of the AKTA. We integrated the hardware and software across systems to allow the UV-Vis to automatically scan after the active mixing phase (see Section 2). We were able to perform the calibration with extremely reduced effort and hands-on time by simplifying preparation to two stock solutions, only one of which required off-line analytical analysis, and by running automated scripts to handle method generation, instrument communication, and data analysis. A stock solution of 2.471 g/L was prepared and mixed 201 times with water to generate calibration UV-Vis scans (Figure 2a). The smoothness of the UV-Vis scan line transitions between one calibration point to the next is notable here, greatly contributing to a more robust and accurate calibration data set. After fitting a linear regression between the UV-Vis scanned absorbance values and the set concentration, we achieved excellent linearity across the scanned wavelengths (Figure 2b), with an overall linearity of  $R^2 = 0.99997$  (Figure 2c). The standard deviation in prediction was

**FIGURE 1** Mixing chromatogram for one automated calibration sample. The UV absorbance, pH, and conductivity are plotted during mixing between stock therapeutic monoclonal antibody product samples of known concentrations. All indicators reached equilibrium near the end of mixing (a). Mixed samples were validated by manual off-line concentration measurements (b) to be equivalent to the algorithmically determined set points. This resulted in the creation of 197 calibration sample points from two manual measurements of starting stock product solutions. The error between the set and empirical concentrations appeared to be normally distributed (c).



**FIGURE 2** Automated in-line calibration of product concentration using UV-Vis absorbance. A custom automatic mixing system was used to collect a series of UV spectra ( $n = 201$ ) from two stock solutions. Absorbance in the 190–320 nm range was visibly correlated to product concentration (a), demonstrating good linear fit over most scanned wavelengths (b). After fitting UV scans to algorithmically set mixed protein concentrations, the integrated automated system achieved linearity (c) superior to that of manual measurement (Figure 1b).



0.003 g/L (0.1%), which is lower and more accurate than what was achievable by hand and with the nanodrop detector (Figure 1b), demonstrating superior accuracy and consistency in the automated system. Thus, we have demonstrated the capability of a mixing system that relies extensively on hardware and software automation, that can generate a calibration model from UV-Vis spectra with only one stock solution of known product concentration.

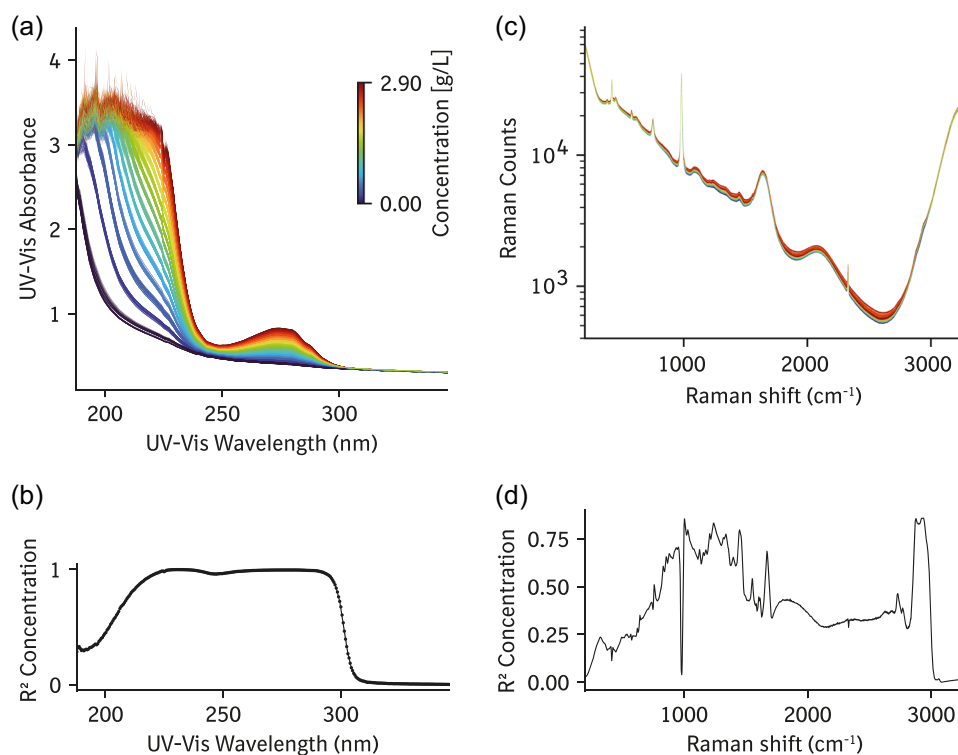
Once the mixing system was validated against existing off-line product protein concentration measurements, we installed an in-line

Raman spectrometer manufactured by Tornado in front of the FlowVPE flow-cell to test its ability to quantify product concentration and aggregation. Again, we made use of extensive automation systems to align all the mixing experiments with the correct UV-Vis and Raman spectra results (Section 2). Recently, the ability to measure several PQAs by Raman spectrometry was demonstrated in a therapeutic protein (Wei et al., 2021), thus we investigated whether it would be possible to calibrate our in-line Raman spectrometer to measure product aggregation with a stock sample of enriched

aggregates. Being able to monitor product quality in a noninvasive in-line sensor would dramatically increase analytical process knowledge and understanding, enabling a wide range of new industrial applications, such as intensified bioprocesses (Pedro et al., 2022). The enriched aggregate stock was made by forced degradation, specifically freeze-thaw cycling (Nowak et al., 2017), and all stocks were buffer exchanged into a cation exchange elution buffer (Section 2). Aggregation content was measured using analytical size exclusion UPLC, by summing all high molecular weight (HMW) peak groups. A full-factorial design-of-experiments (DOE) experiment was planned with product concentration and aggregation as two factors with 10 values each, resulting in a 100-point calibration. The monomer-enriched stock solution contained 2.90 g/L total product with 1.3% aggregation, the aggregate-enriched stock solution contained 2.91 g/L total product with 1.8% aggregation, and the last stock solution contained elution buffer, which was the same buffer used for all stock solutions. The UV-Vis scan continued to show linearity with increasing protein concentration (Figure 3a), and high correlation coefficients between concentration and absorbance in individual wavelengths (Figure 3b), with an overall linearity of  $R^2 = 0.99993$  and a prediction standard deviation of 0.005 g/L (0.2%). The Raman spectra consisted of light intensity counts across 3101 wavenumbers from 200 to 3300  $\text{cm}^{-1}$  and showed a similar correspondence between intensity and protein concentration (Figure 3c) as the UV-Vis scans. This correlation is quantified for

each individual wavenumber (Figure 3d), and although the correlation coefficients are not as high as for the UV-Vis scans, the overall linearity was  $R^2 = 0.9993$ , with a prediction standard deviation of 0.016 g/L (0.6%). Notable peaks in the Raman spectra sensitive to product concentration included the range 998–1043, 1204–1274, 1443–1462  $\text{cm}^{-1}$ , and the plateau at 2863–2970  $\text{cm}^{-1}$ . Although the differences in aggregation were relatively small in magnitude, calibration of product protein concentration was still achievable using the UV-Vis spectra. This capability remained when calibrating the Raman spectra, demonstrating that it was possible for the automated calibration system to generate meaningful Raman calibration models as well.

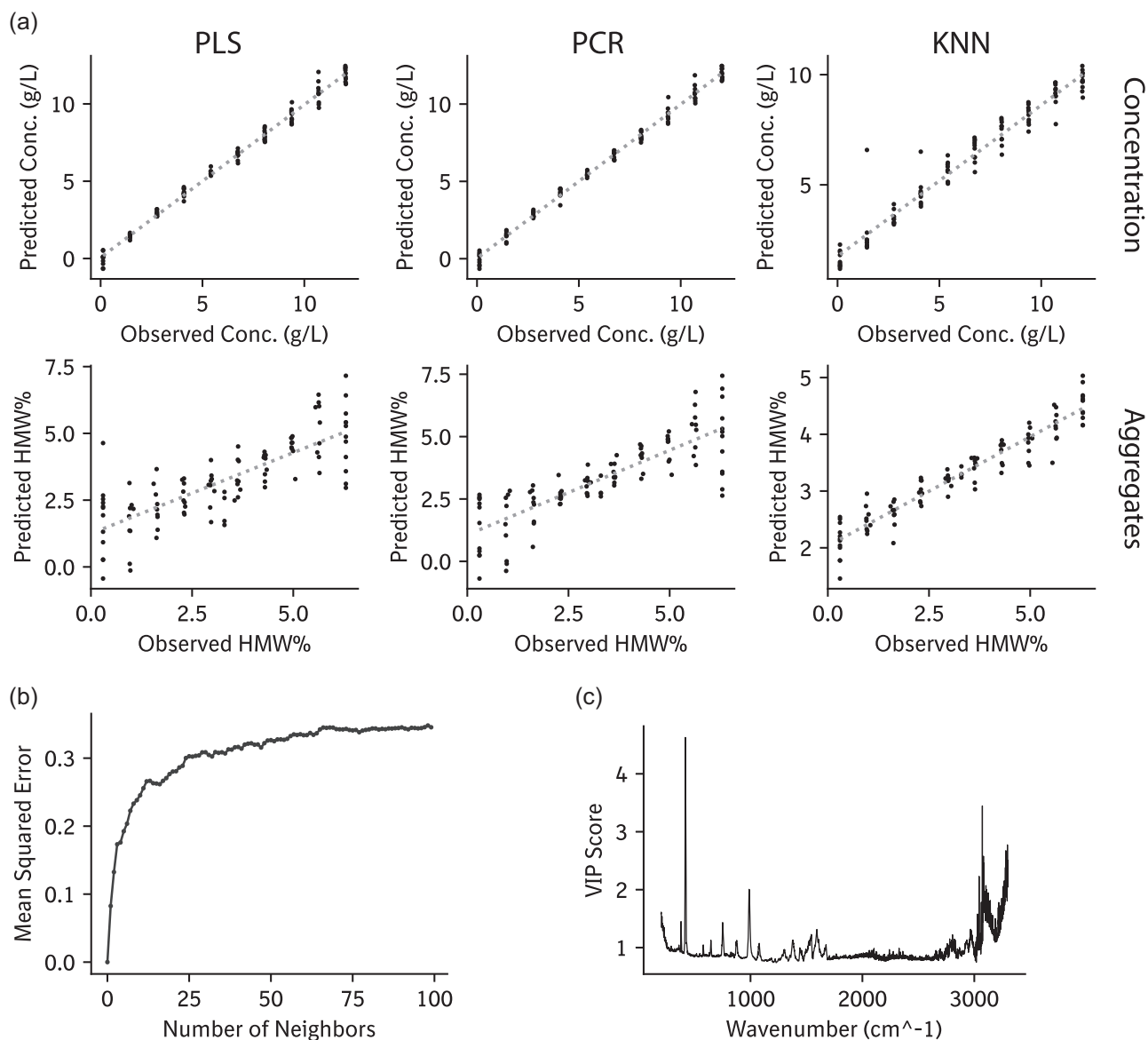
A second full-factorial experiment was planned with a wider range of aggregates to test whether it was possible to calibrate the Raman in-line sensor to measure product aggregation, a quality attribute that is often considered critical in chemistry, manufacturing, and controls (CMC) strategy. Using the same methods as the first combined UV-Vis and Raman calibration, we created new stock solutions containing 12.01 g/L total protein and 0.3% aggregation in the monomer-enriched sample and 12.04 g/L total protein and 6.3% aggregation in the aggregates-enriched sample. We were confident from our previous results (Figure 3) that our system was capable of producing stable mixtures for calibrating both UV-Vis scans and Raman spectra, thus we updated our analysis of the Raman spectra to include methods that have been recently successful for calibration



**FIGURE 3** Automated in-line design-of-experiments calibration of product and aggregate concentration using combined UV-Vis and Raman spectroscopy. The UV-Vis spectra for a 100-point full-factorial run are shown and colored by the expected mixed protein concentrations (a). Product concentrations were accurately modeled over most scanned wavelengths (b). As shown for the UV-Vis spectra, the Raman spectra are similarly colored by the same concentration range (c), and were also able to be modeled over a wide range of Raman shift wavenumbers (d).

model building, including partial-least squares (PLS) regression modeling (Brestrich et al., 2018; Wei et al., 2021) and PCR (McAvan et al., 2020; Ramakrishna et al., 2022; Silva et al., 2020), as well as nonlinear models such as Gaussian processes (GPs) (Tulsyan et al., 2020). Model hyperparameter selection was made by subjective judgment from published values and included: 10 for the number of PLS and PCA latent variables, and 60% as  $k$ , for the KNN model that we used instead of a GP model to reduce risk of overfitting. A total of 10 min of recording was made for each of 100 calibration points, and these recordings were further subdivided into 9.3 s segments. The odd-numbered chunks were averaged together to make up the training set, while the even-numbered chunks made

up the testing set. The three models were then trained and tested using this scheme, and the results of predicting both concentration and aggregate as a percentage of HMW species (HMW%) is shown in Figure 4a. Since we did not have a solid understanding of the impact of the  $k$  hyperparameter in the KNN model, we repeated the training and testing for all values of  $k$  from 1 to 100, showing decreasing performance with increasing  $k$  (Figure 4b). We chose the value of 60 for  $k$  due to its centrality in a relatively stable region of the hyperparameter space with a lower impact on performance. A quantitative summary of the model performance is shown in Table 1, with PLS and PCR models performing well for predicting concentration, achieving linearities of  $R^2 > 0.99$ . The PLS model is a supervised



**FIGURE 4** Performance of different machine learning models in predicting product and aggregate concentration from Raman spectra. The predicted product and aggregate concentrations of the testing set are plotted over values from the training set for three commonly used models. The impact of the hyperparameter  $k$  on KNN performance is plotted in terms of the mean squared error (b). VIP scores identified regions of the spectra most important for the calibration model and are shown in (c). HMW, high molecular weight; KNN,  $k$ -nearest neighbor regressor; PCR, principal component regressor; PLS, partial least squares regressor; VIP, variable importance in projection.



variant of PCR, choosing automatically features that best represent the relationships between the Raman spectra and product quality measures. We demonstrate empirically that the PLS model does perform better than the PCR model for concentration in terms of the mean absolute percent error (MAPE), but the results were not clear for mean absolute and squared errors. For aggregation, we find the same trend that the PCR model performs better than PLS on all metrics except the MAPE. After training the PLS model on both concentration and aggregates, we computed the variable importance in projection (VIP) scores to identify wavenumbers most important for the calibration model (Mehmood et al., 2012). We found that the main sapphire peak  $415\text{--}422\text{ cm}^{-1}$  showed a VIP score of greater

than 2, as well as the regions  $987\text{--}990$ ,  $3043\text{--}3097$ , and  $3266\text{--}3300\text{ cm}^{-1}$  (Figure 4c). The KNN model achieved  $R^2 = 0.91$  on product aggregation over a range of 0.3% to 6.3%HMW, which out-performed both PLS and PCR models ( $R^2 < 0.62$ ). The KNN model is a nonlinear assumption-free model that does not infer any patterns in the data and seems to work quite well for this application. Thus, we have demonstrated the capability of our automated calibration system to create calibration models for measuring product concentration up to  $12.04\text{ g/L}$  and product aggregation up to 6.3%.

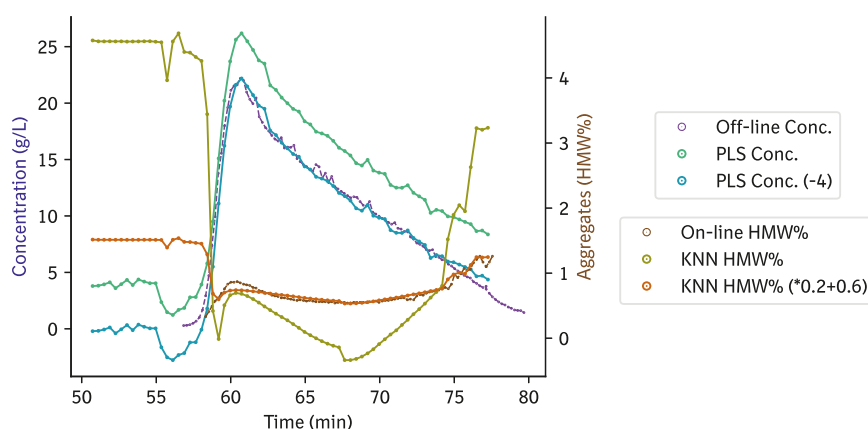
To demonstrate the value of using an in-line sensor, we investigated whether the calibration model created by the automated system could be applied to measure product concentration and aggregation in real-time in a real-world product purification process. We installed our in-line Raman probe used for calibration on an AKTA system and recorded the Raman spectra continuously using the same recording parameters and used the calibration models (Figure 4) to predict the concentration every 22.5 s, or an average of three spectra. The results of off-line analytical analysis of elution fractions and the in-line Raman predictions are shown in Figure 5, where the PLS model was used for product concentration and KNN model was used for aggregation. The model captured the main features of the elution profile and was comparable to off-line results after applying simple linear transformations. The transformed values fit not only the concentration range of calibration, but also the range at the beginning of the elution phase around 51 min that exceeded the calibrated maximum concentration of  $12.04\text{ g/L}$ . It may be possible to bridge the difference between predicted and actual values by using appropriately high stock concentrations for calibration to cover the entire elution range. Thus, we demonstrate the possibility for an off-line calibration model trained to measure product concentration and aggregation after linear adjustment, as well as the ability to predict out-of-range values once fitted to the in-range data set.

**TABLE 1** Performance of different models used for in-line DOE calibration.

Model	Target	$R^2$	R	MMSE	MAE	MAPE (%)
PLS	Conc (g/L)	0.990	0.995	0.15	0.28	13.3
	HMW (%)	0.620	0.790	1.47	0.91	48.5
PCR	Conc (g/L)	0.992	0.996	0.12	0.27	15.5
	HMW (%)	0.679	0.824	1.23	0.79	213
KNN	Conc (g/L)	0.947	0.973	1.83	1.08	24.7
	HMW (%)	0.909	0.953	1.51	1.05	34.6

Note: As illustrated in Figure 4, the performance of the PLS, PCR, and KNN on a 100-point calibration data set is shown using the metrics  $R^2$ , R, mean squared error (MSE), mean absolute error (MAE), and mean absolute percent error (MAPE).

Abbreviations: DOE, design-of-experiments; HMW, high molecular weight; KNN, k-nearest neighbor regressor; MAPE, mean absolute percent error; PCR, principal component regressor; PLS, partial least squares regressor.



**FIGURE 5** In-line monitoring of product and aggregate concentration. Cation exchange chromatography elution was monitored by both off-line fractionation analysis using UPLC and in-line analysis using Raman measurement. Violet, green, and blue lines indicate off-line product concentration, in-line Raman concentration, and adjusted in-line Raman measurements. The coefficients and offsets used for the Raman adjustment are indicated for the corresponding lines in the legend. Following the format for concentration, brown, gold, and orange lines indicate the corresponding measurements for product aggregation. HMW, high molecular weight; KNN, k-nearest neighbor regressor; PLS, partial least squares regressor; UPLC, ultra-performance liquid chromatography.

## 4 | DISCUSSIONS

We investigated the capabilities of the AKTA system and demonstrated controlled, consistent mixing from two or three stock solutions. We achieved this by minimizing sample volume consumption while achieving sufficient mixing. This optimization resulted in a sample consumption of 15 mL, which produced mixtures that were equivalent to that of manual methods (Figure 1). Our resulting automation system can systematically mix samples of known analytical properties to significantly intensify off-line analytical measurement effort by up to 100-fold and is readily adaptable for measuring other PQAs such as fragments or charge variants (Wei et al., 2021). The ability to accelerate analytical data collection is critical since off-line analytical testing is a limiting step for major decision-making milestones in both development and commercial manufacturing. Previously, this analytical bottleneck has been identified as the root cause of the "Low-N" problem in biomanufacturing (Tulsyan et al., 2019), which implies that limited training data set sizes are the root cause of less than ideal performance of machine learning models such as PLS and PCA. Variations in representativeness between training, testing, and real-world data due to changes in process materials or parameters also diminish model performance. While one strategy of solving this data scarcity is by innovating new data analysis methodologies to more effectively make use of precious data (Tulsyan et al., 2020), we demonstrate here the potential of using automated and high throughput mixing to generate new samples of known analytical properties. Theoretically, an unlimited number of product-related impurities could be analyzed and mixed to generate intermediate samples of known analytical values, provided that the mixing hardware automation system can produce the algorithmically set mixing ratios. We validated our hardware system by manually measuring concentrations of 197 mixed samples from one stock solution (Figure 1a) and demonstrated a 201-point calibration of product concentration using UV-Vis scans (Figure 2). If scaled up, this could translate to a theoretical scale-up factor of 200 $\times$  in the analytical output of existing manual calibration experiments, where each calibration point is analyzed by a panel of off-line analytical testing. If all four quaternary valve inlets were filled with stock solutions, three analytical measurements would be needed to generate a calibration for three product quality values, leading to a data collection rate of 201 data points per day. Current methods such as PLS and PCA rely on dimensionality reduction to reduce 3101 Raman variables (Figure 3) down to 10 (Ryabchykov et al., 2018), and this is done because there would not be enough equations (number of calibration experiments) to solve the regression coefficients (number of wavenumber variables). However, with the automated capability presented here, it would be possible to solve all 3101 Raman variables with just three AKTA system-weeks of collection time.

We extended our initial study of product concentration to involve both product and aggregate concentration (Figure 4), demonstrating that our calibration system was able to model the relationship between Raman spectra and product aggregation UPLC

measurements. While Raman has been calibrated for measuring product aggregation previously, and for other PQAs (Wei et al., 2021), we demonstrate here the first use of a nonlinear KNN model for predicting aggregation levels ranging from 0.3% to 6.3%. We also demonstrate the ability to use both PLS and KNN models trained on off-line calibration samples to measure values out of the calibration range in a real-world cation exchange preparatory chromatography operation (Figure 5), and we achieved this with only two off-line analytical test samples for the calibration stock solutions. While several previous studies investigated real-time bioprocessing monitoring at the affinity capture step (Feidl et al., 2019a, 2019b; Rolinger et al., 2021), the ability to monitor aggregates in real-time has not yet been established. Of the studies that look at noninvasive methods for measuring multiple quality attributes (de Faria e Silva et al., 2020; Taraban et al., 2019; Wei et al., 2021; Zhang et al., 2019), the transferability of models trained on off-line calibration data for measuring in-line results remains an open question. We greatly increased the analytical data throughput and data representativeness by performing calibration on the same hardware as the real-world chromatography operation, and we demonstrated how a nonlinear KNN model can leverage this data set to reproduce meaningful product quality patterns in real-time (Figure 5). To achieve the goal of being able to apply the off-line calibration consistently without additional adjustments, more understanding of the Raman spectra and how features are translated to product quality measurements is needed, although this task may be increasingly complex (Guo et al., 2021). The sapphire peak is one of the most prominent features of the Raman spectra and it was also the most important feature in our calibration model (Figure 4c). Since sapphire was the material least likely to change properties in our experiment, this could be a baseline effect that need to be filtered out by one or a combination of baseline correction methods (Byrne et al., 2015; Ryabchykov et al., 2018). Further, normalization or smoothing techniques could be used to pre-process the data before modeling to improve the model quality (Wei et al., 2021). We anticipate the application of baseline removal methods such as those based on frequency filtering would greatly reduce sources of bias originating from confounding effects such as flow rate, background fluorescence, and variations in laser characteristics (Zhang et al., 2010).

The implementation of PAT is expected to dramatically reduce risks involved in both development and manufacturing (Berry et al., 2016; Gillespie et al., 2022; Rathore, 2014), where real-time monitoring would be able to flag out-of-spec intermediate analytical checkpoints. One key aspect of risk-minimization strategy is to understand how in-line sensors such as UV absorbance probes (Gillespie et al., 2022) make measurements, when they become saturated, and how to overcome these limitations by for example varying the pathlength. Chemometric analysis pose severe challenges to this understanding (Guo et al., 2021) due to very large numbers of processing steps and the algorithmic complexity of those steps. Individual amino acids show distinct Raman fingerprints that can be clearly identified (Zhu et al., 2011), thus serving as the chemical foundation for the theoretical possibility of deconvolving every

complex product spectrum down to its component root elements. Until we have sufficient data and mechanistic understanding to do this, we have employed single-variate analyses where possible to try to simplify this algorithmic transparency problem (Figures 2b and 3b,d). The VIP score for PLS model evaluation has additionally proven extremely useful in this respect (Mehmood et al., 2012) (Figure 4c), although this technique does not readily generalize to the PCR and KNN models (Table 1). In addition to single wavenumber effects, protein aggregation may additionally interact with indirect sources of Raman distortions such as the flow of water (Taraban et al., 2019). Until the complex relationship between Raman spectral features and product quality variants is elucidated (Zhu et al., 2011), it will be critical to fully characterize any potential impact of process parameter deviations on calibration model representativeness in an industrial application. The reduction of the number of latent variables for both PLS and PCR has been a common technique for reducing model over-fitting (Wei et al., 2021), due to the relatively higher Raman spectral dimensionality compared to the number of experiments (Tulsyan et al., 2020). The literature seems to accept a wide variety of values for this parameter, ranging from as low as 2 (Brestrich et al., 2018) to 12 (Feidl et al., 2019a). We used a value of 10 (Figure 4a) as the number of variables to balance for enough explained variance while not overfitting characteristics of the off-line calibration training data set that would result in difficulties in predicting the in-line PQAs. We were also careful to mitigate this risk when pioneering the use of the nonlinear KNN model by studying the effect of the number of neighbors parameter on prediction accuracy (Figure 4b), choosing the conservative value of 60 as the number of neighbors which did not necessarily result in the most accurate prediction, but did result in a shallow slope, or the least change in performance for any given variation in the hyperparameter. Although the KNN model was a nonlinear model, it shared similarities to the GP regressor, which has been demonstrated in a bioprocessing monitoring context (Tulsyan et al., 2020), and which does not need large quantities of data for training because it does not assume any class of kernel functions whether linear or nonlinear.

## 5 | CONCLUSIONS

We demonstrate a hardware and software automation system that combines algorithmic control of experiments and machine learning modeling to create calibration datasets at high volume and low variability. The mixing capabilities were first established by manual measurements, then by multiwavelength UV-Vis scans. Once the relationship between algorithmically set concentrations and actual empirical measurements was established, we demonstrated that it was possible to then calibrate a Raman spectrometer to simultaneously measure product concentration and aggregation. We demonstrated the use of a nonlinear KNN model for accurately measuring product aggregation during bioprocessing is promising because the predicted elution profiles were highly correlated to the offline measurements. The expansion of other product qualities such as low molecular species and

charge variants naturally present an opportunity for further work. We envision great potential for in-line monitoring, once implemented, to shape the future of CMC control strategy and advanced manufacturing.

## AUTHOR CONTRIBUTIONS

Jiarui Wang, Jingyi Chen, and Gang Wang conceived and designed experiments and computational analyses. Jiarui Wang and Jingyi Chen performed experiments and composed analytical software. Gang Wang, Joey Studts, and Juergen Hubbuch determined the strategic fit and provided feedback. Jiarui Wang wrote the manuscript and all authors reviewed the manuscript.

## ACKNOWLEDGMENTS

The authors would like to thank Angelika Froeschle and Sabine Welte for the discussion of the main results and automation strategy. This work is funded by Boehringer Ingelheim Pharma GmbH & Co. KG.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author upon reasonable request and require legal agreements before sharing.

## ORCID

Jiarui Wang  <http://orcid.org/0000-0003-0301-5016>

## REFERENCES

- Berry, B. N., Dobrowsky, T. M., Timson, R. C., Kshirsagar, R., Ryll, T., & Wiltberger, K. (2016). Quick generation of Raman spectroscopy based in-process glucose control to influence biopharmaceutical protein product quality during mammalian cell culture. *Biotechnology Progress*, 32, 224–234.
- Brestrich, N., Rüdts, M., Büchler, D., & Hubbuch, J. (2018). Selective protein quantification for preparative chromatography using variable pathlength UV/Vis spectroscopy and partial least squares regression. *Chemical Engineering Science*, 176, 157–164.
- Byrne, H. J., Knief, P., Keating, M. E., & Bonnier, F. (2016). Spectral pre and post processing for infrared and Raman spectroscopy of biological tissues and cells. *Chemical Society Reviews*, 45, 1865–1878.
- Esmonde-White, K. A., Cuellar, M., & Lewis, I. R. (2022). The role of Raman spectroscopy in biopharmaceuticals from development to manufacturing. *Analytical and Bioanalytical Chemistry*, 414, 969–991.
- de Faria e Silva, A. L., Elcoroaristizabal, S., & Ryder, A. G. (2020). Multi-attribute quality screening of immunoglobulin G using polarized excitation emission matrix spectroscopy. *Analytica Chimica Acta*, 1101, 99–110.
- Feidl, F., Garbellini, S., Luna, M. F., Vogg, S., Souquet, J., Broly, H., Morbidelli, M., & Butté, A. (2019a). Combining mechanistic modeling and Raman spectroscopy for monitoring antibody chromatographic purification. *Process*, 7, 683.
- Feidl, F., Garbellini, S., Vogg, S., Sokolov, M., Souquet, J., Broly, H., Butté, A., & Morbidelli, M. (2019b). A new flow cell and chemometric protocol for implementing in-line Raman spectroscopy in chromatography. *Biotechnol Progr*, 35, e2847.
- Gillespie, C., Wasalathanthri, D. P., Ritz, D. B., Zhou, G., Davis, K. A., Wucherpfennig, T., & Hazelwood, N. (2022). Systematic assessment

- of process analytical technologies for biologics. *Biotechnology and Bioengineering*, 119, 423–434.
- Goldrick, S., Umprecht, A., Tang, A., Zakrzewski, R., Cheeks, M., Turner, R., Charles, A., Les, K., Hulley, M., Spencer, C., & Farid, S. S. (2020). High-throughput Raman spectroscopy combined with innovate data analysis workflow to enhance biopharmaceutical. *Processes*, 8, 1179.
- Guo, S., Popp, J., & Bocklitz, T. (2021). Chemometric analysis in Raman spectroscopy from experimental design to machine learning-based modeling. *Nature Protocols*, 16, 5426–5459.
- Hubbuch, J. (2012). Editorial: High-throughput process development. *Biotechnology Journal*, 7, 1185.
- ICH. (2021). *Continuous manufacturing of drug substances and drug products Q13*. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use.
- ICH. (2022). *Analytical procedure development Q14*. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use.
- Khanal, O., & Lenhoff, A. M. (2021). Developments and opportunities in continuous biopharmaceutical manufacturing. *mAbs*, 13, 1903664.
- McAvan, B. S., Bowsher, L. A., Powell, T., O'Hara, J. F., Spitali, M., Goodacre, R., & Doig, A. J. (2020). Raman spectroscopy to monitor post-translational modifications and degradation in monoclonal antibody therapeutics. *Analytical Chemistry*, 92, 10381–10389.
- Mehmood, T., Liland, K. H., Snipen, L., & Sæbø, S. (2012). A review of variable selection methods in partial least squares regression. *Chemometrics and Intelligent Laboratory Systems*, 118, 62–69.
- Nowak, C., K. Cheung, J., M. Dellatore, S., Katiyar, A., Bhat, R., Sun, J., Ponniah, G., Neill, A., Mason, B., Beck, A., & Liu, H. (2017). Forced degradation of recombinant monoclonal antibodies: A practical guide. *mAbs*, 9, 1217–1230.
- Patel, B. A., Gospodarek, A., Larkin, M., Kenrick, S. A., Haverick, M. A., Tugcu, N., Brower, M. A., & Richardson, D. D. (2018). Multi-angle light scattering as a process analytical technology measuring real-time molecular weight for downstream process control. *mAbs*, 10, 945–950.
- Pedro, M. N. S., Klijjn, M. E., Eppink, M. H., & Ottens, M. (2022). Process analytical technique (PAT) miniaturization for monoclonal antibody aggregate detection in continuous downstream processing. *Journal of Chemical Technology and Biotechnology*, 97(9), 2347–2364.
- Ramakrishna, A., Prathap, V., Maranholkar, V., & Rathore, A. S. (2022). Multi-wavelength UV-based PAT tool for measuring protein concentration. *Journal of Pharmaceutical and Biomedical Analysis*, 207, 114394.
- Rathore, A. S. (2014). QbD/PAT for bioprocessing: Moving from theory to implementation. *Current Opinion in Chemical Engineering*, 6, 1–8.
- Rolinger, L., Rüdtt, 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.
- Rolinger, L., Rüdtt, M., & Hubbuch, J. (2021). Comparison of UV- and Raman-based monitoring of the protein A load phase and evaluation of data fusion by PLS models and CNNs. *Biotechnology and Bioengineering*, 118, 4255–4268.
- Ryabchikov, O., Guo, S., & Bocklitz, T. (2018). Analyzing Raman spectroscopic data. *Phys Sci Rev*, 4, 20170043.
- Saleh, D., Wang, G., Rischawy, F., Kluters, S., Studts, J., & Hubbuch, J. (2021). In silico process characterization for biopharmaceutical development following the quality by design concept. *Biotechnology Progress*, 37, e3196.
- Silva, T. C., Eppink, M., & Ottens, M. (2022). Automation and miniaturization: Enabling tools for fast, high-throughput process development in integrated continuous biomanufacturing. *Journal of Chemical Technology & Biotechnology*, 97, 2365–2375.
- Taraban, M. B., Briggs, K. T., Merkel, P., & Yu, Y. B. (2019). Flow water proton NMR: In-line process analytical technology for continuous biomanufacturing. *Analytical Chemistry*, 91, 13538–13546.
- Tulsyan, A., Garvin, C., & Undey, C. (2019). Industrial batch process monitoring with limited data. *Journal of Process Control*, 77, 114–133.
- Tulsyan, A., Wang, T., Schorner, G., Khodabandehlou, H., Coufal, M., & Undey, C. (2020). Automatic real-time calibration, assessment, and maintenance of generic Raman models for online monitoring of cell culture processes. *Biotechnology and Bioengineering*, 117, 406–416.
- Wei, B., Woon, N., Dai, L., Fish, R., Tai, M., Handagama, W., Yin, A., Sun, J., Maier, A., McDaniel, D., Kadaub, E., Yang, J., Saggi, M., Woys, A., Pester, O., Lambert, D., Pell, A., Hao, Z., Magill, G., ... Chen, Y. (2021). Multi-attribute Raman spectroscopy (MARS) for monitoring product quality attributes in formulated monoclonal antibody therapeutics. *mAbs*, 14, 2007564.
- Yilmaz, D., Mehdizadeh, H., Navarro, D., Shehzad, A., O'Connor, M., & McCormick, P. (2020). Application of Raman spectroscopy in monoclonal antibody producing continuous systems for downstream process intensification. *Biotechnology Progress*, 36, e2947.
- Zhang, C., Springall, J. S., Wang, X., & Barman, I. (2019). Rapid, quantitative determination of aggregation and particle formation for antibody drug conjugate therapeutics with label-free Raman spectroscopy. *Analytica Chimica Acta*, 1081, 138–145.
- Zhang, Z. M., Chen, S., Liang, Y. Z., Liu, Z. X., Zhang, Q. M., Ding, L. X., Ye, F., & Zhou, H. (2010). An intelligent background-correction algorithm for highly fluorescent samples in Raman spectroscopy. *Journal of Raman Spectroscopy*, 41, 659–669.
- Zhou, C., Qi, W., Neil Lewis, E., & Carpenter, J. F. (2015). Concomitant Raman spectroscopy and dynamic light scattering for characterization of therapeutic proteins at high concentrations. *Analytical Biochemistry*, 472, 7–20.
- Zhu, G., Zhu, X., Fan, Q., & Wan, X. (2011). Raman spectra of amino acids and their aqueous solutions. *Spectrochimica Acta, Part A: Molecular and Biomolecular Spectroscopy*, 78, 1187–1195.

**How to cite this article:** Wang, J., Chen, J., Studts, J., & Wang, G. (2023). Automated calibration and in-line measurement of product quality during therapeutic monoclonal antibody purification using Raman spectroscopy. *Biotechnology and Bioengineering*, 1–11. <https://doi.org/10.1002/bit.28514>