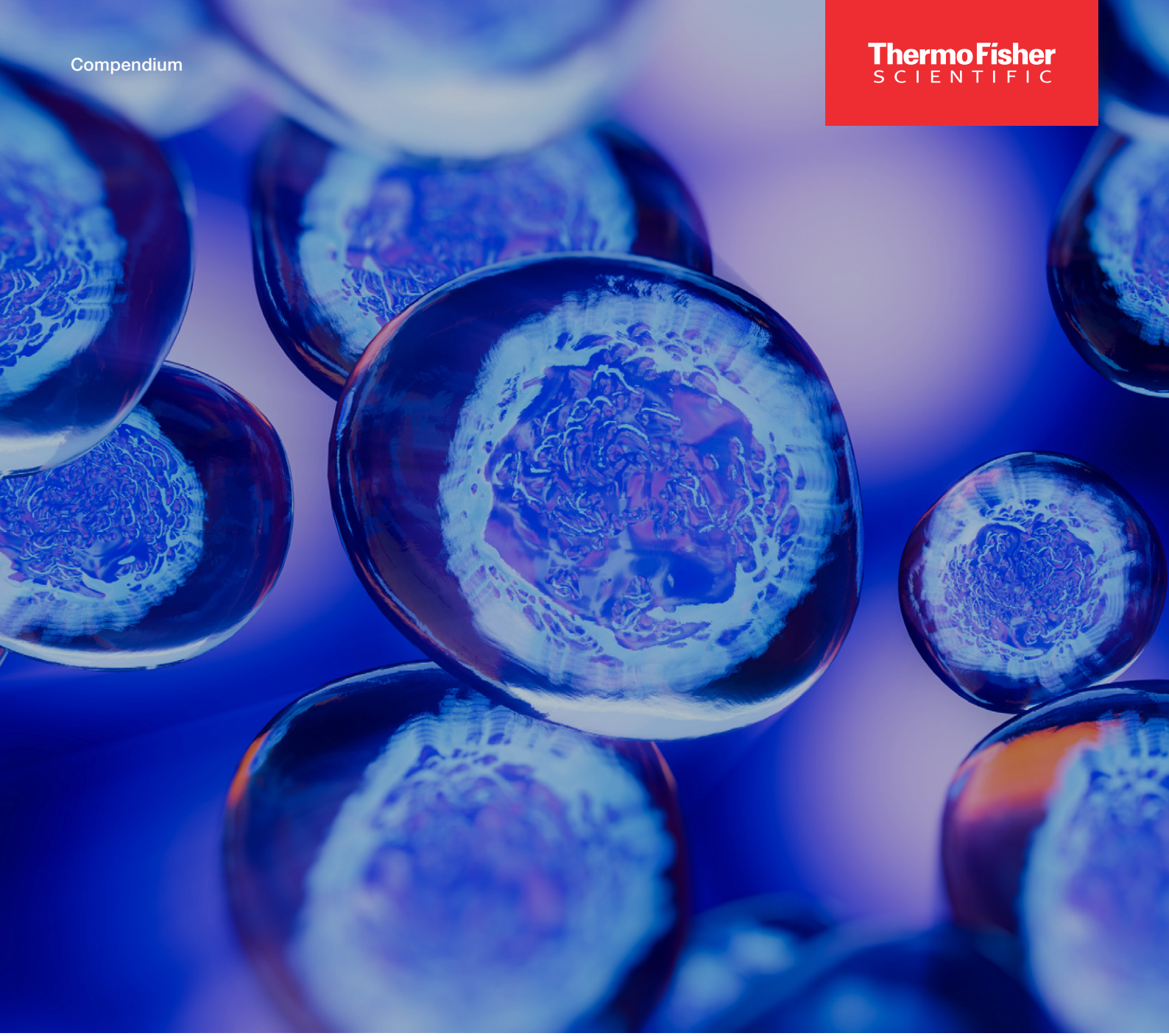
Compendium

ThermoFisher SCIENTIFIC



BioPharmaceutical approach with spectroscopy

thermo scientific

Summary

Heavily-regulated biopharmaceutical manufacturers are increasing their use of molecular spectroscopy techniques, including both vibrational spectroscopy and UV-Visible spectrophotometry. These analytical methods include the use of mid-infrared (MIR), near infrared (NIR), Fourier transform infrared (FTIR), and Raman spectroscopy, along with UV-Vis absorption spectroscopy. All of these techniques provide rapid, accurate analysis capabilities and they are often complementary in nature.

Vibrational spectroscopy plays a major role for biopharmaceutical analysis in upstream, downstream, and fill-finish processes. To support upstream processes, infrared and Raman techniques can be utilized when performing multi-attribute raw material testing. Examples shown in this compendium include the use of FT-NIR to predict protein concentrations in cell cultures, and the application of FTIR to elucidate protein secondary structures.

Ultraviolet-visible (UV-Vis) spectrophotometry has proven to be a very useful method for identifying and quantifying proteins, peptides, and nucleic acids. During upstream stages of biopharma production, UV-Vis is a useful technique for early-stage analyses of raw materials or starter cultures. This is highlighted in the included articles about protein aggregation as identified through UV-Vis absorption spectroscopy, and evaluation of DNA purity for quality control during molecular cloning.

During downstream processes, Raman spectroscopy has become especially useful as a process analytical technology (PAT). Multiple articles towards the latter end of this compendium attest to this. The technique provides rapid feedback that enables immediate process control to enhance the manufacturing of biologics. Process Raman has the potential to be used as a real-time guide to downstream buffer control.

From raw material identification or purity testing in upstream processes, to quality control during the biopharmaceutical production process, and downstream process analysis, UV-Vis and vibrational spectroscopy can provide critical information for biopharmaceutical workflows.

Table of contents

Protein secondary structure elucidation using FTIR spectroscopy 4

Process Raman as a platform solution for automated glucose feeding in fed-batch bioreactors 8

Introduction to method transfer between Raman spectrometers 19

Protein concentration prediction in cell cultures 24

Protein aggregation identified through UV-Visible absorption spectroscopy 27

Evaluating DNA Purity for Molecular Cloning Quality Control 32

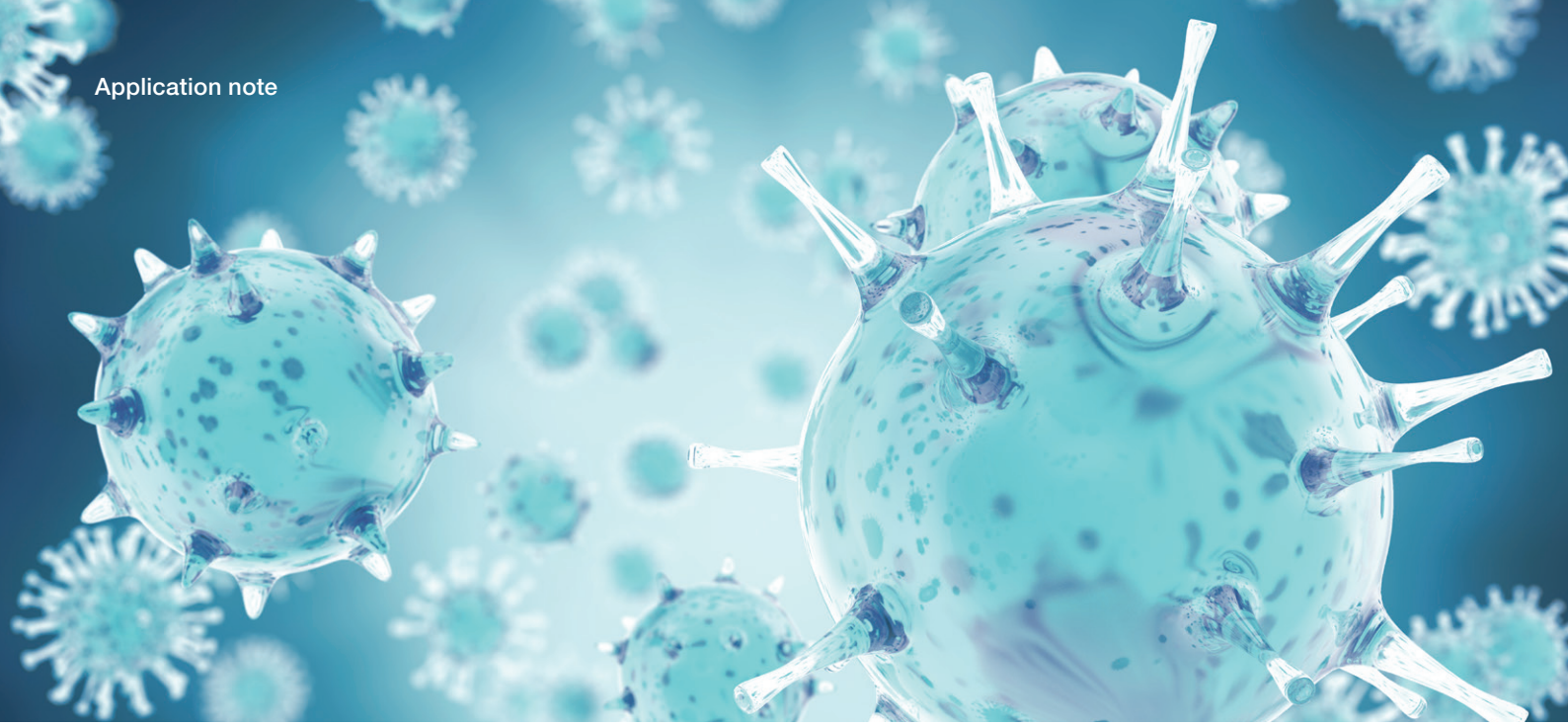
Quantify protein and peptide preparations at 205 nm 36

Observation of gold nanoshell plasmon resonance shifts after bioconjugation 39

The NanoDrop Eight Spectrophotometer detects contaminating nucleic acids in mammalian DNA and RNA preparations 42

Enabling real-time release of final products in manufacturing of biologics 45

Process Raman as a comprehensive solution for downstream buffer workflow 52



Application note

Protein secondary structure elucidation using FTIR spectroscopy

Author

Suja Sukumaran

Thermo Fisher Scientific, USA

Keywords

FTIR, ATR, protein structure elucidation, Biocell calcium fluoride cell, ConcentrateIR2 ATR, transmission

Abstract

Fourier-transform infrared (FTIR) spectroscopy is one of the most versatile analytical tools used across various disciplines. In this study, the Thermo Scientific™ Nicolet™ iS10 and Nicolet™ iS50 FTIR Spectrometers, equipped with attenuated total reflection (ATR) FTIR and transmission FTIR, were used for the determination of protein secondary structures. Structure calculations based on a protein database as well as spectral deconvolution are discussed. The analyses were quick and easy.

Introduction

Protein secondary structure describes the repetitive conformations of proteins and peptides. There are two major forms of secondary structure, the α-helix and β-sheet, so named for the patterns of hydrogen bonds between amine hydrogen and carbonyl oxygen atoms that create the peptide backbone of a protein.[1] Understanding protein secondary structure is important to gain insight into protein conformation and stability. For example, temperature dependent analysis of the secondary structure is critical in determining storage conditions for maintaining active therapeutic proteins.[2] Protein secondary structure is also crucial in understanding the structure–function relationship and enzyme kinetics of various proteins.[3]

FTIR has long been established as a powerful analytical technique to investigate protein secondary structure and local conformational changes.[1, 4] A typical protein infrared (IR) spectrum often contains nine amide bands, with vibrational contributions from both protein backbone and amino acid side chains. Among which, of particular pertinence to protein secondary structure are amide I and amide II bands. The absorptions associated with C=O stretching are denoted as amide I, whereas those associated with N–H bending are amide II. Since both C=O and N–H bonds are involved in the hydrogen bonding between different moieties of secondary structure, the positions of both amide I and amide II bands are sensitive to the secondary structure composition of a protein,[3, 4] although the amide II band is widely viewed as a less useful predictor for quantifying the secondary structure of proteins.

The shifts in the amide I band are often small compared to the intrinsic width of the band, resulting in one broad peak instead of a series of resolved peaks for each type of the secondary structure. Mathematical procedures such as Fourier self-deconvolution and second derivatives can be used to resolve the overlapping bands for the quantitative analysis of protein secondary structure.[3] Table 1 shows the secondary structure band assignments for proteins in water. Note that all assignments are depicted as a range, as the exact position of each peak varies from protein to protein due to the differences in hydrogen bonding interactions and the environment of the proteins.

Secondary structure	Band assignment in water
α-Helix	1,648–1,657 cm^{-1}
β-Sheet (high-frequency component)	1,623–1,641 cm^{-1} 1,674–1,695 cm^{-1}
Random	1,642–1,657 cm^{-1}
Coils	1,662–1,686 cm^{-1}

Table 1. Secondary structure band assignments for protein in water.[2]

With a range of sampling techniques, including transmission, ATR, and infrared reflection absorption spectroscopy (IRRAS), FTIR is particularly advantageous in terms of its versatility and general applicability compared to other analytical techniques for protein secondary structure analysis. Protein sample forms suitable for FTIR analysis include lyophilized powders, water solution, and colloids, to name a few. We report herein two examples of protein secondary structure determination using transmission FTIR and ATR, respectively. Both methods are fast, consume a minute amount of sample, and require minimal sample preparation.

Experiment

All proteins were procured from Sigma-Aldrich (MO, USA) and used as received. For the transmission studies, a BioCell™ Calcium Fluoride Cell (Biotools, Jupiter, FL) was used, and all measurements were carried out at ambient temperature. A 10 µL protein solution was placed at the center of the window, and the protein solution was sandwiched between the two CaF_2 windows, and placed in the holder. The concentration of protein tested was between 6 and 12 mg/mL. A 6 µm path length was created by sandwiching the two CaF_2 windows. CaF_2 windows are suited for water-based sample analysis. As water has a significant absorption peak at 1,645 cm^{-1} region, a small path length of 6 µm can effectively avoid saturated water peaks.

A purged Nicolet iS10 FTIR Spectrometer, equipped with a DTGS detector, was used for transmission analysis. The scan parameters used were 256 scans with a resolution of 4 cm^{-1}. The Thermo Scientific™ Smart OMNI-Transmission™ Accessory allows for a quick purge of the chamber, eliminating the need for water vapor subtraction in most analyses. Secondary structure analysis of the buffer-subtracted spectra was carried out using the built-in feature of the PROTA-3S™ FT-IR Protein Structure Analysis Software. Secondary structure calculation in PROTA-3S software is based on a database of 47 secondary structures (for more information visit **www.btools.com**).

For ATR analysis, a ConcentratIR2™ Multiple Reflection ATR Accessory (Harrick Scientific Products, Inc. Pleasantville, NY) with diamond crystal was used in a Nicolet iS50 FTIR spectrometer equipped with a mercuric cadmium telluride (MCT) detector. The diamond ATR has ten internal reflections with a nominal angle of incidence of 45 degrees. A 10 µL protein solution in phosphate buffer was dried on the surface of the ATR crystal under a stream of nitrogen. Scan parameters used were 256 scans and a resolution of 4 cm^{-1}. Secondary structure determination was carried out using the peak resolve feature of the Thermo Scientific™ OMNIC™ Software.

Results and discussion

Transmission-FTIR with Bio Cell

Figure 1 shows the overlay of three FTIR spectra: phosphate buffer, cytochrome C at 6 mg/mL and 12 mg/mL in phosphate buffer, respectively. At first glance, the spectra are predominantly water bands. The three spectra show little difference, even at a high protein concentration of 12 mg/mL.

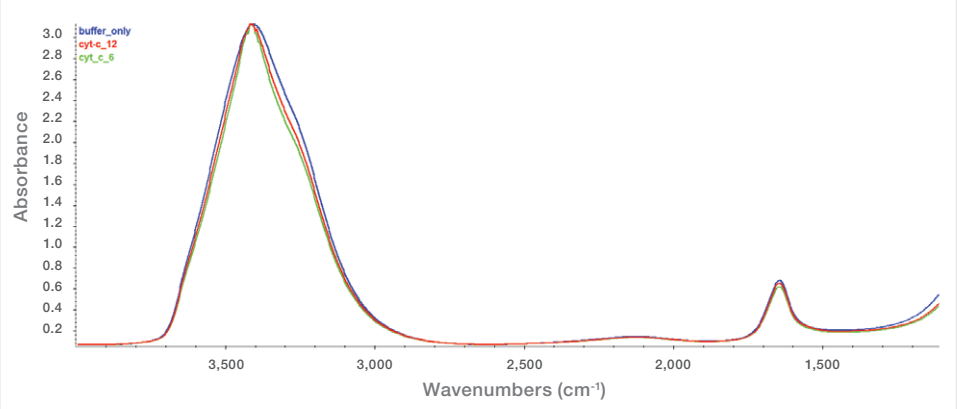


Figure 1. Transmission-FTIR spectra for cytochrome C in phosphate buffer (cytc_12) at 12 mg/mL and 6 mg/mL (cytc_6), and phosphate buffer blank.

Next, the buffer spectrum was subtracted from the raw protein spectra using the PROTA-3S software, and the results are shown in Figures 2A (cytochrome C) and 2B (concanavalin). The amide I and II peaks are clearly discernible in both spectra. The amide I peak position for cytochrome C spectra is 1,654 cm^{-1}, suggesting an α-helix dominant secondary structure. For concanavalin A, the amide I peak centers at 1,633 cm^{-1}, and there is also a noticeable shoulder peak at 1,690 cm^{-1} (red circle), indicative of the β-sheet component and its associated high-frequency component.[2]

Table 2 summarizes the secondary structure prediction using the PROTA-3S software. The cytochrome C has 45% α-helix and 5% β-sheet, whereas concanavalin A has 42% β-sheet and 4% α-helix. Differences in secondary structure composition between X-ray and FTIR data are likely due to the physicochemical state of the protein samples such as crystalline versus solution, temperature, pH, buffer conditions, etc. Furthermore, different prediction algorithms could have slightly varying outputs.[7] Notwithstanding the differences in analytical technique, sample state, and prediction algorithms, the secondary structure elucidation by FTIR using PROTA-3S software is largely in line with that from X-ray. Transmission-FTIR measurements combined with PROTA-3S software offer a facile and fast means to analyze the secondary structure of proteins in solution[2, 3] with minimal sample prep.

ATR-FTIR with ConcentratIR2 Accessory

When the quantity and concentration of protein are limited, FTIR measurements with the ConcentratIR2 Multiple Reflection ATR offer a better alternative than transmission-FTIR spectroscopy. The unique design of this ATR accessory allows for the direct measurement of protein powders, gels, solutions as well as proteins dried on the ATR surface. When concentrating proteins on the crystal surface, caution should be exercised in buffer selection since buffer will also concentrate on the surface of the crystal.

Only those buffers with minimum or no peaks in the amide I and II region should be selected. Figure 3 shows the ATR-FTIR spectra of BSA in phosphate buffer, dried on the crystal from a 1 mg/mL solution. In addition to the amide I and II bands, there are spectral features of the side chain, such as 1,515 cm^{-1} from tyrosine and 1,498 cm^{-1} from aspartic acid. Side chain peaks are critical for the elucidation of protonation and de-protonation states of various amino acids.[2]

	α-Helix (%)		β-Sheet (%)		Random (%)	
Protein	FTIR	X-ray	FTIR	X-ray	FTIR	X-ray
Cytochrome C	45	41	5	0	50	59
Concanavalin A	4	0	42	48	54	52

Table 2. Comparison of secondary structure calculation from FTIR (PROTA-3S) and X-ray data.

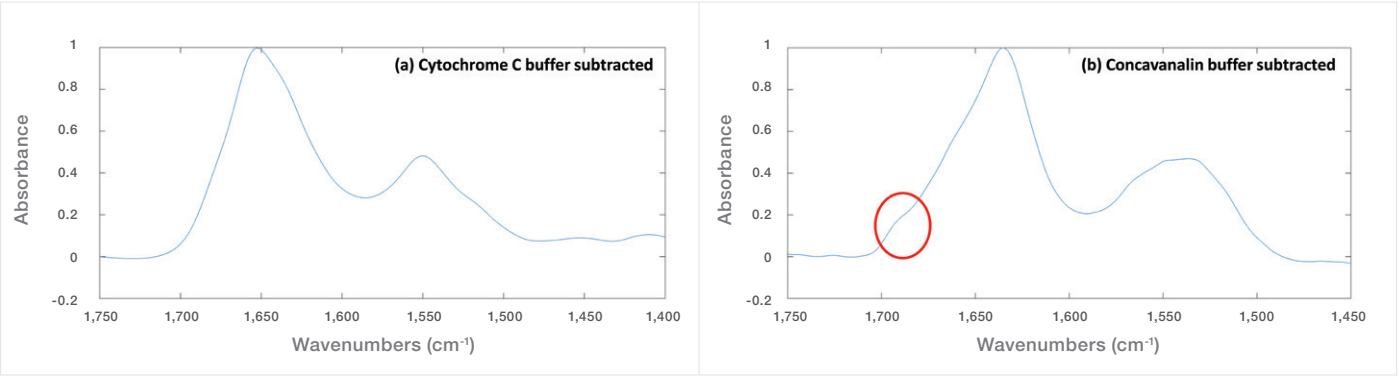


Figure 2. FTIR spectra of (a) cytochrome C and (b) concanavalin A after the buffer spectrum was subtracted using PROTA-3S software.

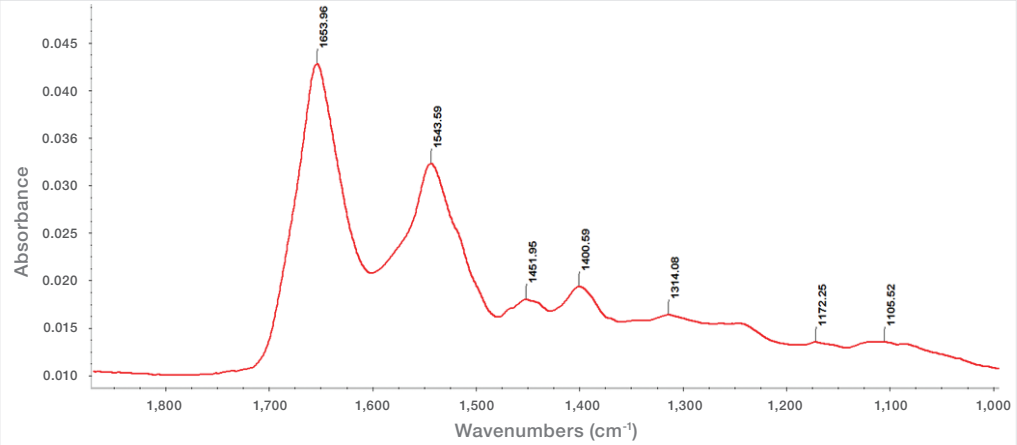


Figure 3. Amide I and II for 1 mg/mL BSA analyzed using ConcentratIR2 ATR on the Nicolet iS50 FTIR Spectrometer equipped with an MCT detector.

Peak deconvolution of the amide I peak (Figure 4) of BSA was carried out using the OMNIC software. It is important to note that second derivative analysis is often performed prior to deconvolution to clearly identify the peaks required for peak fitting.[2] In the current study, the second derivative peaks obtained (results not shown) are well correlated to the secondary structure peak assignments in Table 1. In order to obtain a good peak shape for peak fitting, a baseline correction on the amide I region was also performed. Baseline correction also effectively excluded the contributions from the amide II region. The deconvolution of amide I resulted in 5 peaks, and the area under each peak was then evaluated against the total area. Amide I peak deconvolution shows a secondary structure composition of 47% α-helix, 3% β-sheet, 24% coils, and 26% random, which is to published FTIR5 and X-ray data.

Conclusion

In this note, we have demonstrated two examples of protein secondary structure elucidation using FTIR spectroscopy. Transmission-FTIR measurements combined with PROTA-3S software provides a facile means to analyze secondary structure of proteins in solution with minimal sample preparation. When the quantity and concentration of protein are limited, ATR-FTIR offers a better alternative by drying the proteins in ATR crystals directly. The data were collected using an older model, the Nicolet iS10 Spectrometer. An improved model, the Nicolet iS20 Spectrometer, offers superior speed and performance over this predecessor model.

References

1. Elliott, A., Ambrose, E. J. **Structure of synthetic polypeptides**, Nature (1950) 165, 921-922.
2. Jackson, M., Mantsch, H.H. **The use and misuse of FTIR spectroscopy in the determination of protein structure**, Crit. Rev. Biochem. Mol. Biol. (1995) 30, 95-120.
3. Barth, A. **Infrared spectroscopy of proteins**, Biochim. Biophys. Acta (2007) 1767, 1073-1101.
4. Byler, D.M., Susi, H. **Examination of the secondary structure of proteins by deconvolved FTIR spectra**, Biopolymers (1986) 25, 469-487.
5. Surewicz, W.K., Mantsch, H.H. **New insight into protein secondary structure from resolution-enhanced infrared spectra**, Biochim. Biophys. Acta (1988) 952, 115-130.
6. Sukumaran, S., Hauser, K., Maier, E., Benz, R., Mantele, W. **Tracking the unfolding and refolding pathways of outer membrane protein porin from Paracoccus denitrificans**, Biochemistry (2006) 45, 3972–3980.
7. Klose, D., Janes R.W. 2Struc – **the protein secondary structure analysis server**, Biophysical Journal (2010) 98, 454-455.

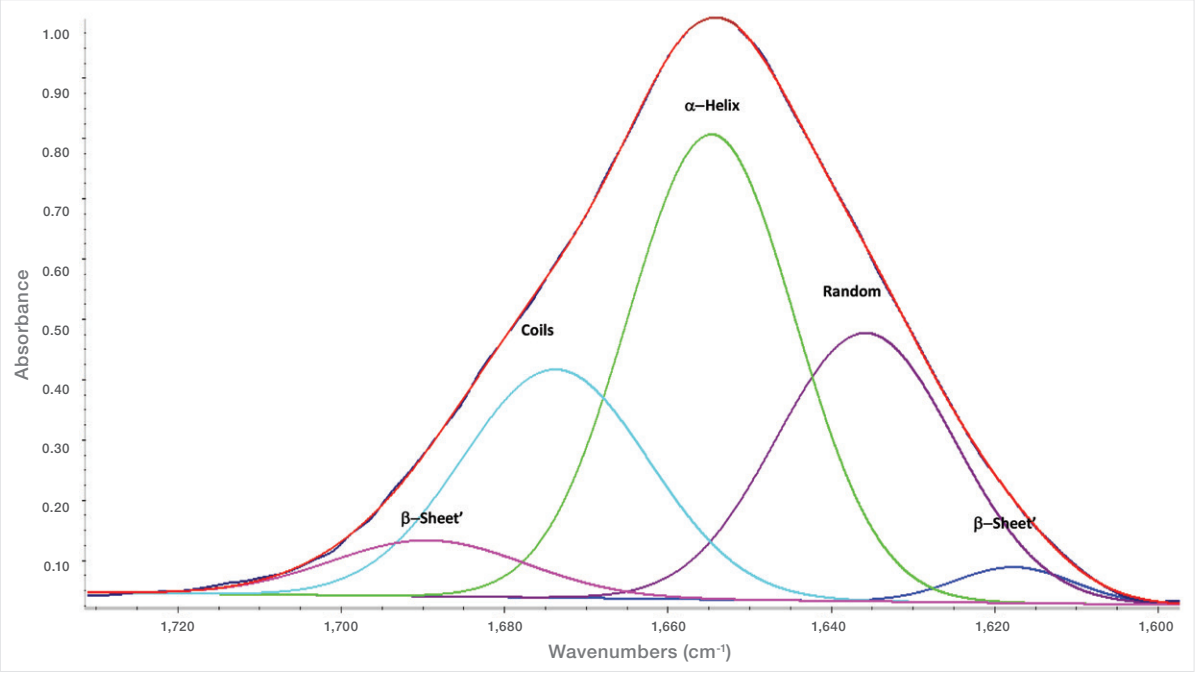


Figure 4. Peak deconvolution of amide I peak of BSA using Peak Resolve function of OMNIC software.

Application note

Process Raman as platform solution for automated glucose feeding in fed-batch bioreactors

Authors

Juan Villa[1], Ph.D., Bioprocess Scientist

Nimesh Khadka[2], Ph.D.,
Senior Application Specialist

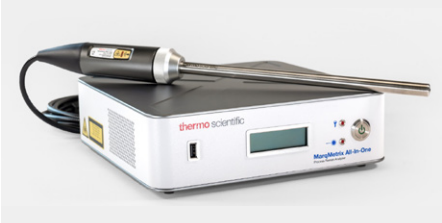
Kevin Keck[1], System Design Engineer

Lin Zhang[2], Ph.D., Manager of Algorithm and Chemometric Development

Matthew Zustiak[1], Ph.D., Director of Bioproduction Collaboration Center

[1]BioProduction, Thermo Fisher Scientific, St. Louis, Missouri USA

[2]Analytical Instrument,
Thermo Fisher Scientific,
Tewksbury, Massachusetts USA



Thermo Scientific MarqMetrix All-In-One Process Raman Analyzer and Thermo Scientific MarqMetrix Performance BallProbe Sampling Optic.

Industry/application

Biopharma PAT / Upstream Bioreactor

Products used

Thermo Scientific™ MarqMetrix™ All-In-One Process Raman Analyzer,
Thermo Scientific™ MarqMetrix™ Performance BallProbe™ Sampling Optic,
Thermo Scientific™ TruBio™ software

Premise and goals

Fed-batch bioreactors are widely used in the biopharmaceutical industry, from process development to manufacturing. Many drugs are currently produced using the fed-batch feeding strategy. The timing and amount of bolus glucose feeding in these processes are determined by manual sampling, offline analytics, and pre-established mathematical relationships.

In this study, we aimed to enhance the efficiency of fed-batch bioreactors by integrating in-line process Raman spectroscopy as a process analytical technology (PAT) tool. Leveraging the feedback control from the process Raman, we automated bolus glucose feeding without altering the existing workflow. The automated glucose feeding was performed once a day at a specified time to reach the defined concentration, mimicking the conventional manual fed-batch approach currently in practice.

Key analytes

Glucose feeding in fed-batch bioreactor

Key benefits

- Efficiency is improved by bringing automation to the widely used fed-batch bioreactor operation, with cost benefits achieved through tighter process control, elimination of the need for laboratory analytics, and improved product repeatability and quality.
- Automation increases the reliability of process control by reducing dependence on analysts and minimizing human errors.
- This approach enables the immediate integration of process Raman into already existing processes with monitoring and control capabilities.

Fed-batch bioreactors are widely used in the biopharmaceutical industry for manufacturing various drugs, including monoclonal antibodies.[1] In a fed-batch bioreactor, glucose is added as a carbon and energy source in defined amounts and at specific times, based on data from laboratory analytics. The fed glucose is consumed by growing cells. When the glucose level falls below a set threshold, the next batch of glucose is added. This fed-batch process continues until the end of the bioreactor run.

Although widely used, the conventional fed-batch strategy has several aspects that can be improved to enhance efficiency and reduce costs. Currently, analysts periodically monitor the bioreactor and perform laboratory analytics to make feeding decisions and to track the progress of the run. These manual processes are time-consuming and prone to human error, leading to inconsistencies in product quality and increased costs. For instance, unexpected rapid cell growth during periods when analysts are not present can result in a sudden depletion of glucose, causing batch failure. Similarly, accidental loading of high glucose concentrations during feeding can lead to the production of unwanted glycated products. To prevent such failures while maintaining all the benefits of the established fed-batch process, in-line Raman spectroscopy offers a viable solution. The analytical technique provides real-time measurements to aid automatic feedback, allowing for tighter process control.

Previously we and others have demonstrated process Raman as a reliable tool to maintain constant glucose concentration in a fully automated fashion throughout the bioreactor run.[2,3] In this study (see Figure 1), we showcase the human-free operation of fed-batch bioreactors by automating the bolus feeding of glucose using feedback control from the in-line Thermo Scientific™ MarqMetrix™ All-In-One Process Raman Analyzer. The automation workflow mimicked the traditional fed-batch process by feeding glucose once a day at a specified time and dosing to a predefined concentration. To compare the performance of automated and manual fed-batch bioreactors, we evaluated the production of lactate profile, the total quantity of titer (product), and the quality attributes of the titer. Lactate is primarily produced during the incomplete oxidation of glucose and is often used as a marker for measuring oxygen availability, which is directly related to cell health.

Our automated approach enhances process efficiency by eliminating the need for laboratory analytics, resulting in significant cost benefits. Automation ensures product repeatability and quality by providing consistent and accurate feeding decisions in real time. Additionally, it reduces dependence on analysts and minimizes human errors, making process control more reliable. The integration of Raman technology into existing workflows offers advanced multimodal monitoring and control capabilities, paving the way for immediate adoption in biopharmaceutical manufacturing.

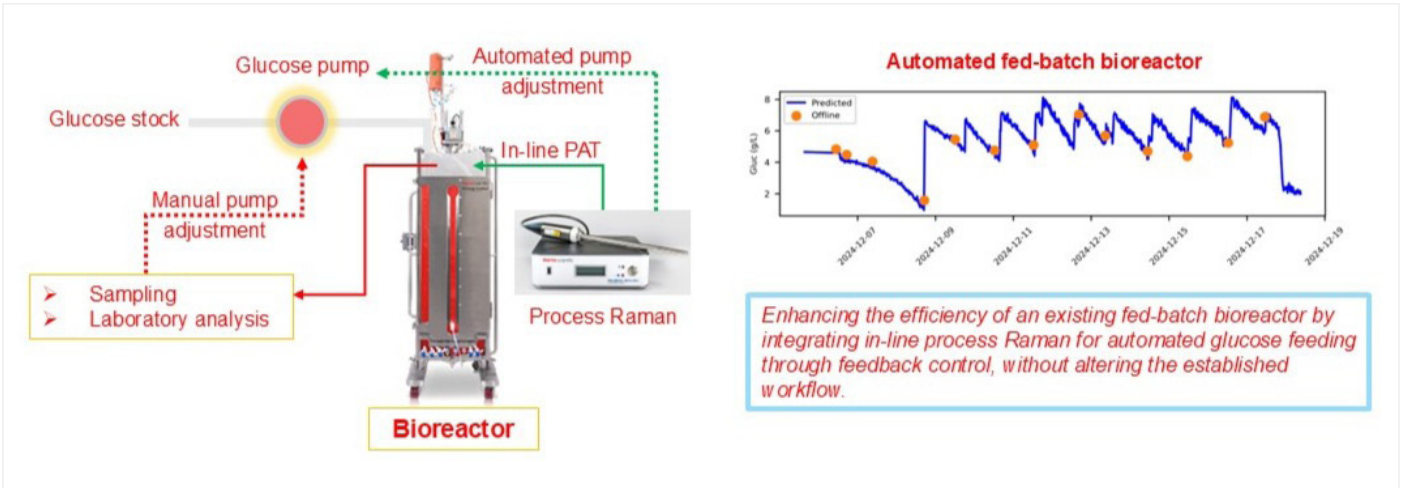


Figure 1. Showing how the manual glucose pump adjustment (red dotted line) workflow in a conventional bioreactor was replaced by automation using in-line process Raman (green dotted line).

Materials and methods

Cell culture

miCHO(TM) CHO-K1 cells (ATUM) were inoculated in Gibco™ Efficient-Pro™ medium supplemented with 1.5 mg/L insulin and 1% anticlumping agent at a density of 0.75 million cells/mL in a 5L glass bioreactor. The bioreactor was operated at a temperature of 37°C, pH 7 ± 0.2, and dissolved oxygen (DO) maintained at 40%. The pH was controlled by the addition of CO_2 for high pH and sodium carbonate for low pH adjustments. The cells were grown in a medium specialized for mammalian cells, in standard fed-batch bioreactors, in duplicate, using a Thermo Fisher Scientific platform process for 12 days. The cells were grown in chemically defined medium and fed daily starting from day 3 with a two-step feeding process using Gibco™ Efficient-Pro™ feed 2 and Enhancer.

Raman data acquisition

Real-time Raman data was acquired using the Thermo Scientific MarqMetrix All-In-One Process Raman Analyzer, integrated with the Thermo Scientific™ MarqMetrix™ Performance BallProbe™ Sampling Optic. The acquisition parameters were set to a power of 450 mW, an integration time of 3000 ms, and an average of 20 scans.

Chemometric model development

Chemometric models were developed using Raman training data collected from previous bioreactors with the same acquisition parameters. The partial least squares (PLS) algorithm was used to develop the glucose and lactate models.

Three spectral regions were selected for the glucose regression model: 1065-1232 cm^{-1}, 1595-1863 cm^{-1}, and 2704-3078 cm^{-1}. The spectral region of 1065-1232 cm^{-1} includes the characteristic Raman peak of glucose at ~1125 cm^{-1}; this peak is attributed to the stretching vibrational modes of CO and CC and the in-plane bending of COH bonds (ν(CO), ν(CC), β(COH)). The spectral region of 1595-1863 cm^{-1} is associated with the symmetric bending of water molecules, while the spectral region of 2704-3078 cm^{-1} includes Raman peaks assigned to the symmetric and antisymmetric stretching vibration modes of CH_2 and CH bonds of biomolecules.

The selected spectral regions were preprocessed in the following order:

1. *Savitzky-Golay (Sav-Gol) filter (1st derivative, order = 2, window width = 13)*
2. *Standard Normal Variate (SNV)*
3. *Mean centering*

The Savitzky-Golay filter removed unwanted baseline information; SNV normalized all spectra to have a mean of zero and a standard deviation of one; and mean centering removed the mean feature from all spectra.

To minimize overfitting, a leave-one-out cross-validation (LOOCV) strategy was used such that each dataset for a given bioreactor run was left out once during the cross validation. The root means squared error of cross-validation (RMSECV) was calculated and used to determine the appropriate number of latent variables (LVs). The optimum LVs—those that minimized the root mean square error of calibration and cross validation while maintaining their ratio close to 1—were selected.

The lactate model was developed using a similar strategy, except a broader spectral region of 800 to 1750 cm^{-1} was used and SNV was substituted with L1 norm as shown in Table 1.

Analyte	Model Type	Region Selection cm^{-1}	Preprocessing
Glucose	PLS	1065-1232; 1595-1863; 2704-3078	Sav-Gol filter (1st Derivative; order = 2; Window width = 13) + SNV + Mean Center
Lactate	PLS	800 - 1750	Sav-Gol filter (1st Derivative; order = 2; Window width = 11) + L1 Norm (Area = 1 for 1540-1750 cm^{-1}) + Mean Center

Table 1. Spectral regions and preprocessing used for model development.

All data management, cosmic ray removal, averaging, and timestamp alignment were performed in an internally developed Python platform. The data were then processed in Python as well as a commercially available software package SOLO 9.3.1 (2024, Eigenvector Research. Inc. Manson, WA USA 98831).

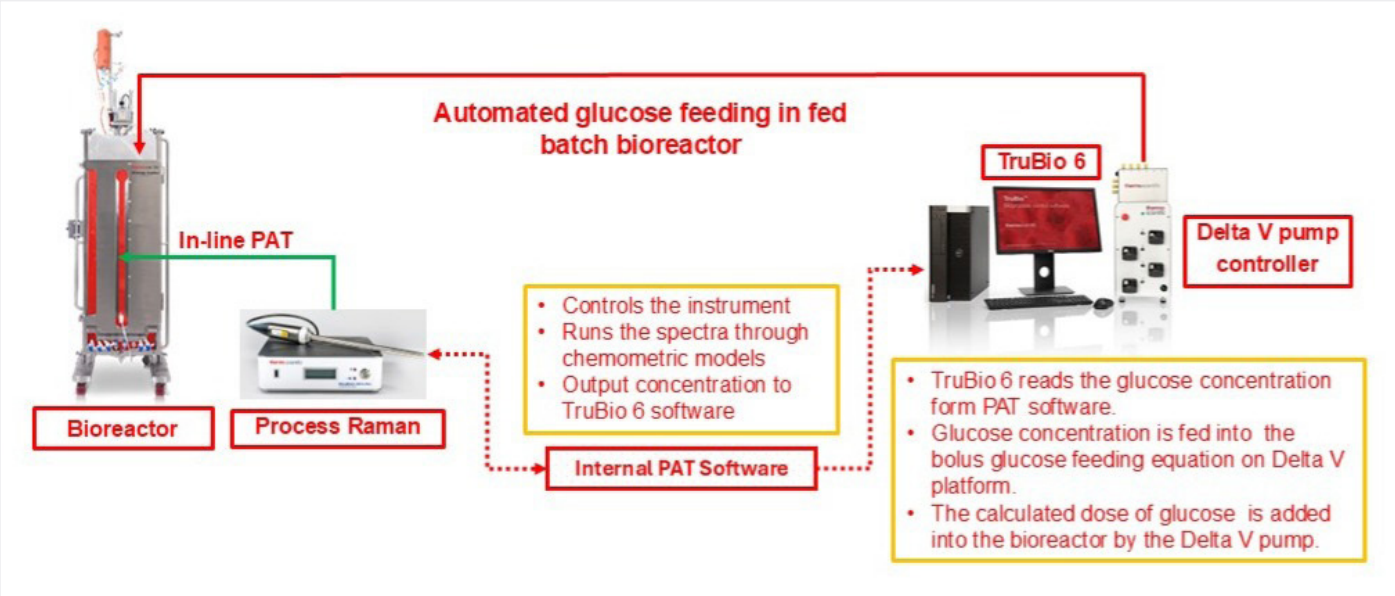


Figure 2. Schematics of workflow for automated feedback control for feeding glucose in a fed-batch bioreactor.

Control strategy

The internally developed PAT software was used to manage the process Raman and mathematically interpret the spectral information into concentration by running it through the chemometric models. The predicted glucose concentrations were communicated to the Thermo Scientific™ TruBio™ software, which in turn controlled the glucose feeding into the bioreactor via the Delta V pump. A simplified schematic is shown in Figure 2.

An automated bolus feeding control strategy was programmed into Finesse Solutions TruBio v 6.0.0 through the logic functions. Comparison statements were first utilized to feed into a logical statement to evaluate multiple criteria simultaneously. (All conditions must be satisfied before the action of turning on a peristaltic pump attached to a G3Lab Universal Controller is performed.)

Using a timer and a sine function, a periodic output was established where the sine function's value exceeded a defined threshold and remained above that threshold for a period of 30 minutes once in a 24-hour period. This can be achieved in many ways, however, the values chosen in this example can be seen in Equation 1. The input variable, x, is a timer that was selected to count in minutes and the constant inside the sine function was chosen such that the period was 24 hours. The constant in front of the sine function was selected to give sufficient precision in the duration above the threshold. In this case that threshold was 991,444 and was set as the first conditional statement.

$$f(x) = 1{,}000{,}000 \cdot \sin\left(\frac{720}{1{,}440} \cdot x\right)$$

Equation 1.

The second condition that needs to be met is for the glucose concentration to be less than a specified value, in this case 4 g/L. A communication protocol enables glucose concentration readings from probe measurements through the MarqMetrix All-In-One Process Raman Analyzer to be sent into TruBio through the PAT software.

Results and discussion

The models were developed in the SOLO software and exported in .xml format to run in the PAT software. The statistics for the developed glucose and lactate PLS model are summarized in Table 2. The models were also replicated in the Python platform, and the results are discussed below.
The details of the model are described below separately in the supplementary section.

Model parameters	Glucose model	Lactate model
Model Range	0-12 g/L	0-12 g/L
Number of Latent Variables	5	5
RMSEC	0.43 g/L	0.23 g/L
RMSECV	0.49 g/L	0.31 g/L
R^2 CV	0.94	0.92

Table 2. Model statistics for glucose and lactate.

The real-time control of glucose feeding in the fed-batch bioreactor is shown in Figure 3a. The control strategy was established at the start of the bioreactor run with the following logic:

1. Based on historical data, glucose concentration was sufficient for the first two days, so no feedback control was active.
2. From day 3 onward, when the predicted glucose concentration from the in-line Raman reached 4 g/L or below, bolus glucose feeding was triggered to achieve a final concentration of 7.5 g/L.
3. If the predicted glucose concentration was above 4 g/L, the predicted glucose concentration from the in-line Raman was fed into the feedback control logic at a specified time (9:00 AM) to initiate glucose feeding, aiming for a final concentration of 7.5 g/L.

Note, only one glucose feeding was performed per day to mimic the conventional fed-batch run.

In Figure 3a, the blue trace represents real-time glucose predictions from the in-line process Raman, averaged over a 10-minute window. The orange dots show pooled samples for offline analytics, used as reference values to calculate the root mean square error of predictions (RMSEP). The initial glucose concentration in the feed media was about 5 g/L. As the cells grew, glucose was consumed, dropping the concentration to 1.5 g/L on day 3, when the control loop was activated.

As demonstrated in Figure 3a, the fed-batch bioreactor run was fully automated and maintained glucose levels between 4 and 7.5 g/L (dotted red line), with an overall RMSEP of 0.45 g/L. Note, on some days the glucose level spiked above 7.5 g/L. This error is within the expected prediction error of the model as indicated by the RMSECV value of ~ 0.50 g/L. These errors can be minimized using appropriate statistical process control logic, although such process is beyond the scope of this work. Nonetheless, the results clearly demonstrate the feasibility of automating glucose feeding in a fed-batch bioreactor using feedback control from in-line process Raman. The green dotted line represents the end of the bioreactor run.

Additionally, lactate production was monitored in real time for the automated fed-batch bioreactor using in-line process Raman. The real-time prediction of lactate, averaged over 10 minute segments, is shown as a blue trace in Figure 3b. The orange dots represent samples pooled for offline analytics. The RMSEP of lactate prediction was 0.24 g/L. Although not shown here, the profile of lactate production in the fed-batch bioreactor with and without automated glucose feeding was similar, indicating no alteration in the process when the manual steps were replaced with the integration of in-line process Raman for automated glucose feeding.

Finally, *titer production* was compared using offline analytics for the fed-batch bioreactor with automated glucose feeding and the conventional approach. Titer production was approximately 8 g/L in both modalities. The *protein qualities* from both processes were also assessed using liquid chromatography coupled with mass spectrometry (LC-MS), and no significant differences were observed.

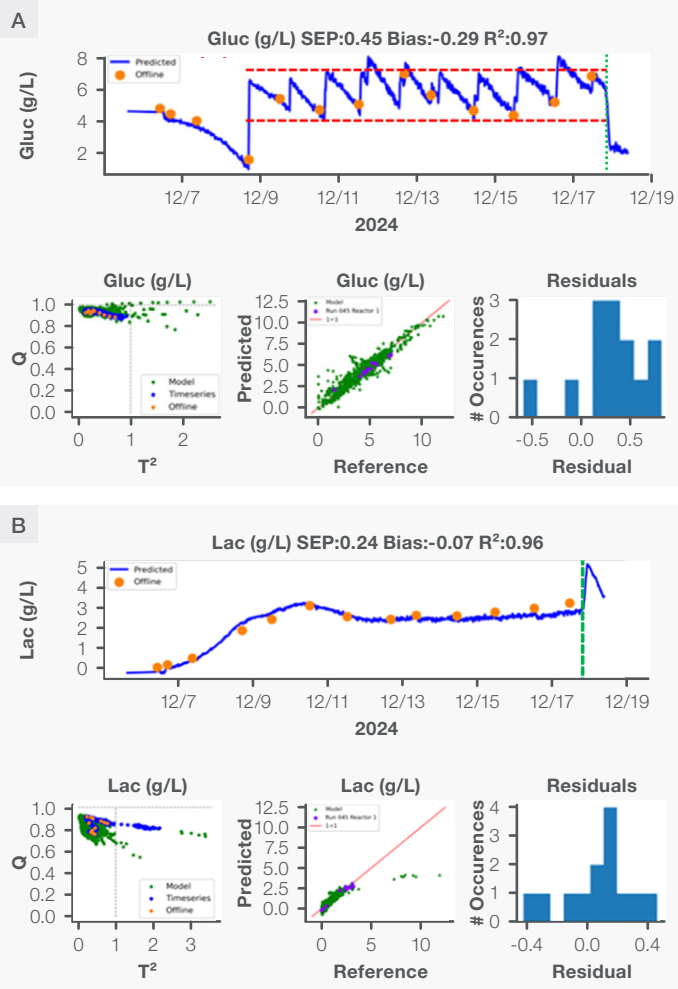


Figure 3. Real-time predictions (blue traces) of glucose (a) and lactate (b) using process Raman, and their correlation to offline reference values (orange dots). Figure (a) demonstrates the human-free glucose feeding for 14 days in a fed-batch bioreactor within the specific limit of 4 to 7.5 g/L (dotted red lines).

Conclusion

Process Raman is a reliable tool that provides a viable option to enhance efficiency in existing workflows by automating glucose feeding in conventional fed-batch bioreactors. The integration of in-line process Raman for automation is unlikely to disrupt established processes, as evidenced by the similar lactate production profile, equivalent titer production, and no significant differences in titer quality. Instead, it offers cost and time benefits, tighter process control, minimization of human errors, and assurance of process reproducibility. This study provides proof of concept, with the aim of encouraging bioprocess scientists and engineers to consider integrating process Raman into their existing fed-batch bioreactors to achieve greater efficiency through automation.

References

4. Yang, J.-D.; Lu, C.; Stasny, B.; Henley, J.; Guinto, W.; Gonzalez, C.; Gleason, J.; Fung, M.; Collopy, B.; Benjamino, M.; Gangi, J.; Hanson, M.; Ille, E. Fed-Batch Bioreactor Process Scale-up from 3-L to 2,500-L Scale for Monoclonal Antibody Production from Cell Culture. *Biotechnol. Bioeng.* **2007**, *98* (1), 141–154. https://doi.org/10.1002/bit.21413.
5. Villa, J.; Zustiak, M.; Kuntz, D.; Zhang, L.; Khadka, N.; Broadbelt, K.; Woods, S. Use of Lykos and TruBio Software Programs for Automated Feedback Control to Monitor and Maintain Glucose Concentrations in Real Time.
6. Zhang, A.; Tsang, V. L.; Moore, B.; Shen, V.; Huang, Y.-M.; Kshirsagar, R.; Ryll, T. Advanced Process Monitoring and Feedback Control to Enhance Cell Culture Process Production and Robustness. *Biotechnol. Bioeng.* **2015**, *112* (12), 2495–2504. https://doi.org/10.1002/bit.25684.

Supplementary information

Details on PLS Raman models for monitoring bioreactor using Thermo Scientific MarqMetrix All-In-One Process Raman Analyzer

Nimesh Khadka, Ph.D.
Lin Zhang, Ph.D.
Analytical Instrument, Thermo Fisher Scientific,
Tewksbury, Massachusetts USA

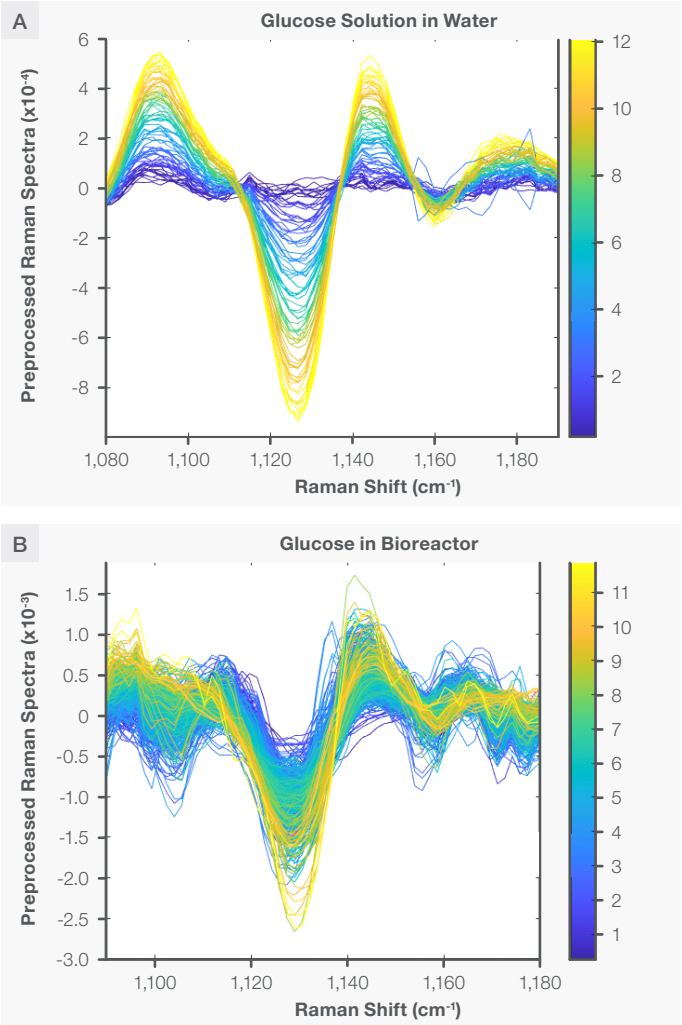


Figure S1. Characteristic Raman peaks of glucose at ~1125 cm^{-1} in aqueous solution (A) and bioreactor (B) shown as second derivative (negative peaks) spectra. The spectra are color coded by the concentration shown as the vertical bar.

Chemometric models development

Glucose PLS model

The PLS regression model for glucose was developed using a spectral region selection approach, focusing on the vibrational fingerprint of the glucose molecule. The Raman spectra of glucose in water and in the bioreactor, after applying the SavGol filter (second derivative, order = 2, window width = 13), are shown in Figures S1A and S1B. The negative peak at ~1125 cm^{-1} is attributed to the stretching vibrational modes of CO and CC and the in-plane bending of COH bonds (ν(CO), ν(CC), β(COH)) in the glucose molecule.

To normalize the spectra and enhance model robustness, three spectral regions were used: 1065-1232 cm^{-1}, 1595-1863 cm^{-1}, and 2704-3078 cm^{-1}. The 1065-1232 cm^{-1} region includes the characteristic Raman peak of glucose. The 1595-1863 cm^{-1} region includes the Raman peak of water due to the symmetrical bending of H-O-H bonds. Since the water concentration remains constant throughout the bioreactor run, this region serves as an internal reference for correcting path length differences (normalization) caused by factors like turbidity or slight variations in laser intensities. This normalization was achieved by applying SNV in the preprocessing step after baseline removal using the Savitzky-Golay filter (first derivative, order = 2, window width = 13). The normalization weight vector calculated from SNV was predominantly influenced by the spectral region corresponding to water.

The 2704-3078 cm^{-1} region includes Raman peaks mainly assigned to the symmetric and antisymmetric stretching vibration modes of various CH bonds, which are present in all biomolecules including glucose. Thus, this region provides information on total biomass in the bioreactor. Including these regions in the PLS glucose model ensures that the latent variables extract glucose-specific contributions from the overall biomass, thereby enhancing the accuracy and selectivity of the glucose PLS model.

The glucose PLS model was developed using five latent variables for the concentration range of 0 to 12 g/L, as shown in Figure S2B. The RMSECV did not improve after five latent variables, as shown in Figure S2C. The loadings for these latent variables containing glucose information are shown in Figure S2D.

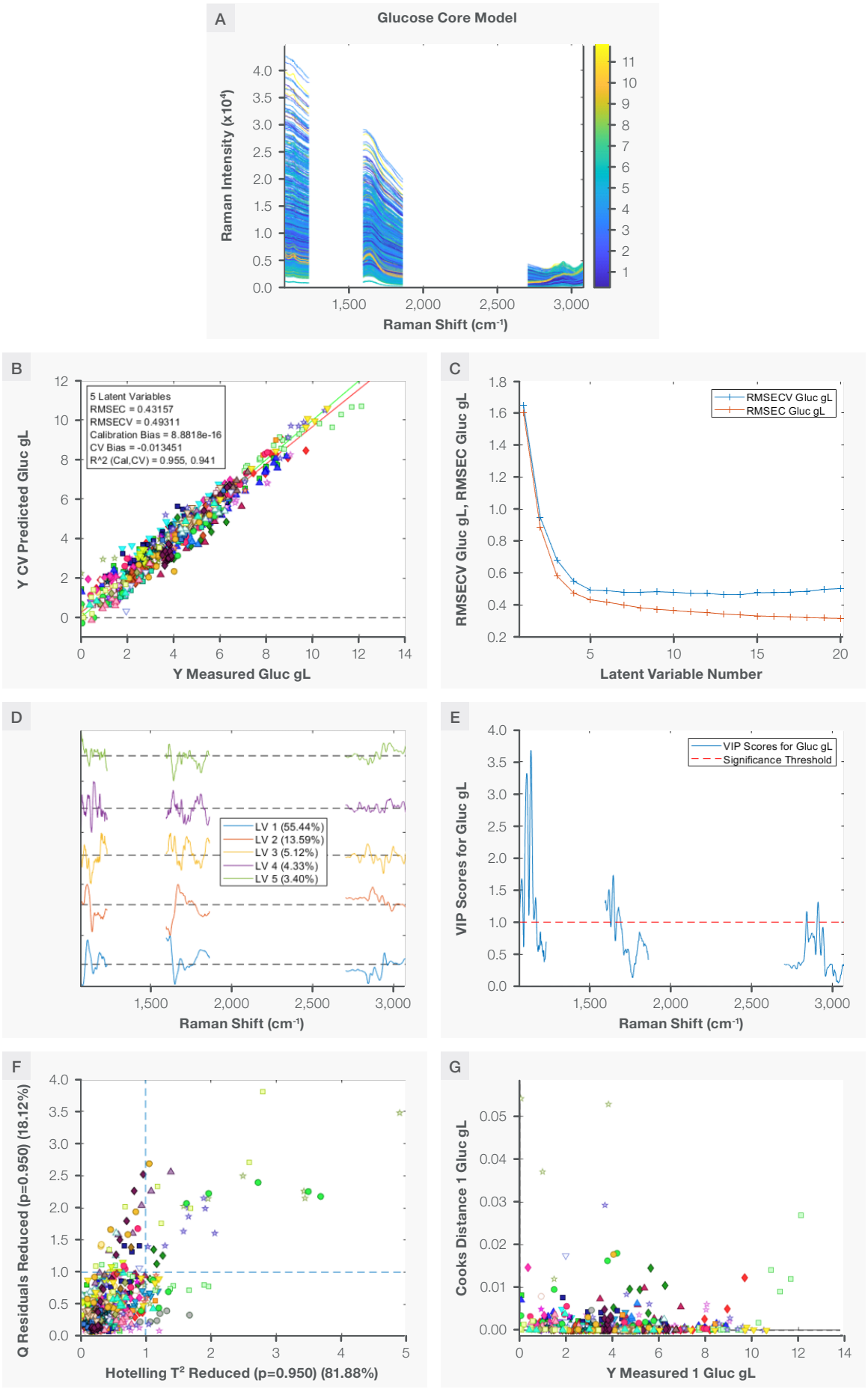


Figure S2. Glucose model: Plot A shows the region used to develop the model; plot B is the correlation plot between the measured and the prediction during cross validation; plot C shows RMSEC and RMSECV vs number of latent variables; plot D shows loading with percent variance captured; plot E shows VIP scores; plot F shows the reduced Q residual vs reduced Hotelling T^2 plot; and plot G is the Crook distance for all training samples.

Evaluation of the glucose model

i. **Specificity:** The specificity of the glucose core model was evaluated using variable importance in projection (VIP) scores. The VIP score plot shows the importance of each Raman shift in explaining glucose concentration. The VIP scores of the glucose core model are shown in Figure S2E, with the red dotted horizontal line representing the significance threshold (VIP score = 1). Raman shifts with scores higher than 1 are considered important for the model. The spectral region around ~1125 cm^{-1} has significantly higher scores, indicating specificity for glucose.

ii. **Accuracy/precision:** The glucose core model was developed with glucose concentrations of 0 to 12 g/L in the training dataset. Accuracy and precision were evaluated using RMSECV and CV bias. As listed in Table 2, the low RMSECV of 0.49 g/L and low CV bias of -0.013 g/L demonstrate high accuracy. The total measurement error is the square root of the sum of squared bias and precision errors.[1] Using RMSECV as total error and CV bias as bias error, the precision error was calculated to be 0.48 g/L. The model is statistically more accurate above glucose concentrations of 1.5 g/L (3*RMSECV) and can monitor or control glucose concentration with a tolerance limit of 0.5 g/L.

iii. **Linearity:** The R^2 for CV for the glucose core model is 0.94, demonstrating the linearity of spectral changes across the range of 0 to 12 g/L.

iv. **Q vs T^2 analysis:** The Q residual and Hotelling T^2 are used to evaluate the model. Q residual measures the residual after projecting data into the model space, while Hotelling T^2 measures sample variation within the model. Reduced Q residual and reduced Hotelling T^2, calculated by dividing Q residual and Hotelling T^2 by their corresponding confidence intervals, are used for normalized statistics. The reduced Q residual and reduced Hotelling T^2 distribution with 95% confidence interval for the core glucose model is shown in Figure S2F. Most data reside within the boundaries (dotted blue line) of reduced Q residual < 1 and reduced Hotelling T^2 < 1. Outliers are expected due to the variety in training data. When applying the model to new processes or instruments, predictions should be validated using orthogonal reference techniques if test samples have high values of reduced Q residual and reduced Hotelling T^2. Boundary limits should be determined by the user after a statistically significant process runs.

v. **Outlier analysis:** Cook's distance was calculated to identify outliers or influential data points. It measures the change in regression estimates when a particular observation is removed. A high Cook's distance indicates substantial impact on regression coefficients. Figure S2G shows all data have low Cook's distances, demonstrating no obvious outliers or influential data in the training set.

Lactate PLS model

The lactate model was developed using the same strategy as described above for the glucose model. Figure S3 shows the characteristic Raman peak for lactate at ~860 cm^{-1} in water (A) and in the bioreactor (B) after applying the SavGol filter (second derivative, order = 2, window width = 13). To ensure specificity for lactate, the model was developed using a single spectral region (800 to 1750 cm^{-1}) that includes the characteristic Raman peak for lactate and the water band for normalization, as explained above.

The lactate model was developed using five latent variables, as shown in Figure S4B. The choice of five latent variables was based on the predicted residual error sum of squares (PRESS) against the number of latent variables, as shown in Figure S4C. The loadings for the model are shown in Figure S4D.

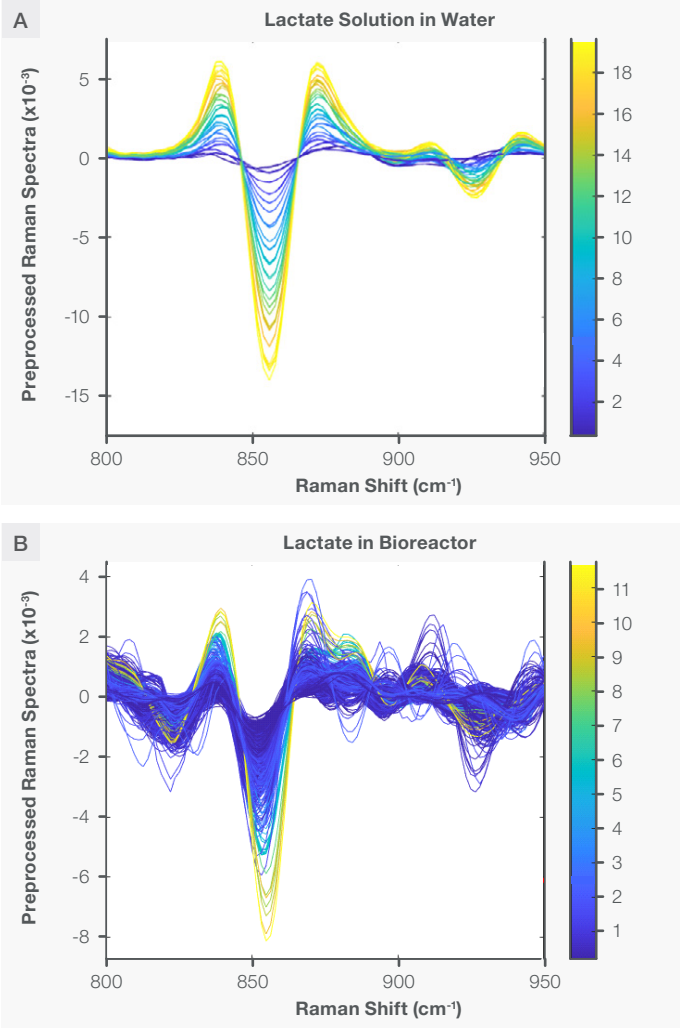


Figure S3. Characteristic Raman peaks of lactate at ~860 cm^{-1} in aqueous solution (A) and bioreactor (B) as second derivative (negative peaks) spectra. The spectra are color-coded by the concentration shown as the vertical bar.

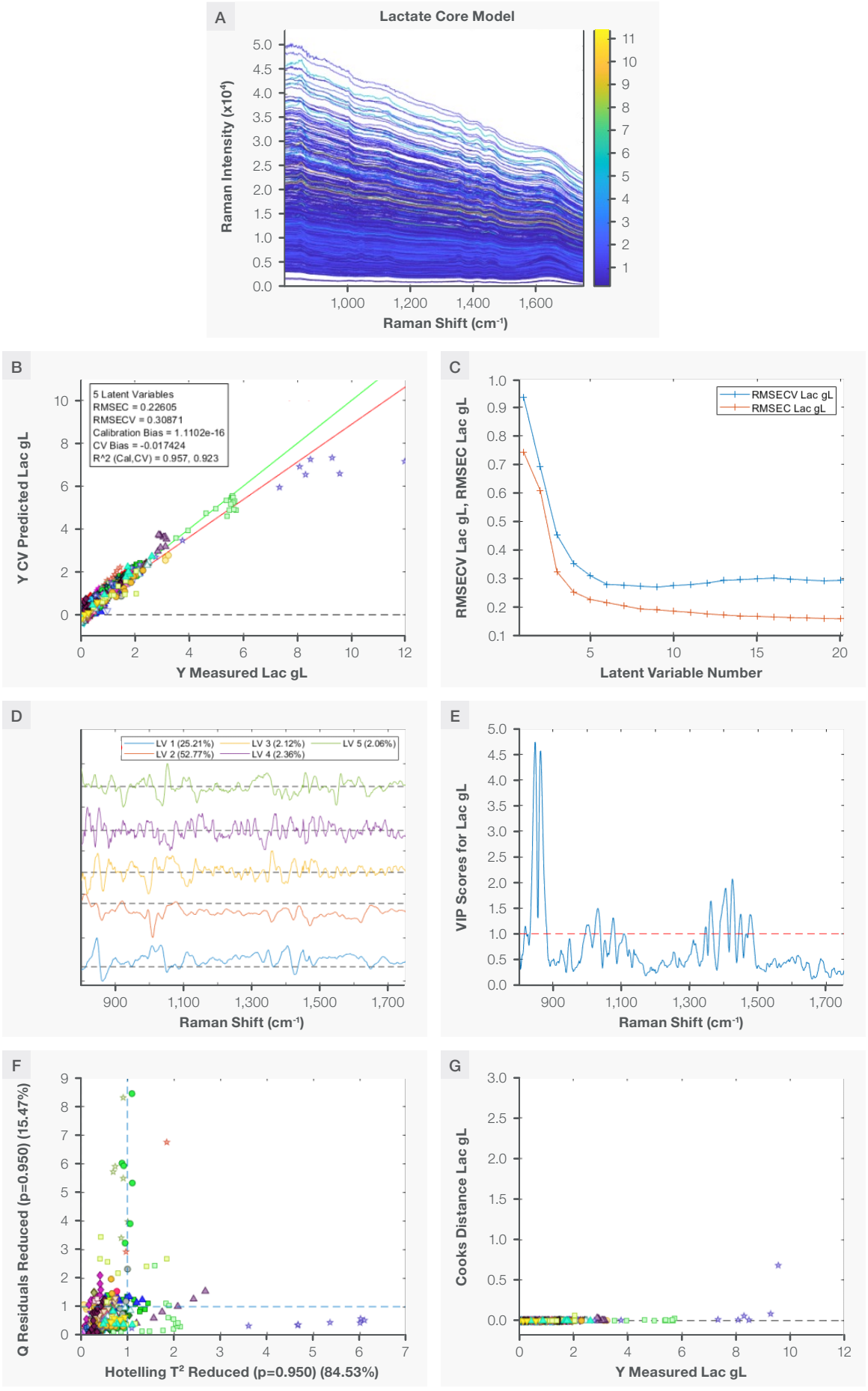


Figure S4. Lactate model: Plot A shows the region used to develop the model; plot B is the correlation plot between the measured and the prediction during cross validation; plot C shows RMSEC and RMSECV vs. number of latent variables; plot D shows loading with percent variance captured; plot E shows VIP scores; plot F shows the reduced Q residual vs. reduced Hotelling T^2 plot; and plot G is the Cook's distance for all training samples.

Evaluation of lactate model

i. **Specificity:** The VIP score plot (Figure S4E) demonstrates that the region ~820 to 880 cm^{-1} is important in the lactate model. This region is associated with the stretching vibrational mode of the C-COO- bond of lactate. Thus, the VIP score plot confirms the model's specificity for lactate.

ii. **Accuracy/precision:** As explained above for the glucose model, the RMSECV of 0.31 g/L, CV bias of -0.0174 g/L, and precision error of 0.31 g/L for the concentration range of 0 to 12 g/L demonstrate that the lactate model has accuracy and precision within acceptable tolerance for typical bioreactor process monitoring.

iii. **Linearity:** The linearity of the lactate model is demonstrated by a CV R^2 of ~0.94 for the concentration range of 0 to 12 g/L. Considering the diverse bioprocess conditions in bioreactors where the training set was collected, a CV R^2 of ~0.94 represents a high correlation between spectral information and measured concentration.

iv. **Q vs. T^2 analysis:** The distribution of the training data in the Q vs. T^2 biplot related to 95% confidence boundaries is shown in Figure S4F. The 95% confidence boundaries of raw Q and T^2 values from the training dataset are used to normalize reduced Q residual and reduced Hotelling T^2 to 1, respectively. These boundaries are represented by the dotted blue line. Predictions from the model are trusted with high confidence if test samples fall within the 95% confidence boundaries. When a test sample generates reduced Q residual or reduced Hotelling T^2 significantly higher than 1, additional reference testing is recommended to validate the prediction.

v. **Outlier analysis:** The low Cook's distance for training data indicates that there were no outliers on the training set, as shown in Figure S4G.

Performance of glucose and lactate model

The performance of glucose and lactate models were tested previously on five different cell lines / media, different scales of bioreactors, and also for the automated feedback control.[2,3] The average root mean square of prediction (RMSEP) for glucose was ~ 0.5 g/L while the average RMSEP for lactate was ~ 0.2 g/L. The low RMSEP demonstrates the accuracy, reliability, and transferability of the models across different processes and scales.

References

1. Bellon-Maurel, V.; Fernandez-Ahumada, E.; Palagos, B.; Roger, J.-M.; McBratney, A. Critical Review of Chemometric Indicators Commonly Used for Assessing the Quality of the Prediction of Soil Attributes by NIR Spectroscopy. *TrAC Trends Anal. Chem.* **2010**, *29* (9), 1073–1081. doi.org/10.1016/j.trac.2010.05.006.
2. Villa, J.; Zustiak, M.; Kuntz, D.; Zhang, L.; Khadka, N.; Broadbelt, K.; Woods, S. Use of Lykos and TruBio Software Programs for Automated Feedback Control to Monitor and Maintain Glucose Concentrations in Real Time.
3. Villa, J.; Zustiak, M.; Ramirez, D.; Kruger, J.; Kuntz, D.; Zhang, L.; Khadka, N.; Broadbelt, K.; Woods, S. Demonstrating Chemometric Model Transferability for 5 Mammalian Cell Lines and 5 Media Types Using the Thermo Scientific MarqMetrix All-In-One Process Raman Analyzer to Monitor Upstream Bioprocesses.

Application note

Introduction to method transfer between Raman spectrometers

Author

Robert Heintz

Industry/application:

Quantitative analysis, transfer of analytical methods (calibrations) between instruments, pharmaceutical manufacturing

Products used:

Thermo Scientific™ DXR3 SmartRaman+ Spectrometer, ASA (Automatic Sampling Array) accessory, TQ Analyst™ Software

Goal:

This document aims to provide a brief introduction to the transfer of analytical methods between Raman spectrometers.

Keywords:

Raman spectroscopy, chemometrics, quantitative analysis, analytical methods, models, method transfer, calibration transfer

Key benefits:

Options for transferring methods (calibrations) between Raman instruments expanding their use.

Introduction

The sampling advantages of Raman spectroscopy, along with the extensive chemical and physical information that can be extracted from Raman spectra, makes it an appealing analytical technique serving a variety of fields. The process of extracting answers from spectra can be relatively simple or involved depending on the information sought, the complexity of the Raman spectrum, and the breadth of the analytical goal. Analysis of these Raman spectra can be performed using “chemometrics,” which is a broad term for the use of mathematical processing to extract information from experimental data. Falling within the chemometrics space is multivariate data analysis. Multivariate data analysis can be used with Raman spectral data as well as data from other forms of vibrational spectroscopy to generate analytical methods (models). These analytical methods utilize previously collected spectral information (databases) to determine properties or characteristics of unknown samples based on acquired spectra. These methods can be qualitative in nature with the goal of identification or classification, or they can be quantitative where the result is a numerical value indicating the amount of a component present in the sample. The diversity of methods is mirrored by the variety of applications.

The development of analytical methods represents an investment of time and resources. For this reason, it is particularly important, especially for larger scale implementations, that once a method has been generated using a primary instrument it can then be transferred for use with secondary instruments. This is referred to as method transfer. Since there are many different types of applications that use analytical methods, there is also a range of what will be required for method transfer. This note offers a brief introduction to the general concepts of method transfer. Different tools for method transfer have been developed and studied quite extensively, but the details and the specifics of implementation are well beyond the scope of this document.[1,2]

The focus here will be on the transfer of methods between instruments, but a second part of method transfer involves changes, unintentional or intentional, in analysis conditions. Just as instrumental differences can affect the results, altering the sampling or measurement conditions can also introduce errors. This is often considered part of method maintenance. It may involve the aging of instruments or repairs to instruments or even changes in the way samples are prepared or presented for analysis. The concepts for method maintenance and method transfer are often related but the source of deviations leading to measurement errors are not necessarily the same.

A Raman spectrum typically consists of two axes. The x-axis units are typically wavelength or wavenumber (cm^{-1}) and are generally displayed as a Raman shifted spectrum (referenced to the wavelength of the laser). The y-axis is an expression of Raman intensity in various units. In an ideal world, Raman instruments would all produce exactly the same spectrum from the same sample. However, this is not the case, so it is necessary to recognize the possible variations and do what can be done to account for them.[2,3]

A calibration of both the spectrograph and the laser wavelength provides accurate and reproducible x-axis values. This typically involves using standards with known wavelengths such as neon bulbs and standards like polystyrene. Shifts in peak positions can have a noticeable effect on chemometric methods. An example of this is shown in Figure 1 where the results of shifting peak positions (+/- 4 cm^{-1}) results in significant deviations (up to 26.5%) in the results obtained from a quantitative method. With this method, a small shift (< 1 cm^{-1}) has a relatively small effect on the results but as the shift gets larger the effect rapidly increases. The magnitude of the effect of peak shifts will depend on the method and the type of preprocessing used.

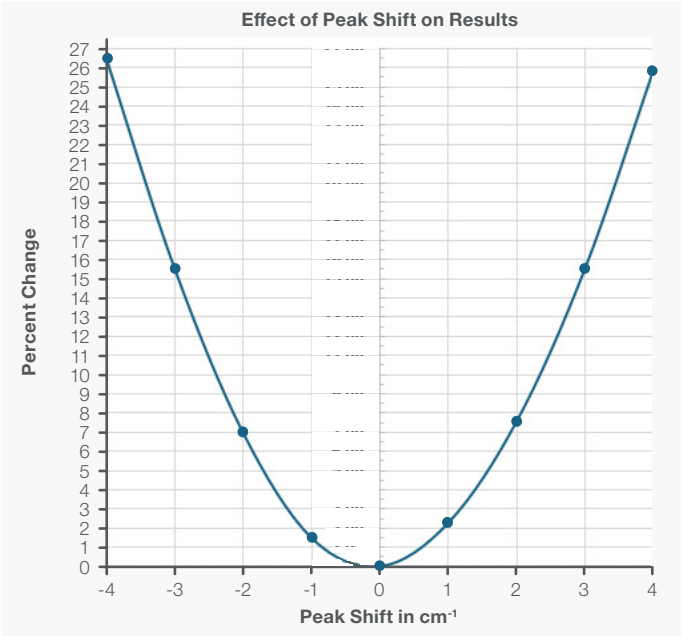


Figure 1. Effect of peak shift on the results obtained from a partial least squares analytical method.

Raman intensity is important particularly for quantitative analysis because it varies directly with the amount of material present. Many factors other than concentration affect intensity: laser power, sample focus, sample preparation, detector response, optical components, etc. While methods for standardizing intensities using NIST standards and calibrated white light sources have been used, that does not solve all the problems of intensity variations between instruments. It is advisable to plan for a way that fits within the specifics of the application, to allow for use of peak ratios or another method to normalize the intensities of the Raman spectra.

Different approaches to method transfer

Direct transfer method: The method is developed solely on the primary instrument and then directly transferred to secondary instruments.

- Direct transfer does not require collection of standards on secondary instruments.
- It involves less time and effort.
- Successful direct method transfer is not always possible.
- Method transfer success depends on the method and analysis requirements.

It is important to optimize methods to minimize the effects of variance through techniques like preprocessing (normalization, baselines, etc.—there are too many preprocessing options to detail here)[2,4] or the selection of other parameters (spectral ranges, etc.).

Global methods (full calibration): This requires collecting reference spectra on both the primary instrument and secondary instruments.

- This is a large all-inclusive method.
- A robust model, it incorporates instrument-to-instrument variations
- Because it is so inclusive, it requires more time to collect the spectral data on each instrument.
- The method must be updated with each new instrument.

Correction or standardization: This method uses a limited number of correction or transfer spectra collected on secondary instruments.

- Many different options exist for this approach—as with direct transfer, there are too many to detail here[3,5,6,7] — but they include univariate and multivariate approaches.
- Correction spectra are used to generate corrections to "adjust" results from the primary method.
- Standardization spectra are used to generate transfer functions for standardizing spectra.
- This method requires fewer secondary spectra than a full Global Method.

To illustrate a couple of these concepts, quantitative methods were developed using a Thermo Scientific™ DXR3 SmartRaman+ Spectrometer. The transfer of methods to other DXR3 SmartRaman+ instruments were investigated.

Experimental

Samples used in this investigation were aqueous solutions of acetaminophen with concentrations between 0 – 10 mg/ml (see Figure 2). The standards and test solution (2.604 mg/ml) were analyzed in glass vials using a custom sample holder for the ASA (Automatic Sampling Array) accessory used with the DXR3 SmartRaman+ spectrometer. The excitation source was a 532 nm laser. This testing involved a primary instrument and 5 secondary instruments. Spectra were collected from the standards as well as the test sample on all 6 of these instruments. A partial least squares method was generated using TQ Analyst software. Preprocessing of the spectra included spectral normalization of peak intensities using the water peak at 3416 $cm^{-1}$and using first derivatives to minimize baseline offsets.

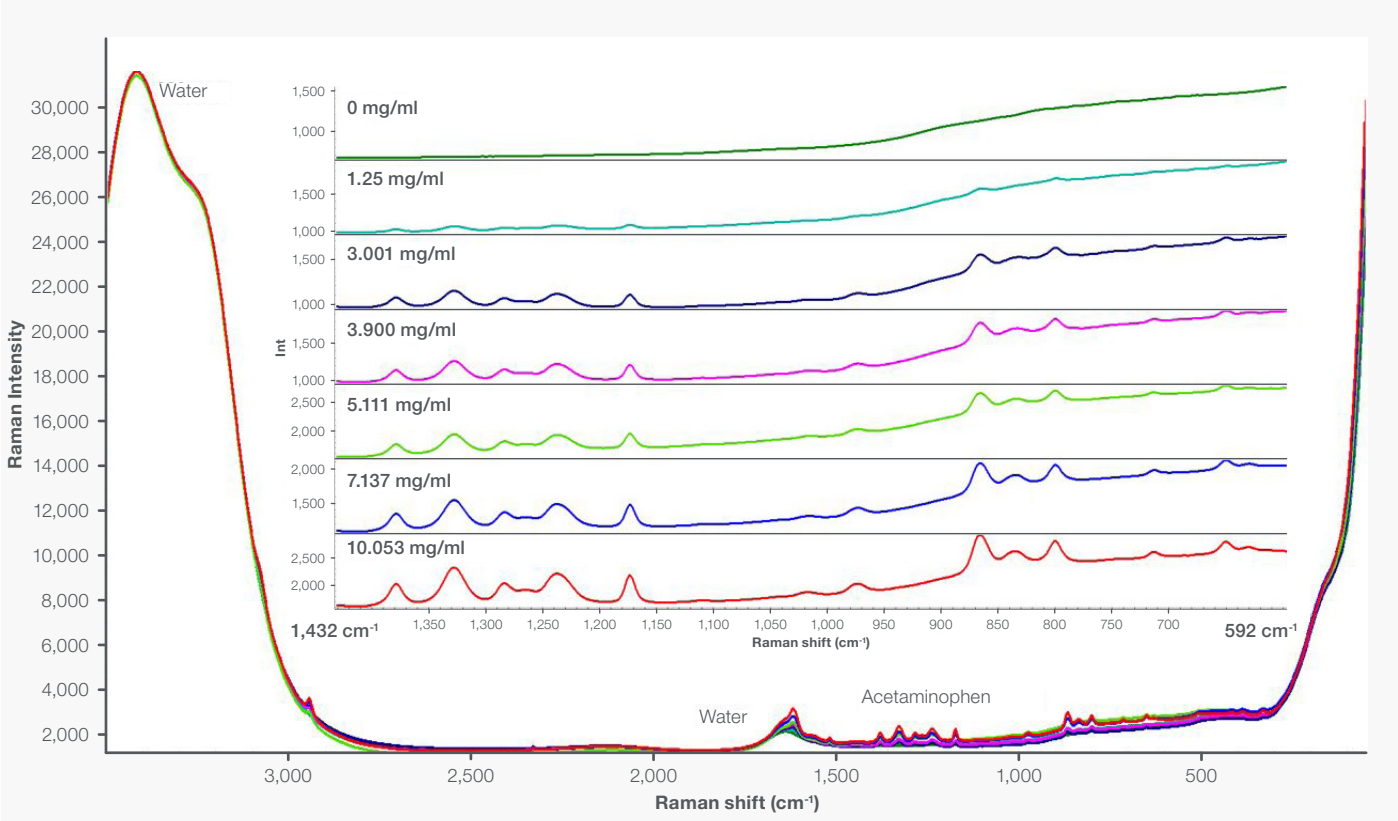


Figure 2. Representative Raman spectra from the acetaminophen solutions.

	Primary	Secondary #1	Secondary #2	Secondary #3	Secondary #4	Secondary #5
Region 1	-1.40%	-3.60%	-17.99%	19.44%	-11.51%	31.22%
Region 2	1.58%	0.00%	-6.52%	-1.33%	-9.17%	-7.44%

Table 1. Percent differences from the expected concentration of the test solution (2.604 mg/ml) for the primary instrument and all 5 secondary instruments. Region 1 is 1432–592 cm^{-1} and Region 2 is 2960–2930 cm^{-1}.

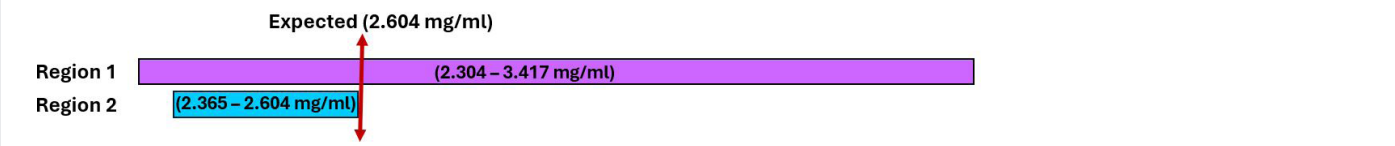


Figure 3. Range of values for the test solution (2.604 mg/ml) determined on the primary instrument and all 5 secondary instruments. Region 1 is 1432–592 cm^{-1} and Region 2 is 2960–2930 cm^{-1}.

	Primary	Secondary #1	Secondary #2	Secondary #3	Secondary #4	Secondary #5
Direct	-1.58%	0.00%	-6.52%	-1.33%	-9.17%	-7.44%
Global	1.58%	6.08%	0.35%	0.09%	-1.92%	-1.96%
Correction	0.83%	2.42%	-4.17%	1.08%	-6.85%	-5.10%

Table 2. Percent differences between the values determined for the test solution and the expected value (2.604 mg/ml) for the primary instrument and all 5 secondary instruments using the direct transfer method, the global method, and the correction method.

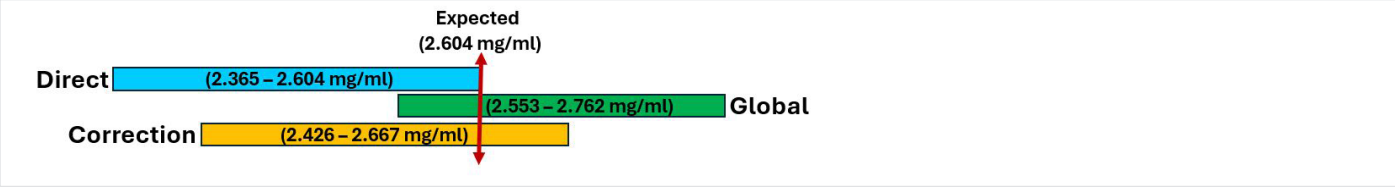


Figure 4. Range of values for the test solution (2.604 mg/ml) determined on the primary instrument and all 5 secondary instruments using the direct transfer method, the global method, and the correction method.

One important parameter that was optimized for method transfer was the selection of the spectral region used for the method. The initial region selected was in the fingerprint region (1432–592 cm^{-1}) and was selected due to the presence of multiple acetaminophen peaks. However, the success of direct transfer of the method from the primary instrument to the secondary instruments was dependent on the preprocessing and specific parameters used. Using the test solution (2.604 mg/ml), the secondary instruments gave values from 2.304 to 3.417 mg/ml (up to a 31% difference). An analysis of the spectral variance in the normalized spectra showed that the C-H stretching region was more consistent, so the method was recalibrated using the region at 2960–2930 cm^{-1}. This time the secondary instruments gave values ranging from 2.365 to 2.604 mg/ml (0–9.2% difference). A comparison of the results is shown in Table 1 and Figure 3. Figure 3 graphically compares the range of values obtained, and Table 1 shows a comparison of how close the values were to the expected value as percent differences. Using the second region shows a considerable improvement in the direct transferability of method.

Multiple vial-holder being loaded into the ASA (Auto Sampling Array) accessory for measurement.

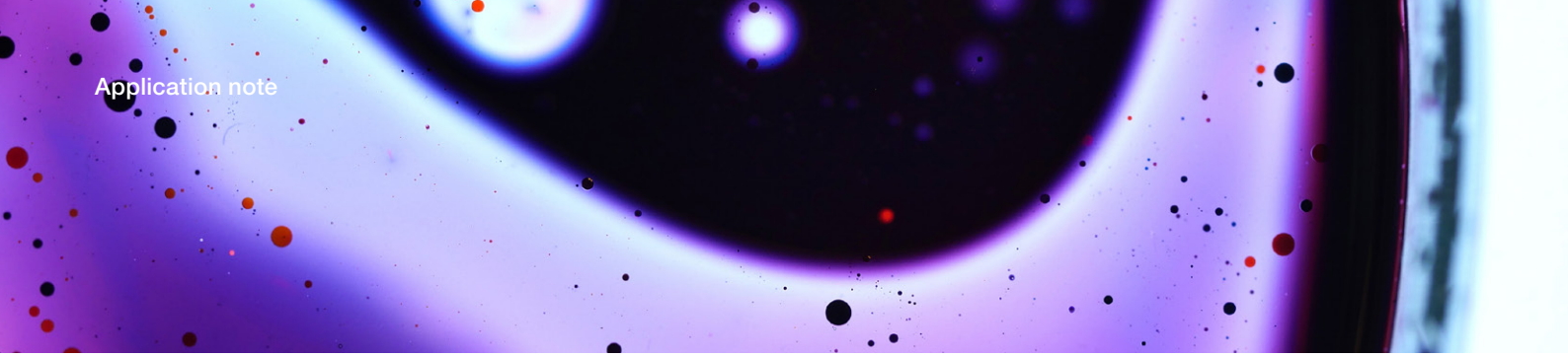
To illustrate how different approaches to method transfer compare, a couple of simple examples were developed. The direct transfer method has already been discussed in the previous section. A global method was also generated by collecting all the standards on the primary as well as all the secondary instruments and using all the reference spectra in the analytical method. The other example is a correction method where the direct transfer method is used but a correction is applied to the results based on a limited number of correction spectra collected on the secondary instruments. The correction is applied post-prediction and is a mathematical fit based on the correction spectra collected on the secondary instruments. In this case the correction was a first order fit. The test solution (2.604 mg/ml) was run on all the instruments, and the calculated intensities based on the different methods were compared. Figure 4 shows a comparison of the range of values obtained from the different method transfer options. Table 2 shows the percent differences between the calculated and expected concentrations of the test solution for each of the instruments. The global method gave the best results in that it has the smallest range and is centered closest to the expected value, but it also required the greatest number of spectra. Using the correction spectra with the direct transfer method improved the results over the direct transfer method itself. Whether the extra work involved in the global method or the correction method is warranted for the enhancement of results depends on the goals of the application, the way methods are implemented, and the analysis requirements.

Conclusions

This was intended as a general introduction to the concepts of method transfer between Raman spectrometers. Method transfer in vibrational spectroscopy has been studied extensively, and a vast number of techniques have been developed and investigated. The examples presented here just illustrate a very small part of a vast subject. An important key to success is to use high quality spectra generated from a Raman spectrometer such as the DXR3 SmartRaman+ spectrometer and then to optimize the analytical method. The next step is to consider what is going to be required for method transfer. While direct method transfer is the most appealing approach because it represents the most straightforward approach with the least effort, it may be the exception rather than the rule. Constructing a global method is a more rigorous approach but is not always practical from both an effort and an implementation standpoint. Correction or standardization method transfer are tempting alternatives because they often require less data compared to a full global method. The example presented here was a correction method transfer approach. The standardization approaches were not addressed here but they use transformations that strive to “standardize” the spectra themselves. There are many options on how to implement method transfer, and in the end, the choice will depend on the specific application.

References

1. Jerome Workman: *A Review of Calibration Transfer Practices and Instrument Differences in Spectroscopy*, **Appl. Spectrosc.** 2018, Vol.73(3) 340-365.
2. B. Barton, J. Thomson, E.L. Diz, R. Portela: *Chemometrics for Raman Spectroscopy Harmonization*, **Appl. Spectrosc.** 2022, Vol 0(0), 1-21.
3. Yu-Cung Lin, Joseph Sinfield: *Characterization of Raman Spectroscopy Systems Transfer Functions in Intensity*, Wavelength, and Time, **Instruments** 2020, 4, 22, 1-52.
4. S. Guo, J. Popp, T. Bocklitz: *Chemometric analysis in Raman spectroscopy from experimental design to machine learning-based modeling*, **Nature Protocols**, 2021, Vol. 16, 5426-5459.
5. J.H. Kalivas, G.G. Siano, E. Andries, H.C. Goicoechea: *Calibration Maintenance and Transfer using Tikhonov Regularization Approaches*, **Appl. Spectrosc.** 2009, Vol. 63(7), 800-809.
6. M. Ross Kunz, J.H. Kalivas, E. Andries: *Model Updating for Spectral Calibration and Transfer Using 1-Norm Variants of Tikhonov Regularization*, **Anal. Chem.** 2010, 82, 3642-3649.
7. D. Brouchaert, J.S. Uyttersprot, W. Broeckx, T. De Beer: *Calibration transfer of a Raman spectroscopic quantification method from at-line to in-line assessment of liquid detergent compositions*, **Analytica Chimica Acta**, 2017, 971, 14-25.



Application note

Protein concentration prediction in cell cultures

The next stage in NIR bioprocess analysis

Author

Todd Strother, PhD,
Thermo Fisher Scientific,
Madison, WI, USA

Figure 1. Antaris MX FT-NIR Process Analyzer used for collecting thespectroscopic information from the cell cultures.

Introduction

Biologically produced materials are an increasingly important aspect in many industrial processes including those related to pharmaceuticals, food, diagnostics, and fuels. Most of these biologicals are produced in fermentors and bioreactors in which specialized cell cultures grow and manufacture the molecule of interest. Many different types of cells are used in culturing and producing biopharmaceutical products including genetically engineered bacterial and yeast cells. However a majority of the products are proteins cultured from mammalian systems such as Chinese hamster ovary (CHO), green monkey (VERO), or human embryonic kidney (HEK) cell lines. Many of these products are large complex proteins, hormones or polysaccharides that are impossible or difficult to manufacture in large quantities any other way. A recent survey of the US Food and Drug Administration noted that there are over 350 biologicals approved for various uses, including vaccines and diagnostic and therapeutically important antibodies.

Bioprocesses that produce the desired materials by nature rely on complex biological systems to synthesize their useful products. While typical chemical manufacturing processes have relatively little variability, the inherent complexity of biological systems makes a great deal of variability from batch to batch inevitable. As a consequence of the complexity and variability of the processes, it has been estimated that 30% of the production batches need to be reprocessed for quality reasons, which results in a tenfold loss in profit. Industries that rely on these complex biological systems benefit greatly from closely monitoring the growth of their cell cultures and production of the target molecule. Process analytical technology (PAT) initiatives in bioprocesses improve the overall product quality, reducing waste by accounting for this inherent variability.

Monitoring and controlling cell culture conditions greatly reduces this variability and results in improved target protein production. Fourier transform near-infrared (FT-NIR) spectroscopy has proven to be a useful technology for monitoring and controlling manufacturing processes including more specific bioprocess applications. It is also part of PAT initiatives across many industries including bioprocessing. Previous work performed on cell cultures using NIR spectroscopy has usually focused on monitoring and controlling nutrients, waste products, cell densities and other parameters related to the health of the cell culture. While these parameters are useful for determining the relative health of the cell culture, the more important parameter of interest is the actual production and concentration of the target molecule. Very few NIR studies have determined and measured protein concentrations in actual cell culture conditions. This application note demonstrates the feasibility of using the Thermo Scientific™ Antaris™ MX FT-NIR Process Analyzer (Figure 1) to predict protein concentrations at biologically relevant concentrations in dynamic cell cultures.

NIR spectroscopy uses light between 10,000 and 4,000 cm^{-1} to determine the identity and quantity of a variety of materials. Most molecules of interest absorb light in this region through combination or overtone vibrations. The advantage of performing spectral analysis on these absorption bands is that the light is able to penetrate more deeply into the material under analysis and does not require dilution or manipulation of the sample. Therefore NIR analyzers can be coupled directly into a process stream or tank where spectral analysis can be performed without human intervention. FT-NIR has been implemented in many different industrial, pharmaceutical and other process settings for many years and has proven to be extremely valuable in collecting real-time analytical data automatically. When used in process environments, the Antaris MX FT-NIR Process Analyzer is easily coupled to process control computers where it is an integral part of maintaining optimal manufacturing conditions. Because of these advantages and the need to control the inherently variable biological systems found in cell culture technologies, NIR is an excellent choice for analyzing different components in bioreactors including proteins.

Methods

Chinese hamster ovary (CHO) cell cultures were grown at optimal conditions until the cell concentrations reached approximately one million per millemeter, representing a typical cell density for a young and growing culture. Samples of the cell culture were tested on a Nova BioProfile™ analyzer to determine concentrations of glucose, glutamine, lactate, and ammonia. The concentrations of these materials changed throughout the experiment and accounted for some variability that might be encountered across multiple cultures. The concentrations were variously and singularly altered by spiking the samples with nutrients or waste products or diluting the samples with unaltered cell culture. Each of those four components was altered so that two or three different concentrations were represented for each. Table 1 lists the concentration ranges for the various nutrient, waste, and protein components of the tested samples. This methodology also has the effect of removing covariance between the different components and protein present.

Component	Range (g/L)
Protein	0.16–5.00
Glucose	7.98–8.12
Glutamine	0.28–0.58
Lactate	0.45–0.90
Ammonia	0.05–2.39

Table 1. Concentration ranges of various components. The solutions represent over 35 different protein concentrations that also vary in concentrations of nutrient and waste components.

Ultrapure bovine albumin protein was added to the solutions to represent target protein synthesized by the cells. Genetically modified cell cultures are designed to produce the target protein in large quantities almost exclusively to all other cellular proteins. As a result, the protein concentrations in the cell culture media will often approach and exceed 5.0 g/L and consist almost entirely of the target molecule. Albumin protein is an excellent mimic for recombinant proteins because it is available in extremely pure form and contains NIR active groups essentially identical to a typical target protein from a cell culture. In this case, purity is extremely important because any extraneous material present will also have a NIR signal and would lead to confounding results. The albumin protein material was carefully weighed and added to the cell cultures in concentrations ranging from 0.16 to 5.0 g/L. Over 35 different solutions were produced that had a range of nutrient and waste as well as protein concentrations. These varied solutions resulted in 54 spectra that were used to build the chemometric method and 20 spectra that were used to validate that method.

The cell culture samples were scanned with an Antaris MX FT-NIR Process Analyzer in the range between 10,000 and 4,000 cm^{-1}. The analyzer was coupled to a transflectance probe with an adjustable path length. The gap distance was set to 1.25 mm for a total path length of 2.5 mm. Sixteen scans were averaged per spectrum and were collected using eight wavenumber resolution with a gain of 0.1. Sample time took approximately 15 seconds. Two spectra were collected per sample. Figure 2 shows images of the probe before insertion into a cell culture sample and during spectral collection.

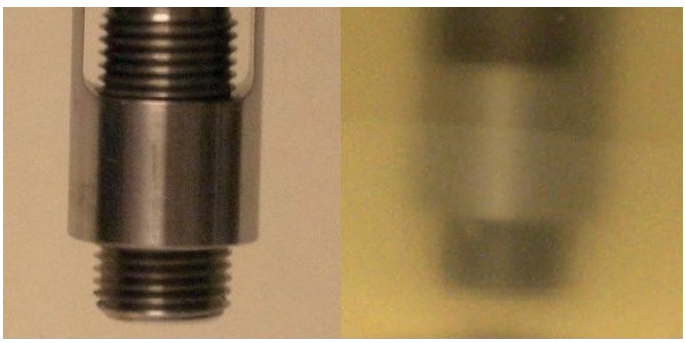


Figure 2. Transflectance probe used for data collection. Left panel shows the design of the probe with the adjustable pathlength. Right panel shows probe inserted into cell culture during data collection.

The sample spectra were loaded into the Thermo Scientific TQ Analyst™ Pro Edition Software for chemometric analysis using a partial least squares (PLS) method with a constant pathlength. The spectra were analyzed in the first derivative using a Norris smoothing filter. Two regions were used for the analysis: 8,910 to 5,340 cm^{-1} and 4,830 to 4,340 cm^{-1} These two regions collected information across a wide range of data points while avoiding the totally attenuating water peak centered around 5,100 cm^{-1}. Figure 3 shows representative raw spectra and the first derivative spectra of the samples.

Results

PLS analysis of the protein concentrations in the various cell culture samples revealed excellent predictive capabilities within the range of materials tested. The 54 spectra used to develop the PLS method are shown on a calibration plot (Figure 4) that compares the calculated protein concentrations versus the actual concentrations.

The calibration plot can be used to determine how well the method predicts the actual protein concentrations in the samples. The plot developed by the chemometric method resulted in a correlation coefficient of 0.977. Root mean square error of calibration (RMSEC) was 0.33 g/L and the Root mean square error of prediction (RMSEP) calculated from the 20 validation samples was 0.31 g/L. Additionally, the Root mean square error of cross validation (RMSECV) was 0.51 g/L. These errors indicate that the protein concentration in the cell culture samples can be predicted to 0.5 g/L or less. Approximately 1/3 of this error was attributed to the balance used to weigh the protein material.

Conclusions

Measuring protein concentrations in living dynamic cell cultures was successfully performed with the Antaris MX FT-NIR Process Analyzer. Protein concentration is a critical parameter in determining the success and quality of a cell culture in manufacturing a viable end product. This NIR technique successfully demonstrates the ability to measure and monitor protein concentrations in real time at relevant concentrations. The developed method shows excellent correlation with actual protein concentrations between 0.16 and 5.0 g/L and with errors of less than 0.5 g/L.

This application demonstrates the continued capability of the Antaris MX FT-NIR Process Analyzer to be successfully used in bioprocess environments where it can safely, accurately and automatically monitor and control cell cultures. While previous NIR studies have monitored cell culture conditions to promote optimal protein production, few have actually monitored and predicted protein concentrations. This feasibility study shows the power of the Antaris MX FT-NIR Process Analyzer to correctly predict target protein concentrations in a live and dynamic cell culture.

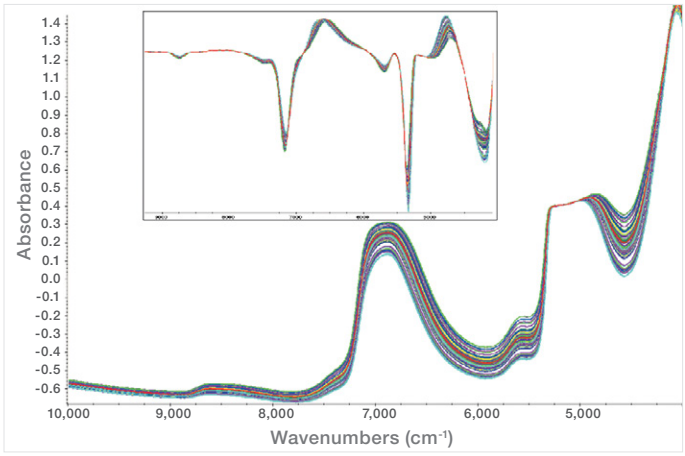


Figure 3. Representative raw spectra showing the variability present in the cell culture samples. Regions of analysis avoided the attenuated water peak at 5,100 cm^{-1}. Inset shows the first derivative spectra used for the PLS chemometric method.

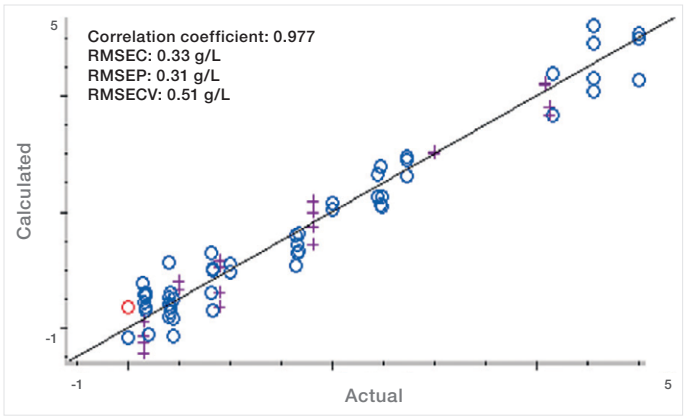
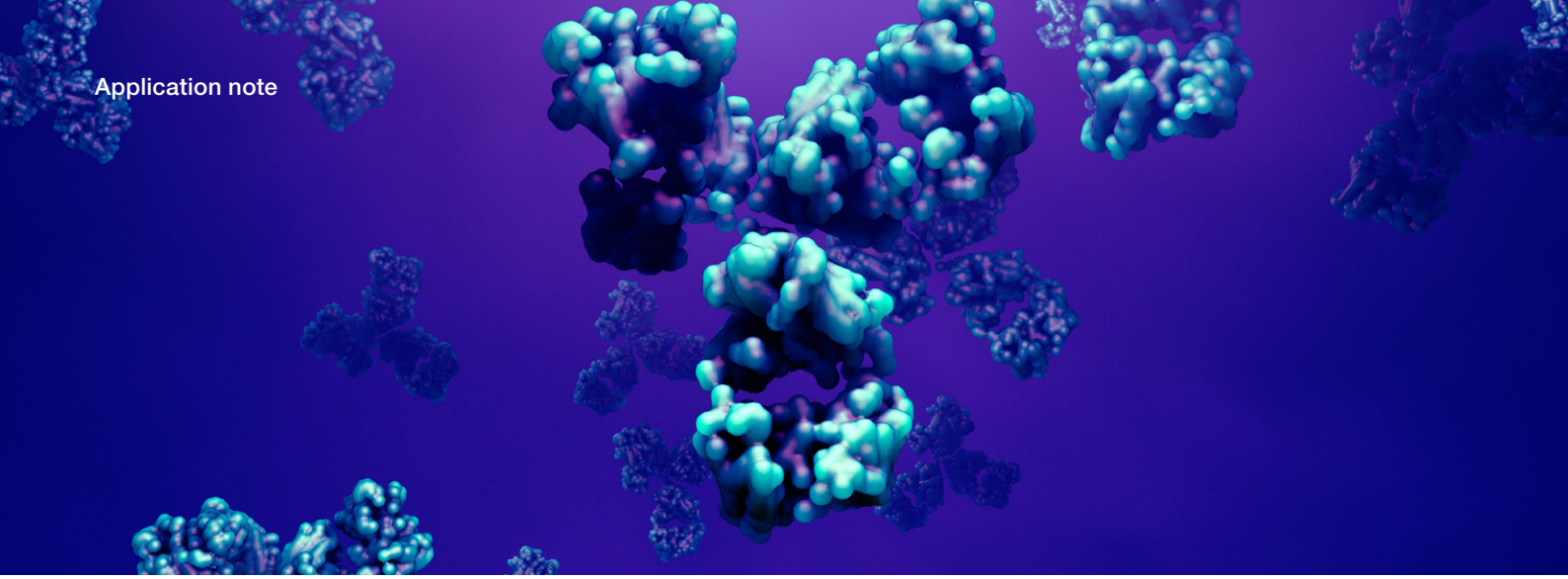


Figure 4. Calibration plot comparing the calculated protein concentrations to the actual concentrations from the PLS method. Root mean square errors are approximately 0.5 g/L or less. Blue circles (o) represent spectra used to create the method, purple crosses (+) are spectra used to validate the method.

Learn more at **thermofisher.com/brighteroutcomes**



Application note

Protein aggregation identified through UV-Visible absorption spectroscopy

Introduction

Misfolded or denatured proteins can associate in solution,[1] forming insoluble aggregates (Figure 1). This process is often irreversible, effectively removing useful proteins from solution and making the detection of aggregates critical for further downstream use of protein solutions. This is particularly important when studying unstable or abnormal proteins, which are more likely to form aggregates.[2,3]

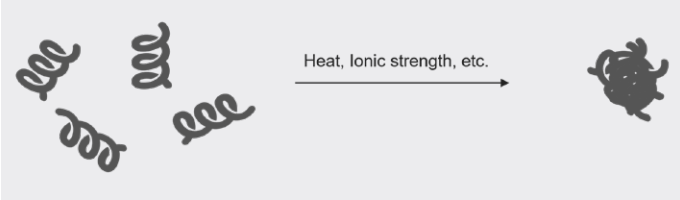


Figure 1. Visualization of protein aggregation induced by heat or changes in ionic strength.

The formation of protein aggregates in the body has also been linked to several diseases, including Alzheimer's and Parkinson's disease.[1,4,5] In the pharmaceutical industry, protein therapeutics, such as insulin,[6] have been developed to effectively treat a variety of diseases but have been difficult to synthesize.[7] The presence of aggregates in these products can lead to lower product yields and can reduce the efficacy of the final therapeutic.[5, 8] For example, protein therapeutics that undergo aggregation have been linked to lowered immune responses and, in some cases, can even induce allergic reactions.[8]

In the food industry, protein composition can have a large impact on the palatability of the final product. Protein aggregates can significantly change a food's organoleptic properties (e.g., taste, smell, etc.), as well as the digestibility of the material.[5]

Size-exclusion chromatography has previously been used to identify the presence of aggregates in a sample.[9] This characterization method is time-consuming, however, and sample retrieval can be difficult. An alternative method for the detection of protein aggregates uses UV-visible (UV-Vis) absorption spectroscopy, a technique that measures a sample's light absorption. Aggregates in solution are known to scatter incoming light, resulting in an apparent absorption artifact across the entire spectrum.[5, 10] This scattering artifact does not represent the true absorption of the sample and instead indicates that the solution contains aggregates large enough to scatter the incoming light.

In this application note, UV-Vis absorption spectroscopy was used to identify the presence of protein aggregates in aqueous bovine gamma globulin (BGG) samples. Aggregation was induced in these samples using heat or the addition of NaCl. An integrating sphere was further used to measure the scatter-free spectra of the samples. Scatter-correction methods were used to determine the concentration of free, non-aggregated BGG in solution.

Experimental

Absorption spectra were collected using a Thermo Scientific™ Evolution™ One Plus UV-Vis Spectrophotometer. Samples were held in a 10 mm quartz cuvette, and measurements were collected between 220 and 400 nm. A stock 1.1 mg/mL BGG solution was made by diluting standard Thermo Scientific™ Pierce™ BGG Standard (2.0 mg/mL, Lot Number MH162604) with phosphate buffer (PBS, 1×) to achieve the appropriate concentration. A 5.3 M NaCl solution in phosphate buffer was made by dissolving 1.5 g NaCl (Fisher Scientific) in 6.0 mL of phosphate buffer. BGG samples were prepared as described in Table 1.

	BGG sample		Volume of 1.1 mg/mL BGG (mL)	Volume of PBS (mL)	Volume of 5.3 M NaCl (mL)
	Temperature (°C)	NaCl concentration (M)			
1	25.0	0.00	1.0	1.0	0.0
2	25.0	2.65	1.0	0.0	1.0
3	75.0 (60 min incubation)	0.00	1.0	1.0	0.0
4	75.0 (30 min incubation)	0.00	1.0	1.0	0.0

Table 1. BGG solution preparation.

BGG samples were heated using a single-cell Peltier accessory at 75°C for 30 or 60 minutes. Sample measurements were collected using a Thermo Scientific™ Evolution™ ISA-220 Integrating Sphere Accessory in transmission geometry. The collected data was reported using the Kubelka-Munk transformation. An 8° wedge was used for optimized light collection. After integrating sphere measurements were completed, Sample 4 (Table 1) was filtered using a syringe filter. The absorption spectrum of the filtrate was then measured using the Evolution One Plus Spectrophotometer, without the Evolution ISA-220 Accessory.

Results

The absorption spectrum of BGG (not aggregated), depicted in Figure 2a (blue curve), is in agreement with literature values.[11] Upon addition of NaCl, the entire spectrum appears to have a higher absorbance, an artifact resulting from the presence of larger particulates. Increased ionic strength of a protein solution (due to high salt concentration) has been shown to induce protein aggregation;[4] this scattering signal can therefore be attributed to the presence of small BGG aggregates. Scattering is observed regardless of the visual (clear, non-turbid) appearance of the solution (Figure 2c). This indicates that, while it is difficult to confirm through visual observation alone, aggregate scattering can be measured using UV-Vis absorption, and the technique can be used as a test for protein aggregation.

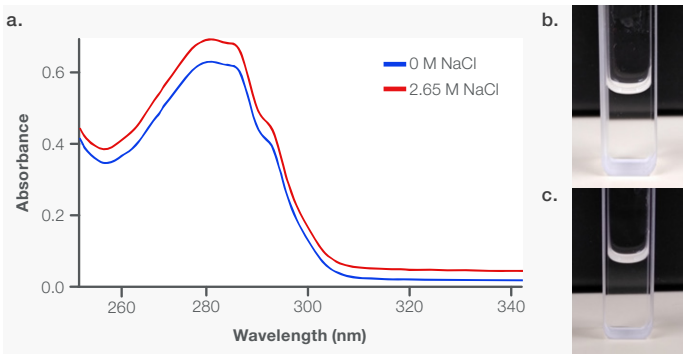


Figure 2. Absorption spectra of 0.55 mg/mL BGG in PBS with (red) and without (blue) 2.65 M NaCL. Images of a solution of BGG with (b) and without (c) 2.65 M NaCl.

Scattering appears as a raised baseline at longer wavelengths but also influences the apparent absorption across the entire spectrum and is highly dependent on the wavelength of the incident light. This influence can be estimated using the following equation:

$$A_{scatter} = \log (I_0/I_{no\ scatter}) + A_{offset} = \log (I_0/I_0\text{-}(f/\lambda^4))) + A_{offset} \quad (1)$$

In the equation above, $A_{scatter}$ is the scattering artifact/apparent absorption due to scattering, I_0 is the intensity of the light before it interacts with the sample, $I_{no\ scatter}$ is the intensity of the light that reaches the detector (not scattered by the solution), f is an arbitrary scaling factor, λ is wavelength in nanometers, and A_{offset} is an offset. This equation uses Beer's law,

$$A = \log (I_0/I) \quad (2)$$

and the relationship between the wavelength of light and the intensity of the scattered light, which is defined by the Rayleigh equation,[12]

$$I_{scatter} \propto (1/\lambda^4) \quad (3)$$

to determine an estimated intensity of the scattered light ($I_{scatter}$). Assuming I_0 is 1 and the intensity of the scattered light is less than 1, Equation 2 includes only two parameters that must be fit to determine the scattering contribution. The relationship between scattering intensity and wavelength indicates that there is a larger effect in the UV region (Figure 3a), where there are prominent absorption features for proteins. This effect must therefore be carefully corrected.

Figure 3b shows the data corrected using two different methods. The first, referred to as "baseline correction," involves taking the average of the absorption reported in the spectral region in which the sample should not absorb. The calculated average is then subtracted from each point in the spectrum, as described by:

$$A_{corrected,\lambda} = A_{measured,\lambda} - A_{average,(330\text{-}350\ nm))} \quad (4)$$

In this equation, $A_{measured\lambda}$ is the absorption spectrum collected, $A_{average,\ (330–350\ nm)}$ is the average of the absorption measured between 330 and 350 nm, and $A_{corrected\lambda}$ is the corrected absorption spectrum. The resulting spectrum is shown in Figure 3b (green curve); the maximum absorption from the band is still higher than that of the untreated BGG sample. This does not match the expected result, as formation of aggregates should remove free BGG from solution, leading to a lower concentration and lower absorbance in the region of interest. Consequently, the "baseline correction" does not properly account for the scattering artifact present in the collected spectrum.

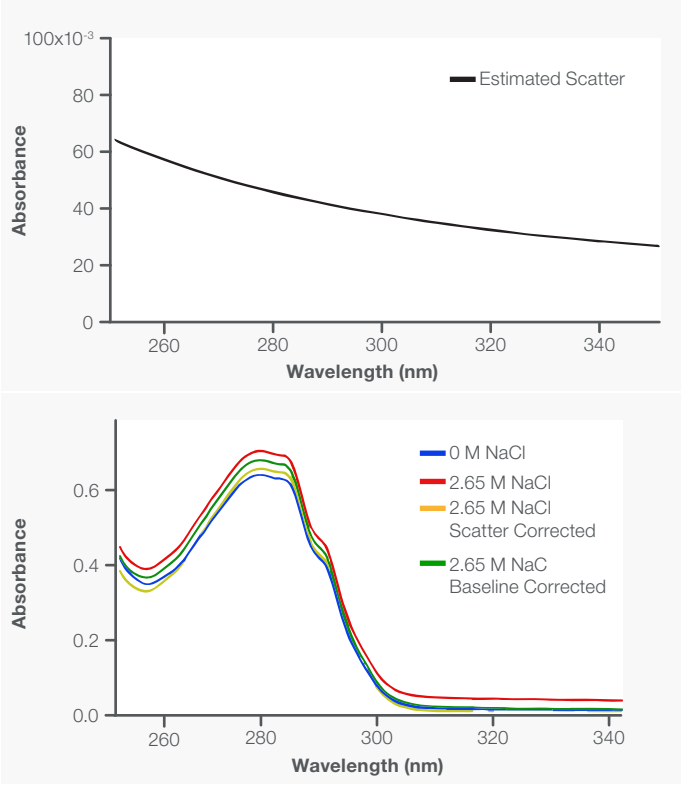


Figure 3. a) Estimated scattering calculated using Equation 1. b) Absorption spectra of BGG with and without NaCL. Baseline-corrected data is shown in green, calculated using Equation 4. Scatter-corrected data is shown in yellow, calculated using Equation 5.

The second method, called "scatter corrected", fits the long wavelength baseline to Equation 1, where f and A_0 are fit such that the resulting function matches the long wavelength signal well. The scattering function described in Figure 3a was fit using f = 6.1 x 10^8 and A_0 = 0.006. The resulting scatter function was then subtracted from the absorption spectrum, as shown in the following equation,

$$A_{(corrected,\lambda)} = A_{measured,\lambda} - A_{scatter,\lambda} \quad (5)$$

where $A_{scatter\lambda}$ is the calculated scatter estimate. This correction results in the yellow spectrum in Figure 3b. Unlike the baseline corrected spectrum (green curve, Figure 3b), the maximum absorption of the scatter-corrected spectrum is below the absorption maximum of the spectrum for untreated BGG, as expected.

The concentration of free, non-aggregated BGG in the sample was found to be 0.54 mg/mL using Beer's law:

$$A = c l \varepsilon \quad (6)$$

In the equation above, A is the measured absorbance, c is the concentration, l is the path length (1 cm), and ε is the extinction coefficient of the protein. Therefore, the concentration of proteins that contribute to aggregation in this sample is 0.01 mg/mL.

For samples with a relatively low scatter contribution, the mathematical scatter-correction method works well. However, for samples that are visibly cloudy/turbid, this correction is not ideal, as only a small portion of the light is allowed to interact with the detector. To study a sample that is turbid, a 0.55 mg/mL BGG sample was held at 75°C for 60 minutes using a single-cell Peltier accessory for the Evolution One Plus Spectrophotometer, producing a cloudy solution (Figure 4b). The resulting absorption spectrum is depicted in Figure 4a. The scattering artifact present indicates that ~30% of the light is transmitted through the sample at 310 nm, where BGG begins to absorb, and even less is transmitted at shorter wavelengths. This suggests there is a high concentration of aggregates present in this heated sample.

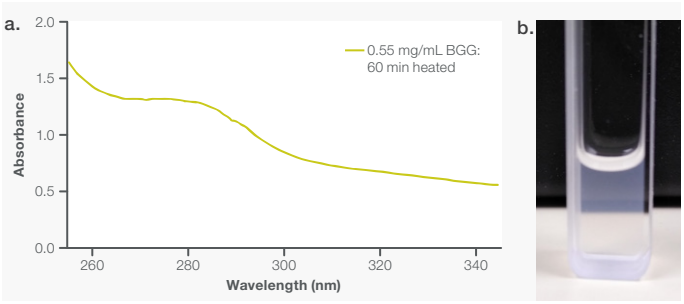


Figure 4. Absorption spectrum of 0.55 mg/mL BGG following a 60-minute incubation at 75°C.

As mentioned previously, the small amount of light reaching the detector makes it difficult to mathematically correct for scattering. Instead, an integrating sphere can be used—this accessory allows for the collection of scattered light diffusely reflected off the inner walls of the sphere. As the diffuse light reflects many times, it can be uniformly collected, removing the scattering artifact. To correct for the scatter shown in Figure 4a, a spectrum for the aggregated BGG sample (Table 1, Sample 3) was collected using an Evolution ISA-220 Accessory. Through the instrument software, the signal was reported using Kubelka-Munk units, F(R), which is proportional to both the absorption coefficient, k, and scattering coefficient, s, of the material:

$$F(R)=k/s \quad (7)$$

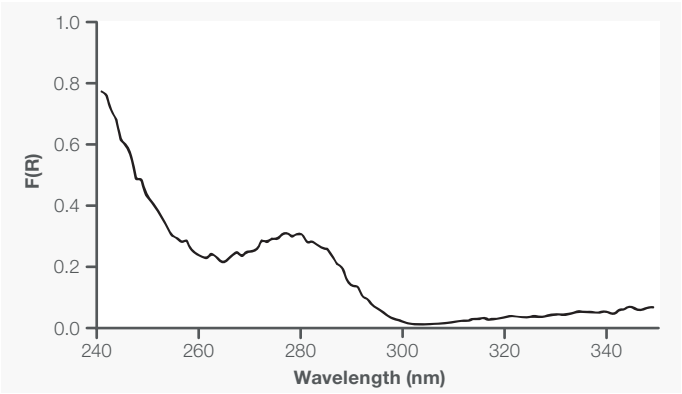


Figure 5. Kubelka-Munk spectrum of 0.55 mg/mL BGG after a 60-minute incubation at 75˚C.

Figure 5 demonstrates the Kubelka-Munk spectrum of the BGG solution shown in Figure 4; the scattering signal is largely removed from the spectrum.

F(R) is not equivalent to absorbance, indicating Beer's law cannot be used to determine concentration from the collected results. However, as F(R) is proportional to the absorption coefficient, it is also proportional to the absorbance, A, and the concentration, c, of the free proteins in solution:

$$F(R) \propto A \propto c. \quad (8)$$

To determine the concentration of aggregated and non-aggregated proteins in solution using the Kubelka-Munk formula, the fully non-aggregated sample (control) was measured using the integrating sphere. The resulting Kubelka-Munk spectrum collected is shown in Figure 6a (gray curve). A second BGG sample heated to 75˚C for 30 minutes (Table 1, Sample 4), which also resulted in a large scattering artifact, was analyzed using the Evolution ISA-220 Accessory as well.

If the collected F(R) of the sample at a given wavelength is assumed to be equivalent to the concentration of the proteins in solution multiplied by some constant, b, that is shared between all BGG samples, then we can construct a series of equations:

$$F_{control}(R) = c_{control} b \quad (9)$$

$$F_{sample}(R) = c_{sample} b \quad (10)$$

$$c_{sample} = c_{control} * F_{sample}(R)/F_{control}(R) \quad (11)$$

The equations above can be used to relate the concentration of non-aggregated BGG in the sample that was incubated at 75˚C (c_{sample}) to the concentration of the non-aggregated BGG control ($c_{control}$), the Kubelka-Munk signal of the sample ($F_{sample}(R)$), and the control ($F_{control}(R)$). For more complex samples, constructing a standard curve with multiple control samples of differing concentration would be a more effective analysis tool.

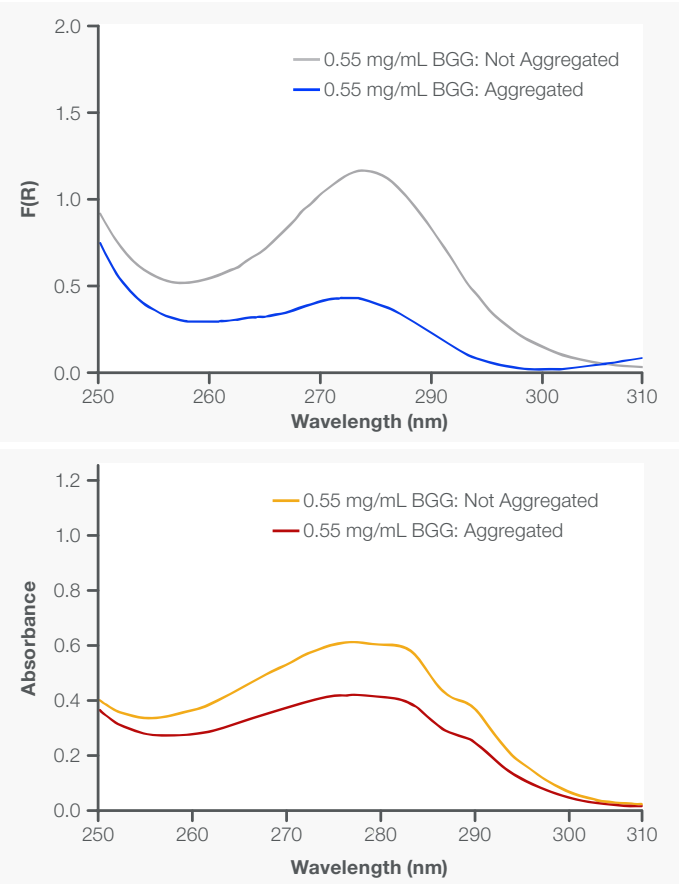


Figure 6. a) Kubelka-Munk spectra of 0.55 mg/mL BGG after a 30-minute incubation at 75˚C (blue) and 0.55 mg/mL non-aggregated BGG (gray). b) Absorption spectra of filtered 0.55 mg/mL BGG after a 30-minute incubation at 75˚C (brown) and 0.55 mg/mL non-aggregated BGG (orange). The incubated BGG sample was filtered using a Millipore Millex-GV PVDFA filter.

Using Equation 11, the concentration of non-aggregated materials in the BGG sample was found to be 0.20 mg/mL, implying 0.35 mg/mL of BGG contributed to the formation of aggregates in this sample. To verify this equation, the BGG sample containing aggregates was filtered using a syringe filter and the absorption spectrum of the filtrate was collected using a traditional cell holder. Using Beer's law, the concentration of the BGG filtrate was found to be 0.20 mg/mL, matching the calculated concentration determined using the integrating sphere. This further implies that BGG aggregates in solution do not absorb an appreciable amount of light in the spectral region of interest for this sample.

Conclusion

Protein aggregates in solution can quickly be detected using the Evolution One Plus UV-Visible Spectrophotometer. For samples with a low concentration of aggregate present, the resulting scattering artifact can be corrected by estimating the scattering contribution and subtracting that estimate from the measured spectrum. For highly scattering solutions, the Evolution ISA-220 Integrating Sphere Accessory works well in removing the scattering artifact from the spectrum. The concentration of free proteins in solution can then be solved for the corresponding spectrum of a known standard or a series of known standards.

References

1. Weids, A.J., Ibstedt, S., Tamás, M.J.; Grant, C.M., Sci. Rep. 2016, 6, 24554.
2. Chen, Z.; Huang, C.; Chennamsetty, N; Xu, X.; Li, Z.J., J. Chromatogr. A, 2016, 1460, 110 – 122.
3. Jubete, Y.; Maurizi, M.R., Gottesman, S., J. Biol. Chem., 1996, 271, 48, 30798-30803.
4. Da Vela, S.; Roosen-Runge, F.; Skoda, M.W.A.; Jacobs, R.M.J.; Seydel, T.; Frielinghaus, H.; Sztucki, M.; Schweins, R.; Zhang, F.; Schreiber, F., J. Phys. Chem. B 2017, 121, 23, 5759 – 5769.
5. Pignataro, M.F.; Herrera, M.G.; Dodero, V.I.; Molecules 2020, 25, 20, 4854.
6. Johnson, I.S., Science, 1983, 219, 4585, 632 – 637.
7. Mitragotri, S.; Burke, P.A.; Langer, R., Nat. Rev. Drug Discov., 2014, 13, 9, 655 – 672.
8. Lundahl, M.L.E.; Fogli, S.; Colavita, P.E.; Scanlan, E.M., RSC Chem. Biol., 2021, 2, 1004-1020.
9. Hawe, A.; Friess, W.; Sutter, M.; Jiskoot, W., Anal. Biochem. 2008, 38, 115 – 122.
10. Hall, D.; Zhao, R.; Dehlsen, I.; Bloomfield, N.; Williams, S.R.; Arisaka, F.; Goto, Y.; Carver, J.A., Anal. Biochem., 2016, 489, 78 – 94.
11. Smith, E.L.; Coy, N., J. Biol. Chem., 1946, 164, 1, 367 – 370.
12. Yao, G.; Li, K.A.; Tong, S.Y., Anal. Chim. Acta, 1999, 398, 319 – 327.

Application note

Evaluating DNA purity for molecular cloning quality control

NanoDrop Lite Plus Spectrophotometer

Molecular cloning requires two main components to create recombinant DNA: 1) a DNA vector; and 2) one, or several, DNA fragment(s) containing the gene(s) of interest (GOI). To generate a fragment containing a GOI, digestion with a restriction endonuclease cuts the DNA at specific sites and yields base pairing complementarity with a similarly digested vector through their cohesive ends (Figure 1A). The vector and the DNA fragment are then joined at their cohesive ends by a ligase enzyme to form the recombinant DNA product (Figure 1B).[1-3]

Quality control (QC) in molecular cloning is important before and after the restriction digestion step. Once extracted from the cell line, plasmid DNA purity should be checked before beginning the digestion. Common contaminants from DNA extraction include phenol, ethanol, and salts, which are known to inhibit the restriction endonuclease and prevent proper cleavage.[4] After digestion, the efficiency can be evaluated with agarose gel electrophoresis by verifying the presence of expected bands and little to no smearing.[5]

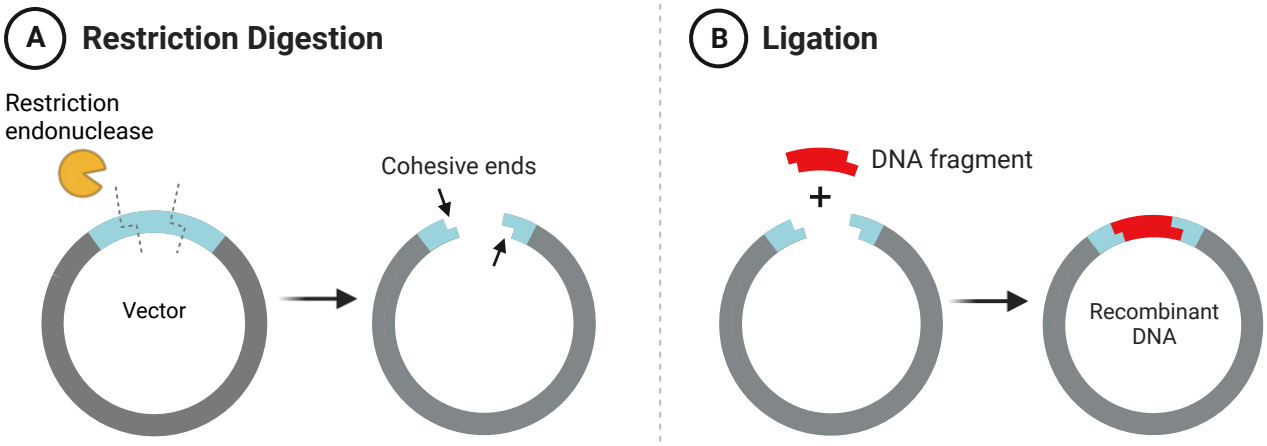


Figure 1. A) Restriction digestion of a vector with an endonuclease to produce cohesive ends for complementary binding with a DNA fragment containing a gene of interest. B) Ligation of the digested DNA fragment and vector to produce recombinant DNA. Figure created with BioRender.com.

Using restriction endonucleases produces many fragments of varying lengths when attempting to clone a single gene from genomic DNA.[3] For this reason, restriction digested DNA is commonly analyzed via gel electrophoresis after digestion, and the correct fragment band containing the GOI can be excised and purified from the gel for downstream ligation and transformation. Purity and concentration should also be determined after gel purification to ensure highly efficient ligation and transformation.

The purity and concentration of extracted plasmid DNA is typically evaluated with ultraviolet-visible (UV-Vis) spectrophotometry as it is a quick and simple technique. DNA absorbs light at 260 nm in the UV range, while salts absorb below 230 nm and proteins and phenolics absorb around 280 nm. The measured absorbance can be related to the concentration using the Beer-Lambert Law, shown in the following equation, where "c" = concentration, "A" = absorbance at a 1.0 cm pathlength, "ε" = sample-specific extinction coefficient, and "b" = pathlength (typically 1.0 cm):

$$A = \varepsilon bc$$

The purity ratios, A260/A230 and A260/A280, are lowered by the presence of salts and proteins, respectively. This makes the ratios key tools for assessing purity of DNA. For dsDNA, the expected A260/A280 ratio is ~1.8 and the expected A260/A230 ratio range is 2.0 - 2.2.

The Thermo Scientific™ NanoDrop™ Lite Plus Microvolume UV Spectrophotometer can provide purity ratios and DNA concentrations for samples of just 1.0 – 2.0 µL volumes. The NanoDrop instrument's ability to analyze such small volumes and allow conservation of sample material for downstream experiments is critical, since most extractions elute or resuspend DNA in volumes below 50 µL.

Experimental methods

Three samples of plasmid pUC19 DNA (Thermo Scientific, SD0061) were prepared. Sample 1 was pure pUC19 and samples 2 and 3 were pUC19 spiked with 150 ppm phenol and 20 mM EDTA, respectively, to mimic the contamination from a typical DNA extraction. The concentration, A260/A280 purity ratio, and A260/A230 purity ratio of all three samples were determined using 2.0 µL volumes on a NanoDrop Lite Plus spectrophotometer using the dsDNA sample type.

The contamination effect on restriction endonuclease cleavage was evaluated by digesting both pure and contaminated pUC19 with HindIII (Thermo Scientific, ER0501) and incubating for one and a half hours at 37°C. After incubation, HindIII was inactivated by heating at 80°C for 20 minutes. HindIII has one cleavage site in the pUC19 sequence, producing a linear plasmid of 2686 base pairs upon digestion (Figure 2).

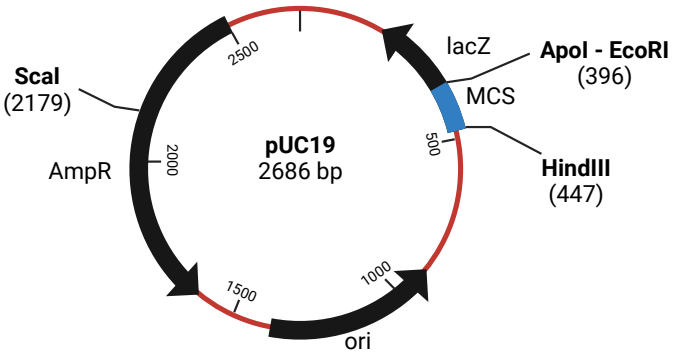


Figure 2. Plasmid map of pUC19 with several restriction endonuclease cleavage sites. Figure created with BioRender.com.

Gel electrophoresis of the digested pUC19 samples was performed with a 1.2% agarose gel to evaluate the endonuclease digestion efficiency and to confirm that the expected fragment length of 2686 bp was formed. The concentration of DNA loaded per well was equivalent to about 20 ng/µL. The bands on the gel corresponding to the digested plasmids were excised with a sterile scalpel. Excess gel was removed from each sample to yield 50 mg for downstream extraction.

Plasmid DNA was extracted from the three gel-excised samples by following manufacturer's instructions from the Invitrogen™ PureLink™ Quick Gel Extraction Kit (Invitrogen, K210012). Two minor changes were applied to the manufacturer's instructions to improve yield: 1) warming the elution buffer to 65°C before loading on the column; and 2) incubating the elution buffer on the column for 10 minutes before elution. The eluted DNA was analyzed for concentration and purity with the NanoDrop Lite Plus spectrophotometer using the dsDNA sample type.

	Concentration (ng/µL)	Standard Deviation (ng/µL)	A260/A280	A260/A230
Pure pUC19	241.8	0.6	1.90	2.03
pUC19 + Phenol (150 ppm)	358.3	0.5	1.77	1.76
pUC19 + EDTA (20 mM)	231	3	2.04	0.28

Table 1. Concentration and purity results of pure pUC19 and contaminated pUC19 measured in replicates of five on the NanoDrop Lite Plus spectrophotometer.

Results

The concentration and purity results of the pure and contaminated pUC19 samples determined by the NanoDrop Lite Plus instrument are outlined in Table 1. Pure pUC19 displayed an average concentration of 241.8 ng/µL and purity ratios in the expected range for dsDNA. Phenol contamination caused an increase in the reported concentration to 358.3 ng/µL due to the additional absorbance contribution of phenol at 260 nm.

The A260/A230 purity ratio of 1.76 fell below the expected range as phenol also contributes to increased absorbance below 230 nm. EDTA contamination increased the A260/A280 ratio and significantly reduced the A260/A230 ratio. Since salts are highly absorbing below 230 nm, contamination is clearly revealed in the A260/A230 purity ratio of 0.28.

When the purity ratios are outside of the expected range, the concentration should be considered unreliable due to the contamination effect on absorbance. At this stage, samples should be further purified with a column extraction or alcohol precipitation to ensure an accurate concentration measurement.

Figure 3 displays the gel electrophoresis results on the digestion efficiency of HindIII in the presence of EDTA and phenol contamination. Undigested pUC19 in column B remains in its supercoiled topology and travels further down the gel than its linear counterpart in column C.[6] The linear pUC19 control band was at the expected 2686 bp location in relation to the ladder in column A. Phenol contamination (column D) had a minimal effect on the digestion efficiency but higher concentrations of phenol or other organic solvents can further inhibit the restriction enzyme.[7-8] EDTA contamination inhibited HindIII digestion, as shown by the alignment of the band in column E with the undigested control in B. The results of the digestion confirm the importance of incorporating a purity checkpoint prior to running a restriction endonuclease digestion.

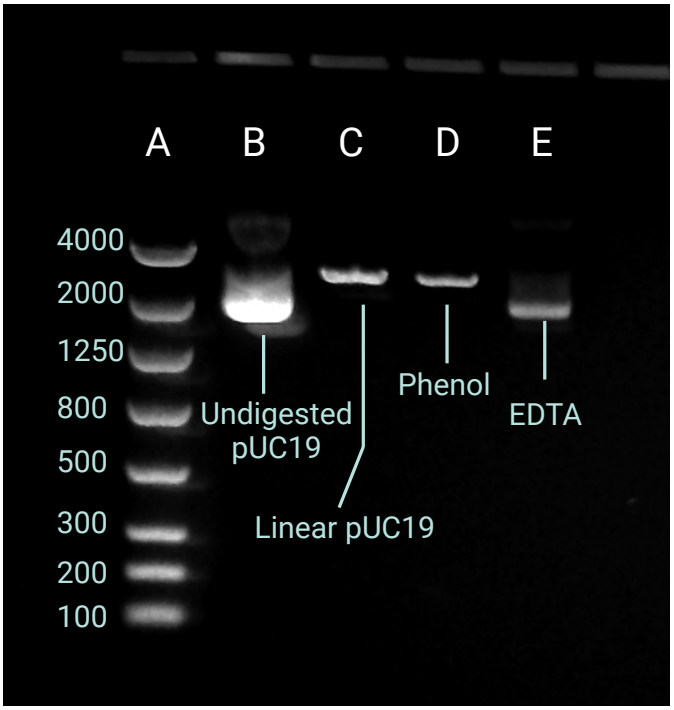


Figure 3. Gel electrophoresis image capture of supercoiled, undigested pUC19 (Lane B); linear pUC19 digested with HindIII (Lane C); pUC19 spiked with phenol (150 ppm) and HindIII digested (Lane D); pUC19 spiked with EDTA (20 mM) and HindIII digested (Lane E). Ran on a 1.2% agarose gel. Ladder in Lane A displayed as base pairs.

After performing a gel extraction of the three samples digested with HindIII, the purified samples were measured again on the NanoDrop Lite Plus spectrophotometer to calculate the sample concentration, percent recovery, and purity ratios. These results are presented in Table 2. The pure pUC19 sample was washed with wash buffer once while the phenol and EDTA contaminated samples were washed twice, which reduced the percent recovered DNA. The average concentration of recovered DNA from all samples was 15.04 ng/µL with a recovery of up to 89%. The A260/A280 and A260/A230 purity ratios for all samples were within the acceptable range except for the EDTA contaminated pUC19. With the reduced A260/A230 purity ratio, this is indicative of residual EDTA or guanidine salt from the extraction reagents. Incorporating isopropanol prior to loading on the extraction column aids in precipitating DNA away from salts and would lead to a purified DNA sample.[9]

Conclusions

The molecular cloning workflow requires QC checkpoints before and after the restriction endonuclease digestion to reduce the failure of downstream reactions. Contaminants such as phenol and salts have been shown to inhibit or reduce the efficiency of endonucleases, which highlights the need for performing a purity check before digestion. After digestion and gel extraction, purity should again be checked for ensuring successful ligation and transformation. With the NanoDrop Lite Plus spectrophotometer, the A260/A280 and A260/A230 purity ratios provide a fast and easy method for completing QC steps without the need for dilutions that require large volumes of extracted DNA.

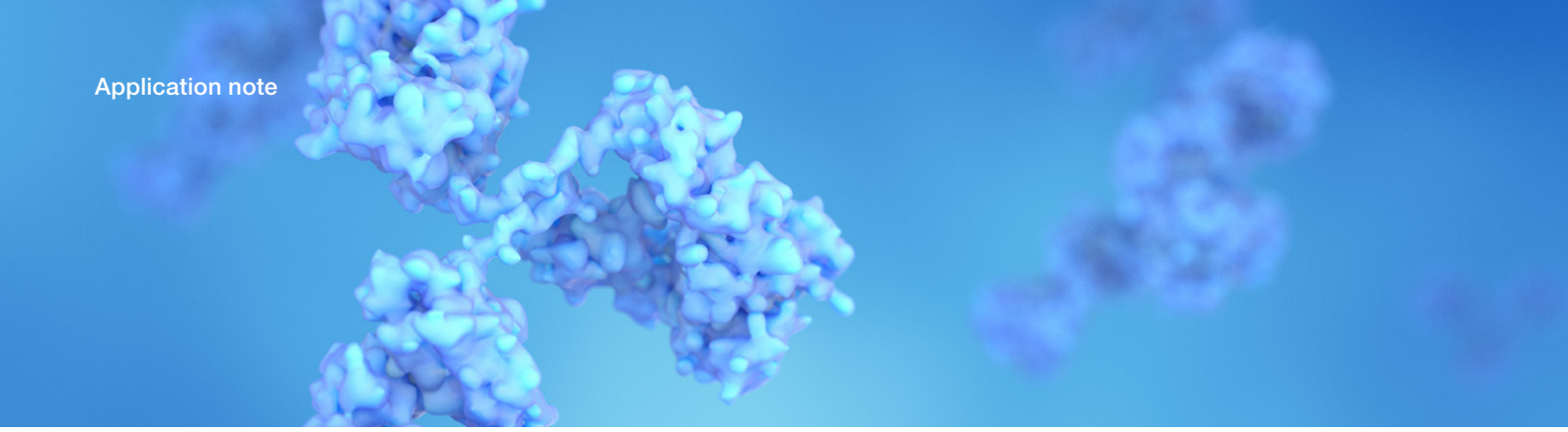
	Concentration (ng/µL)	Standard Deviation (ng/µL)	% Recovery from Gel	A260/A280	A260/A230
Pure pUC19	17.83	0.09	89%	1.92	2.10
pUC19 + Phenol (150 ppm)	13.5	0.3	67.5%	1.84	2.04
pUC19 + EDTA (20 mM)	13.8	0.2	69%	1.95	1.69

Table 2. Concentration and purity results of gel extracted pUC19 measured in replicates of five on the NanoDrop Lite Plus spectrophotometer.

References

1. Cooper, G. M. (2000). Recombinant DNA. In The Cell: A Molecular Approach. 2nd edition. Sinauer Associates. ncbi.nlm.nih.gov/books/NBK9950/
2. Mather, M. W., Keightley, J. A., & Fee, J. A. (1993). Recovery and cloning of genomic DNA fragments from dried agarose gels. Methods in Enzymology, 218, 695–704. doi.org/10.1016/0076-6879(93)18052-e
3. Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., & Walter, P. (2002). Isolating, Cloning, and Sequencing DNA. In Molecular Biology of the Cell. 4th edition. Garland Science. ncbi.nlm.nih.gov/books/NBK26837/
4. Optimizing Restriction Endonuclease Reactions | NEB. Retrieved January 8, 2024, from neb.com/en-us/protocols/2012/12/07/optimizing-restriction-endonuclease-reactions
5. Gel Electrophoresis Applications. Thermo Fisher Scientific. Retrieved December 12, 2023, from thermofisher.com/us/en/home/life-science/cloning/cloning-learning-center/invitrogen-school-of-molecular-biology/na-electrophoresis-education/na-electrophoresis-applications.html
6. Bates, A. D., & Maxwell, A. (2005). DNA Topology. Oxford University Press.
7. Using NanoDrop One/OneC to determine phenol and protein contaminants in nucleic acids for RT-qPCR quality control. Thermo Fisher Scientific. Retrieved January 29, 2024, from assets.thermofisher.com/TFS-Assets/MSD/Application-Notes/nanodrop-one-phenol-protein-contaminants-rtqpcr-en-tn53472.pdf
8. Tan, S. C., & Yiap, B. C. (2009). DNA, RNA, and Protein Extraction: The Past and The Present. Journal of Biomedicine and Biotechnology, 2009, 574398. doi.org/10.1155/2009/574398
9. Green, M. R., & Sambrook, J. (2017). Precipitation of DNA with Isopropanol. Cold Spring Harbor Protocols, 2017(8), pdb.prot093385. doi.org/10.1101/pdb.prot093385

Application note



Quantify protein and peptide preparations at 205 nm

NanoDrop One Spectrophotometer

Figure 1. NanoDrop One Proteins Home screen showing available preprogrammed applications for protein quantitation.

Abstract

The presented data herein was performed on a Thermo Scientific™ NanoDrop One™ Spectrophotometer and has since been replaced by the Thermo Scientific™ NanoDrop Ultra™ Spectrophotometer and Fluorometer with enhanced detection limits, using the A205 preprogrammed direct absorbance application. The new A205 application offers a choice of methods for peptides that contain Tryptophan and Tyrosine residues in their sequence as well as peptides that completely lack aromatic amino acids. The A205 application offers enhanced sensitivity for peptide quantification in seconds from only 2 µL of sample.

Introduction

Researchers have always needed ways to quickly quantify various biomolecules (e.g., protein and nucleic acid preparations) as a routine part of their workflows. This information helps them make informed decisions before proceeding with downstream experiments. There are many protein quantification methods to choose from including gravimetric approaches, colorimetric assays, direct spectrophotometric UV measurements (such as A280), and amino acid analysis. All of these methods have their strengths and weaknesses. Direct spectrophotometric microvolume UV measurements are a popular choice for researchers because they are simple to perform, require no reagents or standards, and consume very little sample. The NanoDrop One Spectrophotometer has preprogrammed applications (Figure 1) for direct quantification of proteins using absorbance measurements at 280 nm and 205 nm. This application note specifically describes how to use the Protein A205 application to quantify protein samples.

A protein's peptide backbone absorbs light in the deep UV region (190 nm-220 nm), and this absorbance can be used for protein sample quantitation. The A205 protein quantitation method has several advantages over the direct A280 protein method such as lower protein-to-protein variability (because A205 extinction coefficients are not based on amino acid composition) and higher sensitivity (because of the high molar absorptivity proteins have at 205 nm). However, technical limitations made it difficult to obtain these measurements in the past. Spectrometers' stray light performance, deep UV linearity, and protein buffers containing UV-absorbing components have all added to the challenge of obtaining A205 data. The NanoDrop One patented sample-rentention technology and low stray light performance have simplified quantification of small amounts of protein by A205 methods.

In this application note, we discuss the three A205 measurement options included in the NanoDrop One Protein A205 application and present performance data for each option.

A205 extinction coefficients for peptide and protein measurements

The Thermo Scientific™ NanoDrop One™ Protein A205 application allows customers to choose from three different options (Figure 2). The selected option will automatically determine the extinction coefficient that will be used to calculate the protein concentration based on the sample absorbance at 205 nm.

- ε_{205}=31 method
- Scopes method[2]
- Other = custom method $\varepsilon_{205}^{1mg/mL}$

Previous studies showed that most protein solutions at 1 mg/mL have extinction coefficients ($\varepsilon_{205}^{1mg/mL}$) ranging from 30 to 35[2]. The ε_{205} of 31 mL $mg^{-1}cm^{-1}$ is an extinction coefficient often used for peptides lacking tryptophan and tyrosine residues[1]. The Scopes method gives a more accurate ε_{205}, especially for proteins containing a significant amount of tryptophan (Trp) and tyrosine (Tyr) residues. The increased accuracy of this method takes into account the significant absorbance at 205 nm contributed by the aromatic side chains of Trp and Tyr. This method uses an A280/A205 ratio in its equation to correct for Trp and Tyr side-chain absorbance[3]. Recently, Anthis and Clore proposed the use of a sequence-specific ε_{205} calculation (e.g., custom/Other method), which is suitable for a wide range of proteins and peptides[1]. This method is appropriate for pure preparations of proteins or peptides whose amino acid sequences are known.

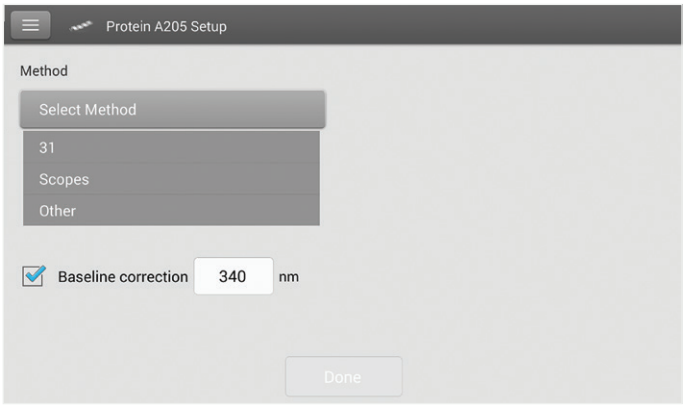


Figure 2. NanoDrop One Protein A205 methods selection screen.

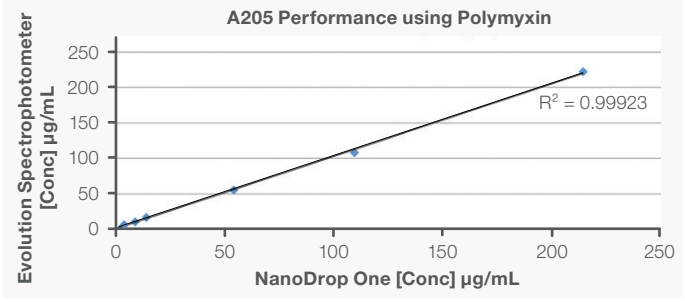


Figure 3. Polymyxin concentrations calculated with the Evolution Spectrophotometer and NanoDrop One instruments were plotted. Regression line shows that protein concentration measurements on the NanoDrop One instrument are in good agreement to those obtained on a traditional high end UV-Vis spectrophotometer.

A205 performance on the NanoDrop One

Preparations of polymyxin, a cationic detergent antibiotic with a peptide backbone, but no Trp or Tyr residues, were made in 0.01% Brij™ 35 buffer and measured on the NanoDrop One and the Thermo Scientific™ Evolution™ UV-Vis Spectrophotometers. To ensure the validity of the measurements taken with the Evolution Spectrophotometer instrument, the polymyxin preparations were diluted in 0.01% Brij buffer to ensure that the measurements taken were within the linear range of the detector. For measurements on the NanoDrop One instrument 2 µL of sample were pipetted directly on the sample pedestal, while a 10 mm quartz cuvette was used for measurements on the Evolution Spectrophotometer. The polymyxin concentration data obtained on both instruments (Table 1) were plotted (Figure 3). Regression line shows that protein concentration results from the NanoDrop One instrument are in good agreement to the results obtained on a traditional high end UV-Vis spectrophotometer using a cuvette.

Target [Conc] mg/mL	NanoDrop One			Evolution
	A205	Std. Dev.	[Conc] mg/mL	[Conc] mg/mL
0	-0.01	0.04	-0.18	-0.02
5	0.11	0.01	3.60	5.05
10	0.27	0.01	8.84	10.53
15	0.44	0.02	14.08	17.09
50	1.68	0.01	54.14	55.32
100	3.39	0.01	109.44	108.48
200	6.64	0.03	214.16	222.50

Table 1. Various preparations of Polymyxin were measured on the NanoDrop One and Evolution Spectrophotometers. Five(5) replicates of each solution were measured on the NanoDrop One instrument using the 205=31 application. Solutions with absorbance over 1.0A were diluted and measured in triplicate on the Evolution Spectrophotometer instrument.

Protein preparation	# of Trp of Tyr					
	Trp	Tyr	A205	STDV	[Concentration] ε_{205}=31 (µg/mL)	[Concentration] Scopes method (µg/mL)
BSA 1	3	21	3.960	0.013	127.73	131.80
BSA 2	3	21	37.271	0.218	1202.30	1261.71
BSA 3	3	21	70.044	0.239	2259.48	2387.91
BSA 4	3	21	129.170	1.458	4166.77	4345.20
BSA 5	3	21	271.027	0.851	8742.81	9198.13
Lysozyme 1	6	3	29.069	0.169	937.71	795.95
Lysozyme 2	6	3	53.651	0.545	1730.68	1459.05
Lysozyme 3	6	3	102.713	0.668	3313.32	2814.79
Polymyxin 1	0	0	0.112	0.015	3.60	3.12
Polymyxin 2	0	0	0.274	0.014	8.84	10.12
Polymyxin 3	0	0	0.437	0.021	14.08	16.03
Polymyxin 4	0	0	1.678	0.014	54.14	60.99
Polymyxin 5	0	0	3.393	0.014	109.44	125.16
Polymyxin 6	0	0	6.639	0.034	214.16	244.87

Table 2. Comparison of different A205 methods for various protein and peptide preparations on the NanoDrop One Spectrophotometer.

To assess the effect that the extinction coefficients used at 205 nm (i.e., Scopes and ε_{205}=31 methods) would have on the result, we prepared dilutions of three different proteins with varied amounts of aromatic residues: bovine serum albumin (BSA, 3 Trp and 21 Tyr residues), lysozyme (6 Trp and 3 Tyr residues) and polymyxin (no Trp, no Tyr). These preparations were measured on the NanoDrop One instrument using the ε_{205}=31 and Scopes methods (Table 2).

Conclusion

To assess NanoDrop One Spectrophotometer performance at A205, we compared polymyxin concentration results obtained with the NanoDrop One and the Evolution benchtop Spectrophotometers, which have excellent stray light performance. Table 1 shows that the NanoDrop One instrument provided very consistent results between replicate measurements at 205 nm with standard deviations below 0.04A. In addition, the results obtained with both instruments were comparable (Figure 3). Comparison between the A205 methods (Scopes and ε_{205}=31 methods) offered in the NanoDrop One A205 application shows that the number of tryptophan and tyrosine residues has a large effect on the calculated concentration (Table 2). This is because tryptophan is the largest contributor to A280 absorbance, and the Scopes method uses the A280/A205 ratio to correct for aromatic side-chain absorbance at A205.

Our results show that A205 quantification using the ε_{205}=31 method gives comparable results when proteins have only a few tryptophan residues.

One limitation of the A205 method is that many of protein buffers commonly used have absorbance at 205 nm. Before using this technique, we recommend checking the protein buffer for any contribution to the absorbance at 205 nm.

References

1. Anthis, NJ and Clore, GM 2013. Sequence-specific determination of protein and peptide concentrations by absorbance at 205 nm. *Protein Science* 22:851-858.
2. Goldfarb, AR, Saidel, LJ, Mosovich E 1951. The ultraviolet absorption spectra of proteins. *Journal of Biological Chemistry* 193(1):397-404.
3. Scopes, RK 1974. Measurement of protein by spectrophotometry a 205 nm. *Analytical Biochemistry* 59:277-282.

Application note

Observation of gold nanoshell plasmon resonance shifts after bioconjugation

Using the NanoDrop One Microvolume UV-Vis Spectrophotometer

Authors

Kejian Li[1], Megan N. Dang[1], Alexis B. Duffy[1] and Emily S. Day[1,2,3]

[1] University of Delaware, Dept. of Biomedical Engineering, Newark, DE, USA

[2] University of Delaware, Department of Materials Science and Engineering, Newark, DE, USA

[3] Helen F. Graham Cancer Center and Research Institute, Newark, DE, USA

RNA interference (RNAi)-based therapy has shown great potential in improving the study and treatment of diseases whose genetic underpinnings are known. However, challenges such as susceptibility to nuclease degradation, low cellular uptake, or rapid clearance from circulation impede the successful preclinical and clinical application of RNAi therapeutics.[1] To overcome these limitations, small interfering RNAs (siRNAs) or microRNAs (miRNAs) can be conjugated to nanoparticles (NPs), such as nanoshells (NS), to improve their stability, cellular uptake, and blood circulation time, thus resulting in increased effectiveness.[2, 3, 4]

Prior to using RNA-NP conjugates in therapeutic applications, it is critical to confirm successful RNA conjugation to the NP. One common method to confirm molecule loading onto gold-based NPs involves evaluating the surface plasmon resonance (SPR) spectra of the NPs before and after functionalization; successful RNA attachment will typically cause a slight red-shift in the peak SPR wavelength. Traditionally, UV-Vis spectrophotometers are used to analyze the optical properties of gold-based NPs. For example, the peak absorbance can be utilized to determine NP concentration via Beer's Law and to evaluate changes due to any surface modification. However, conventional cuvette-based UV-Vis spectrophotometers have limited linear range due to the use of a standard fixed pathlength (10 mm) cuvette, and they often require relatively large sample volumes (ranging from 0.5 mL to 3 mL). This is not ideal for conserving precious samples such as NPs coated with expensive RNA molecules. Furthermore, the need to dilute samples to fit the operating range of the instrument is time-consuming and increases the likelihood for inaccurate measurements. Alternative measurement techniques that require less volume and allow analysis of concentrated samples without dilution would be ideal.

The presented data herein was performed on a Thermo Scientific™ NanoDrop One™ Spectrophotometer and has since been replaced by the Thermo Scientific™ NanoDrop Ultra™ Spectrophotometer and Fluorometer with enhanced detection limits. The NanoDrop One has shown that it can be used to accurately measure highly concentrated NP samples without dilution, owing to its surface tension system and auto-ranging pathlength technique.[5, 6] For example, 150 nm diameter NS can be measured at concentrations up to 100 pM with high reproducibility.[5] In this application note, the use of the Nanodrop One instrument to observe shifts in the SPR of NS after conjugation to thiol-modified siRNA duplexes and methoxy-poly(ethylene glycol)-thiol (mPEG-SH; a passivating agent) was investigated. The results indicate that the Nanodrop One instrument can serve as a microvolume alternative to traditional cuvette-based spectrophotometers for qualitatively confirming RNA and PEG loading on gold-based NPs via plasmon resonance shifts.

Experimental procedures

NS were synthesized by published protocols via the Oldenburg method.[7] First, 3-5 nm diameter gold colloid was made by the Duff method[8] from hydrogen tetrachloroaurate (III) hydrate (HAuCl4) (VWR), tetrakis(hydroxymethyl)phosphonium chloride (VWR), and 1 N sodium hydroxide (Fisher Scientific). The gold colloid was then combined with 120 nm diameter silica spheres functionalized with 3-aminopropyltriethoxysilane (Nanocomposix) and 1 M sodium chloride (NaCl) and rocked for 3-4 days at room temperature to create "seed" nanoparticles. The seed was purified twice via centrifugation at 3000 rpm for 30 minutes each and resuspended in Milli-Q™ water (Sigma) to an optical density at 530 nm (OD530nm) of 0.1, as determined using a cuvette-based UV-Vis spectrophotometer. The diluted seed was mixed with additional HAuCl4 diluted in potassium chloride followed by addition of a small volume of 37% formaldehyde (VWR). The mixed solution was rapidly agitated to form complete gold shells and purified twice via centrifugation at 500 g for 15 minutes each. Additionally, NS were treated with 0.1% diethyl pyrocarbonate (DEPC) (Sigma) for 3 days rocking at 37°C to render the NS RNase-free. All materials described were purchased or treated with DEPC to be RNase-free prior to use.

siRNA oligonucleotides were purchased as single strands from IDT DNA, with sequences listed in Table 1. Thiolated sense strands were mixed with complementary non-thiolated antisense strands in equimolar amounts, boiled at 95° C for 5 min in a thermomixer, and then slowly cooled to 37° C over 1 hour to facilitate siRNA duplexing. RNase-free NS were diluted to OD800nm = 1.5 in Milli-Q water (as measured on a cuvette-based spectrophotometer). Next, 10% Tween-20 and 5 M NaCl were added to final concentrations of 0.2% and 12 mM, respectively, and the NS incubated for 5 min at room temperature. Then, siRNA duplexes were added to a final concentration of 200 nM, and the solution was bath sonicated and rocked at 4° C for 3 hours. NaCl was then added incrementally to a final concentration of 400 mM prior to rocking overnight at 4° C. The following day, 5 kDa mPEG-SH was diluted in Milli-Q water to 1 mM and added to NS to a final concentration of 10 µM. After rocking for 4 hours at 4° C, the NS solution was purified via centrifugation at 500 g for 5 minutes 3 times, resuspended in RNase-Free 1X phosphate buffered saline (PBS) with 100 X less volume of the starting NS, and stored at 4° C until use.

For conventional spectrophotometry, bare NS and siRNA-NS (diluted 100-fold in water) were placed in 1-cm pathlength disposable cuvettes and analyzed on a reference UV-Vis spectrophotometer from 1,100 nm to 400 nm. The NS concentrations were calculated from Beer's Law using the peak extinction (OD at ~800 nm) as determined by the spectrophotometer and the theoretical extinction coefficient of NS with 120 nm diameter silica cores and 15 nm thick gold shells. This revealed the initial bare NS and siRNA-NS had a concentration of 6.9 pM and 150 pM, respectively. To prepare samples for measurement with the NanoDrop One Spectrophotometer, the bare NS were concentrated by centrifugation at 500 g for 15 minutes, followed by removal of the supernatant and dilution in water to 100 pM. The siRNA-NS were directly diluted in water to 100 pM. The 100 pM bare NS and siRNA-NS solutions were measured on a NanoDrop One Spectrophotometer from 850 nm to 190 nm by pipetting 2 µL aliquots directly onto the sample pedestal. Between measurements, the NanoDrop One instrument sample pedestal was cleaned using a lint-free lab wipe. The auto pathlength option was turned on in the NanoDrop One Spectrophotometer software for each measurement.

Name	Sequence
siRNA sense	GCU GAU AUU GAC GGG CAG UAU / iSpPC//iSpPC//3ThioMC3-D/
siRNA antisense	AUA CUG CCC GUC AAU AUC AGC

Table 1: siRNA sense and antisense RNA sequences used in this work, denoted 5' to 3'. iSpPC is a photo-cleavable 10-atom spacer molecule, while 3ThioMC3-D is a thiol modification that facilitates attachment to gold NS.

Results

The absorption spectra of 150 nm NS, before (bare NS) and after (siRNA-NS) conjugation to thiolated siRNA and mPEG-SH at concentrations of 100 pM are shown in Figure 1. These spectra reveal the bare NS and siRNA-NS have a peak plasmon resonance at ~795 nm and ~804 nm, respectively, which is consistent with the spectra obtained using a reference spectrophotometer. The slightly red-shifted peak post functionalization, which maintains the overall shape and intensity of the spectra, provides evidence of successful siRNA and mPEG-SH conjugation. This was corroborated by dynamic light scattering and zeta potential measurements, as well as by siRNA loading quantification via OliGreen assay.[2, 4, 9] Notably, the spectra produced by the NanoDrop One Spectrophotometer were highly accurate and reproducible. Very little sample volume (2 µL) was used in the measurement, and no dilution was required for the analysis of highly concentrated samples (100 pM).

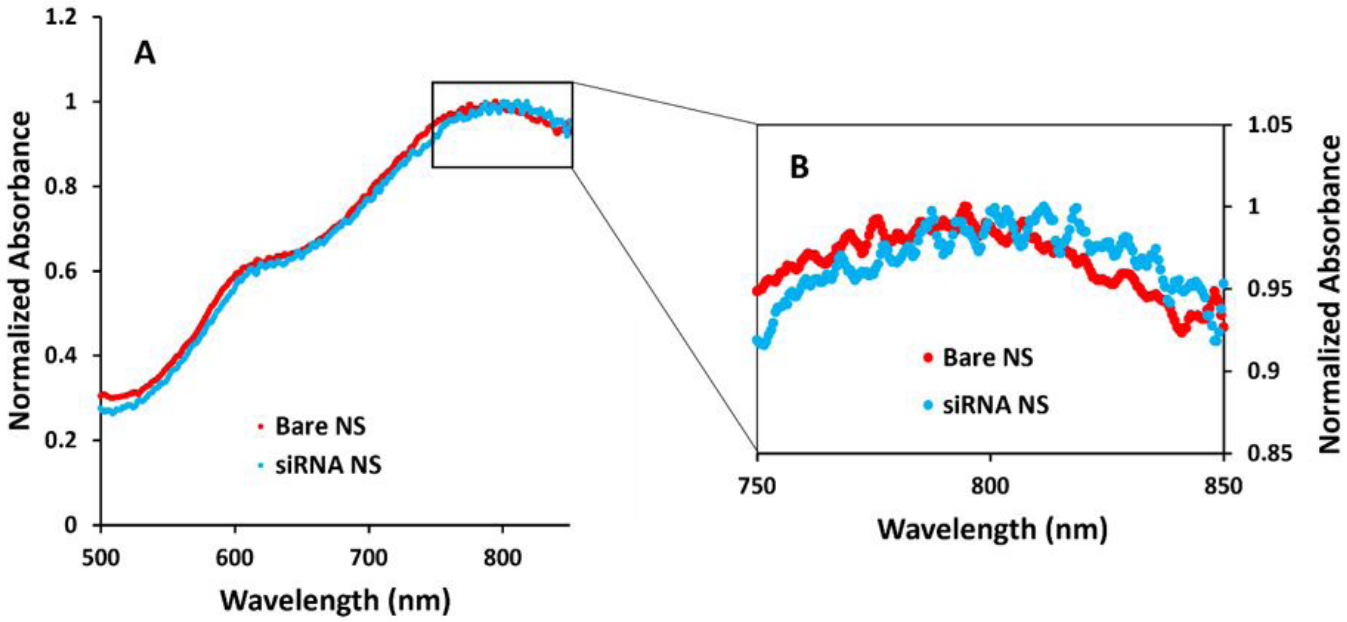


Figure 1: (A) UV-Vis spectra for Bare NS and siRNA-NS at concentrations of 100 pM, as measured on the NanoDrop One Spectrophotometer. n=3. (B) Zoom in of the UV-Vis spectra peak for Bare NS and siRNA-NS (750 nm to 850 nm).

Conclusions

This study demonstrates that the NanoDrop One Spectrophotometer can be used as a simple and reliable method to evaluate the surface modification of NS. The NanoDrop One Spectrophotometer can produce highly reliable results due to its built-in Thermo Scientific™ Acclaro™ Sample Intelligence Technology, which identifies common contaminants or other anomalies that may impact measurement accuracy. Additionally, the NanoDrop One Spectrophotometer allows the users to measure highly concentrated samples in 1–2 µL without dilution and produce full spectral data in seconds compared to a traditional cuvette-based spectrophotometer. These advantages save valuable time and money and help determine the quality and quantity of the sample before use in downstream applications. The ease of operation and small sample size requirement make the NanoDrop One Spectrophotometer an ideal and valuable instrument to characterize the properties of surface-modified NPs.

References

1. Wang, T.; Shigdar, S.; Shamaileh, H. A.; Gantier, M. P.; Yin, W.; Xiang, D.; Wang, L.; Zhou, S. F.; Hou, Y.; Wang, P.; et al. **Challenges and Opportunities for siRNA-Based Cancer Treatment.** *Cancer Lett.* 2017, 387, 77–83.
2. Riley, R. S.; Dang, M. N.; Billingsley, M. M.; Abraham, B.; Gundlach, L.; Day, E. S. **Evaluating the Mechanisms of Light-Triggered SiRNA Release from Nanoshells for Temporal Control Over Gene Regulation.** *Nano Lett.* 2018, 18, 3565–3570.
3. Artiga, Á.; Serrano-Sevilla, I.; De Matteis, L.; Mitchell, S. G.; De La Fuente, J. M. **Current status and future perspectives of gold nanoparticle vectors for siRNA delivery.** *J. Mater. Chem.* B 2019, 7, 876-896.
4. Dang, M. N.; Gomez Casas, C; Day, E. S. **Photoresponsive miR-34a/Nanoshell Conjugates Enable Light-Triggered Gene Regulation to Impair the Function of Triple-Negative Breast Cancer Cells.** *Nano Lett.* 2021, 21(1), 68-76.
5. Li, K.; Kapadia, C. H.; Dang, M. N.; Day, E. S. **Quantification of gold nanoshells using the NanoDrop One Microvolume UV-Vis Spectrophotometer.** https://www.thermofisher.com/document-connect/document-connect.html?url=https%3A%2F%2Fassets.thermofisher.com%2FTFS-Assets%2FMSD%2FApplication-Notes%2Fquantification-gold-nanoshells-nanodrop-one-uv-vis-spectrophotometer-an53464.pdf
6. Kapadia, C. H.; Melamed, J. R.; Day, E. S. **Quantification of gold nanoparticles using the NanoDrop One Microvolume UV-Vis Spectrophotometer.** http://assets.thermofisher.com/TFS-Assets/MSD/Application-Notes/AN53100-quantification-gold-nanoparticle.pdf
7. Oldenburg, S. J.; Averitt, R. D.; Westcott, S. L.; Halas, N. J. **Nanoengineering of optical resonances.** *Chem. Phys. Lett.* 1998, 288 (2–4), 243–247.
8. Duff, D. G.; Baiker, A.; Edwards, P. P. **A new hydrosol of gold clusters. 1. formation and particle size variation.** *Langmuir.* 1993, 9, 2301–2309.
9. Melamed, J. R.; Riley, R. S.; Valcourt, D. M.; Billingsley, M. M.; Kreuzberger, N. L.; Day, E. S. **Chapter 1: Quantification of siRNA duplexes Bound to Gold Nanoparticle Surfaces. In Biomedical Nanotechnology** *Humana Press*: New York, NY, USA, 2017.

Learn more at **thermofisher.com/nanodrop**

The NanoDrop Eight Spectrophotometer detects contaminating nucleic acids in mammalian DNA and RNA preparations

Introduction

Understanding nucleic acid sample quality and quantity is integral for many life science applications, reducing the occurrence of costly delays caused by troubleshooting downstream experimental failures. The Thermo Scientific™ NanoDrop™ Eight Microvolume UV-Vis Spectrophotometer measures eight samples at a time and provides you the ability to measure the concentration of biomolecules for high-throughput assays using a 1–2 µL sample size without the need for dilutions. With a measurement time of less than 20 seconds, you can easily insert the NanoDrop Eight Spectrophotometer into your high-throughput workflows.

The Thermo Scientific Acclaro™ Sample Intelligence Technology integrated within the NanoDrop Eight Spectrophotometer's software utilizes chemometrics to detect RNA in dsDNA sample preparations and dsDNA in RNA preparations to then calculate a corrected dsDNA or RNA concentration, respectively. Historically, the A260/A280 purity ratio has been utilized to assess nucleic acid sample purity; however, nucleic acid contaminants at low concentrations, such as RNA contamination in dsDNA samples, have a negligible effect on the purity ratio, and the contaminant identity is not easily determined by a change in the A260/A280 purity ratio or by visualizing the UV-Vis spectrum. Acclaro Technology's contaminant analysis capability eliminates the need for purity ratio assumptions and reports the contaminant present, contaminant absorbance, and a corrected sample concentration (Figure 1).

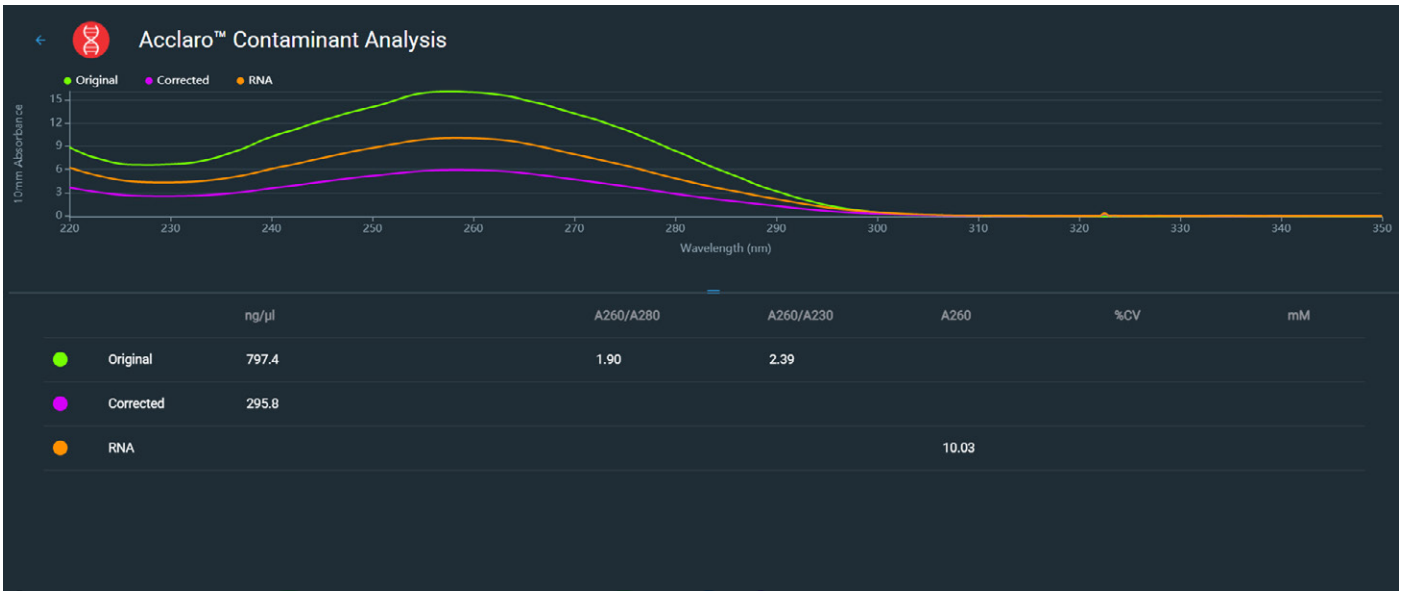


Figure 1. Acclaro Technology's contaminant analysis screen outlining the original concentration, corrected dsDNA concentration, and the absorbance contribution of RNA contamination. The original spectrum is shown in green, the corrected spectrum in pink, and the contaminating RNA spectrum in orange.

Materials and methods

Total RNA and genomic DNA from mouse tissue (BioChain Institute Inc., R1334035-50 and D1334999-G01) and RNA and genomic DNA from the MCF-7 cell line (BioChain Institute Inc., R1255830-50 and D1255830) were dialyzed and diluted in tris-EDTA buffer (TE pH 8.0, Fisher Scientific, BP2473500) and made into various DNA/RNA mixtures according to percentage of absorbance contribution. Triplicates of each mixture were measured on the NanoDrop Eight Spectrophotometer using fresh 1.0 µL aliquots per replicate for the dsDNA and RNA applications.

The NanoDrop Eight Spectrophotometer's Acclaro Technology-corrected results from the mouse and MCF-7 DNA/RNA mixtures were compared with the theoretical concentration and the original, uncorrected concentration in Figures 2 and 3 using the dsDNA and RNA applications, respectively. Acclaro Technology calculated an original, uncorrected concentration and a corrected concentration based on a modified Beer's Law equation and the absorbance contribution at 260 nm.

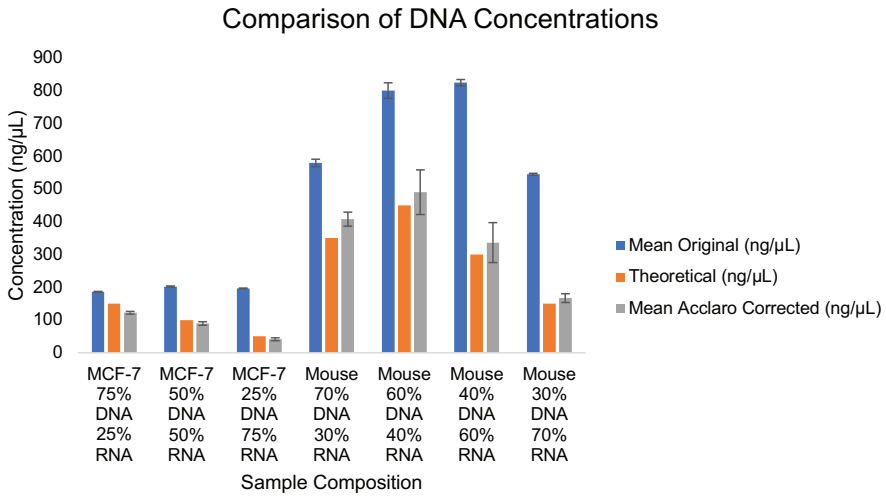


Figure 2: Comparison of the concentration reported by the Acclaro Technology for different sample compositions of DNA and RNA based on percentage of absorbance contribution. DNA and RNA from either the MCF-7 cell line or mouse tissue were mixed according to absorbance percentage and were measured using the dsDNA application. The mean original concentration (blue bars), the theoretical concentration (orange bars), and the mean Acclaro Technology software-corrected concentration (gray bars) were reported by the NanoDrop Eight Spectrophotometer's software. Error bars represent the standard deviation.

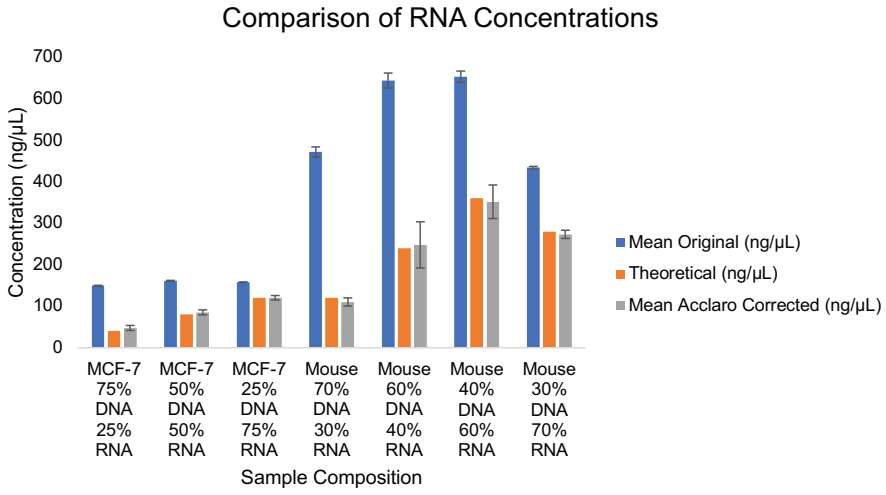


Figure 3: Comparison of the concentration reported by the Acclaro Technology for different sample compositions of DNA and RNA based on percentage of absorbance contribution. DNA and RNA from either the MCF-7 cell line or mouse tissue were mixed according to absorbance percentage and were measured using the RNA application. The mean original concentration (blue bars), the theoretical concentration (orange bars), and the mean Acclaro Technology software-corrected concentration (gray bars) were reported by the NanoDrop Eight Spectrophotometer's software. Error bars represent the standard deviation.

Results

In Figures 2 and 3, the Acclaro Technology's software-corrected mean concentration from the NanoDrop Eight Spectrophotometer was graphed against the original, uncorrected concentration and the theoretical concentration for the mouse and MCF-7 DNA/RNA mixtures with standard deviation shown as error bars. Since nucleic acids absorb at 260 nm, the original, uncorrected concentration is inflated compared to the Acclaro Technology's software-corrected concentration when DNA and RNA are both contributing to absorbance.

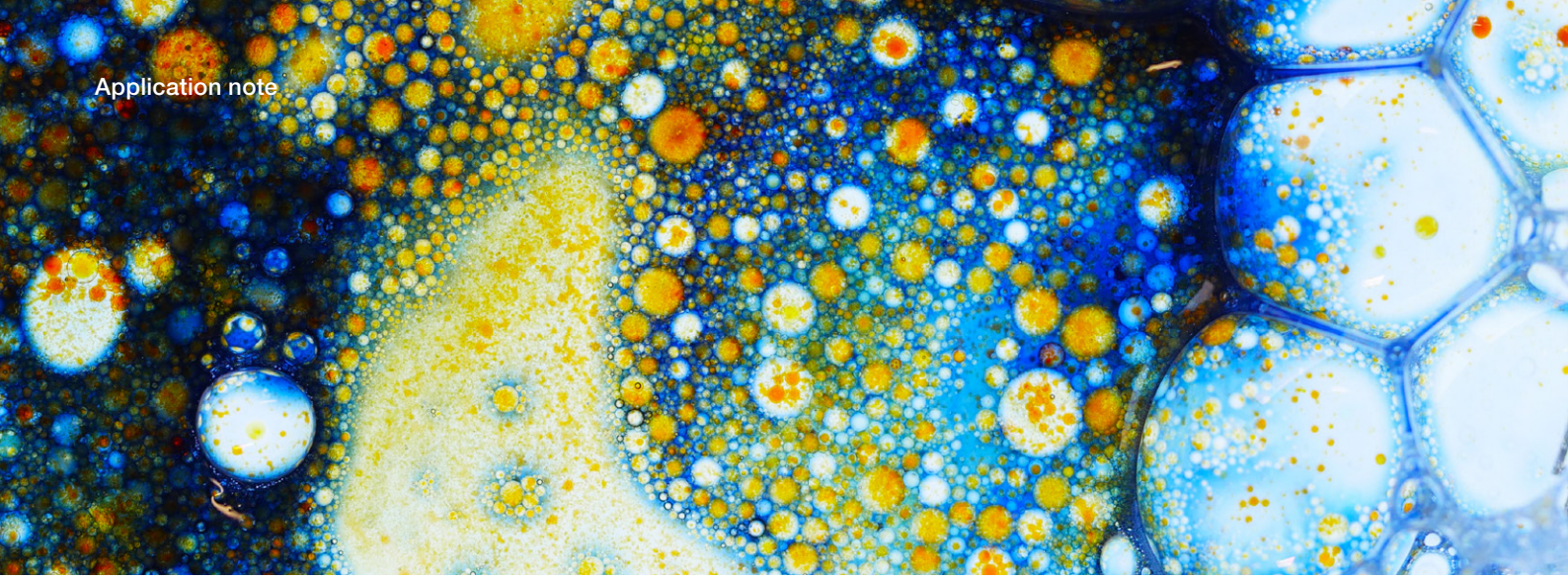
With the inclusion of the Acclaro Technology in the NanoDrop Eight Spectrophotometer's software, the corrected nucleic acid concentration was calculated after correcting for the contaminant absorbance contribution. This feature allows for simultaneous nucleic acid purity and quantity assessments. All the Acclaro Technology's software-corrected concentrations fall within ±20% of the theoretical concentration, with most samples within ±10%.

Conclusion

Contaminating nucleic acids in dsDNA or RNA preparations can cause costly delays in applications such as qPCR, where exact quantitation is crucial for a successful experiment. Since RNA and dsDNA both absorb at 260 nm, the true nucleic acid concentration will be overestimated with a copurified contaminant present. This overestimation can lead to experimental failures and require extensive troubleshooting. The ease with which the Acclaro Technology corrects for contaminating nucleic acids will save time, effort, and associated costs by improving sample purity and quantity assessments.

The function of the Acclaro Technology makes the nucleic acid purity assessment clear and simple. With each measurement of a nucleic acid sample, the NanoDrop Eight Spectrophotometer takes quality assessment a step further by outlining the contaminant identification, absorbance contribution, and the corrected sample concentration. The results from the experiments above indicate the NanoDrop Eight Spectrophotometer, which includes the Acclaro Technology in its software, can be implemented into many molecular biology workflows to obtain an accurate and advanced nucleic acid evaluation for downstream success.

Learn more at **thermofisher.com/nanodropeight**



Application note

Enabling real-time release of final products in manufacturing of biologics

Authors

Shaileshkumar Karavadra,
David James, and Arnaud Di Bitetto,
Hemel Hempstead, UK

Keywords

DXR3 SmartRaman, spectrometer, biopharmaceutical, GMP, real-time release testing, QbD, RTRT, manufacturing, multi-attribute testing

Introduction

Biopharmaceuticals (or biologics) are manufactured using biological-expression systems (such as mammalian, bacterial, and insect cells) and have spawned a large and growing biopharmaceutical industry (BioPharmaceuticals). The structural and chemical complexity of biologics, combined with the intricacy of cell-based manufacturing, imposes a huge analytical burden to correctly characterize and quantify both processes (upstream) and products (downstream). In small-molecule manufacturing, advances in analytical and computational methods have been extensively exploited to generate process analytical technologies (PAT) that are now used for routine process control, leading to more efficient processes and safer medicines.

Raman spectroscopy is a vibrational spectroscopy technique with several useful properties (non-destructive, non-contact, high molecular-specificity, and robustness) that make it particularly suited for PAT applications in which molecular information (composition and variance) is required.

Typical good manufacturing practice (GMP) operations involve performing an extensive set of tests according to approved specifications before the material is released to the market or for further processing. Recent ICH guidelines (ICH Q8, Q9, Q10, and Q11), however, suggest an alternative real-time release strategy to provide assurance of product quality prior to release. Real-time release testing uses the principles of the pharmaceutical Quality by Design (QbD) to optimize release and stability testing. A combination of manufacturing process understanding, process control, and product knowledge can be used to demonstrate that the material was made according to GMP.

The exact approach to real-time release testing (RTRT) will vary depending on the process requirements. The RTRT strategy may be based on control of process parameters, monitoring of product attributes, or on a combination of both at appropriate steps throughout the process. Critically, the RTRT strategy should be based on a firm understanding of the process and the relationship between process parameters, in-process material attributes, and product attributes.

Quality, cost, and speed are the major drivers for implementing in-line monitoring, at-line monitoring, and real-time release.

Here, we review some of the most important applications of Raman spectroscopy to the manufacturing and analysis of biopharmaceuticals. This article covers two aspects of the biopharmaceutical-manufacturing process: identity/variance testing of raw materials and cell culture media; and multi-attribute product testing of a biologic drug product or final product testing of a biologic drug product.

Raw material characterization

Acceptance of raw materials today is often predicated on small-scale functional testing and/or limited analytical methods, which may not be representative of at-scale performance. This leads, in some cases, to fluctuating process outputs and, in extreme cases, not meeting predefined release criteria. Furthermore, many clinical products are developed using a small number of batches resulting in a narrow range of raw material variation and thus a limited process understanding. Especially in upstream cell culture, the unforeseen variability of various components of the cell culture media can impact a product's micro-heterogeneity and its critical quality attributes (CQA).

Multi-attribute tests for high-risk raw materials may include identity test, quantitative test for the concentration of key ingredients in a raw material, batch-to-batch variability test, and degradation tests.

One high-risk raw material encountered in biologics manufacturing is cell culture media. Identification of cell culture media samples by traditional liquid chromatography (LC) methods, such as amino acid or vitamin analysis, has high costs and requires significant analytical expertise and laboratory space. Raman spectroscopy offers many potential benefits, such as low cost, portability, and potentially limited skill required to operate the instruments.

Buffers are another set of critical raw materials used in downstream manufacturing. Osmolality is a measure of concentration and is considered a critical quality attribute and critical process parameter in bioprocessing. The yield and quality of a biologic are highly dependent on the optimization of the downstream process. Identity testing along with osmolality of buffers can be carried out using a multi-attribute method based on principal component analysis and partial list squares. Rapid testing of buffers through single-use flexi bags can be carried out using the fiber optics probe of the Thermo Scientific™ DXR3 SmartRaman Spectrometer at the point of use with no need for sample preparation.

Final product identity testing

Final product identification of biologics pre- and post-shipment is another regulatory requirement. Product testing for identity through different kinds of primary packaging (glass vials, syringes, glass bottles) poses a significant analytical challenge in the manufacturing of biologics. Fill finish sites may not have the necessary analytical expertise to carry out the tests and may have to send the samples to the parent site or external lab for testing, incurring time and money.

Moreover, biologics or small molecule drug products would also have to undergo retesting upon importation either from a third country in the EU member state or the USA when drug products have been sent to the USA from other countries. A full list of tests is typically carried out, including final product identity testing. For biopharma manufacturers, this involves either sending the samples back to the parent site for analysis or employing third-party labs in the country of import. This increases significant costs and delays in the delivery of highly needed drug products.

End product identity testing/final product identity testing of biologics after fill-finish or pre-shipping to the fill-finish line is carried out by a variety of analytical techniques depending on the molecule/registration dossier.

For example, the verification test for biologic proteins is peptide mapping—a long-established workflow for protein identification using LC/mass spectrography (MS). This complex separation technique requires protein extraction and clean-up, enzyme digestion, one or more stages of liquid chromatography, and two phases of mass spectrometry before the final spectrum is matched against protein databases. Although it is a standard methodology, peptide mapping necessitates an analytical lab with qualified technical resources, entails extensive time for preparation, and introduces significant costs in solvents, columns, and analytical equipment.

The DXR3 SmartRaman Spectrometer, with its high sensitivity and resolution, allows characterization of the drug product by evaluating the fingerprint region of the molecule. Therefore, the DXR3 SmartRaman Spectrometer's unique capability with sampling flexibility ensures repeatable measurements, and subsequent analysis allows rapid method development and deployment.

We ran a feasibility study for multinational drug manufacture whereby the primary goal was to set up a rapid multi-attribute end product test to differentiate 15 different types of drug products and determine the concentration of the two preservatives in the drug products.

For this feasibility test we were given 15 different types of biologic drug products that varied in concentration from 0.5 mg/mL to 6 mg/mL. Concentration of two preservatives A and B ranged from 0.85 mg/mL to 5.0 mg/mL and 0.42 mg/mL to 3.91 mg/mL respectively.

These commercial drug products were supplied in their native glass vials varying in size and volume. A picture of such glass vials is shown below (Figure 1).

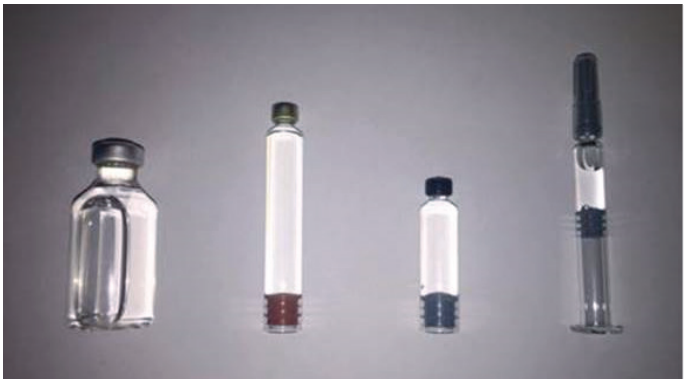


Figure 1. Typical native glass vials.

Reversed-phase high-performance liquid chromatography (HPLC) is currently used for the final product identity test and quantitative measurement of two preservatives in the final drug product.

DXR3 SmartRaman Spectrometer with universal sampling plate and 180-degree sampling module was used to acquire spectra of 15 drug products. To acquire each spectrum, a 532 nm laser with 40 mW power and 1 minute of scanning time was used. Ten spectra were acquired for each sample to accommodate the variability of glass vials and scattering effects.

DXR3 SmartRaman Spectrotometer offers excellent selectivity, repeatability, and full wavelength range to characterize biologics based on the characteristic band assignment (Table 1 and Figure 2).

Band frequency (cm^{-1})	Region	Vibrational mode	Protein structure assignments
870–1,150	Backbone, skeletal stretch	C_α-C, C_α-C_β, C_α-N	Secondary structure elements: α-helix, β-sheets, less-ordered structure
1,200–1,340	Amide III	N-H in-plane, Cα-N stretch	Hydrogen bonding, secondary structure
1,400–1,480	Side chain deformations	CH_2 and CH_3 deformations	Local environments, intermolecular interactions of side chains
1,510–1,580	Amide II	N-H deformations and C-N stretch (observed in UVRR and not conventional Raman spectra)	Local environments, intermolecular interactions of side chains
1,630–1,700	Amide I	C=0 stretch N-H in-plane bending	Secondary structure elements: α-helix, β-sheet, less-ordered structure

Table 1. Characteristic Raman band assignment.

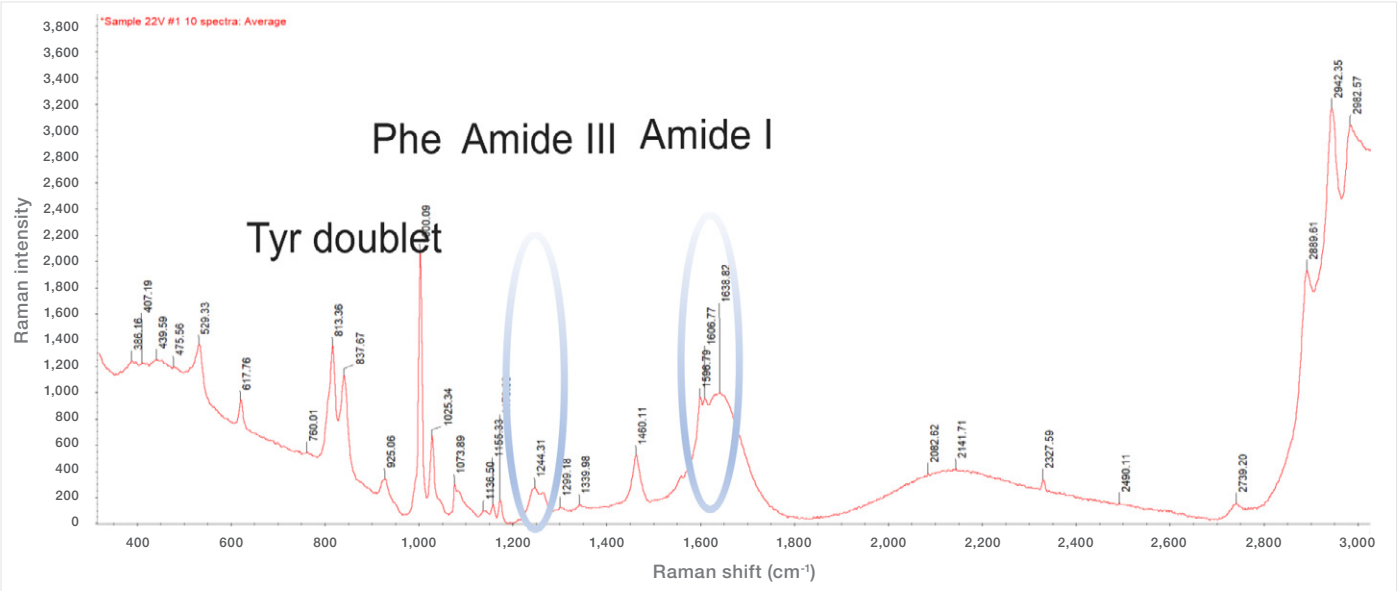


Figure 2. DXR3 SmartRaman spectrum showing characteristic bands of a biologic drug product.

Figure 3 shows the spectra of a sample containing a drug product against its placebo. It is imperative to establish that technique chosen for a feasibility study. In this case, Raman spectroscopy is sensitive enough to detect the differences between the drug product and its placebo. DXR3 SmartRaman Spectrometer offers high sensitivity to determine the significant differences between placebo and actual drug products.

Figure 4 is showing spectra of different classes of drug products. These spectra were utilized to build the discriminant analysis method on the Thermo Scientific™ TQ Analyst™ Software. TQ Analyst Software is a validated qualitative and quantitative method building software offering full compliance for pharmaceutical applications.

The discriminant analysis classification technique can be used to determine the class or classes of known materials that are most similar to an unknown material by computing the unknown's distance from each class center in Mahalanobis distance units. The discriminant analysis technique is typically used to screen incoming materials or final products to determine if they are compound/molecule a, b, or c.

Discriminant analysis methods typically specify at least two classes of known materials, but the method also works with only one class. Multiple standards may be used to describe each class (at least one class must contain two or more standards). Multiple regions of the spectrum may be used for the analysis.

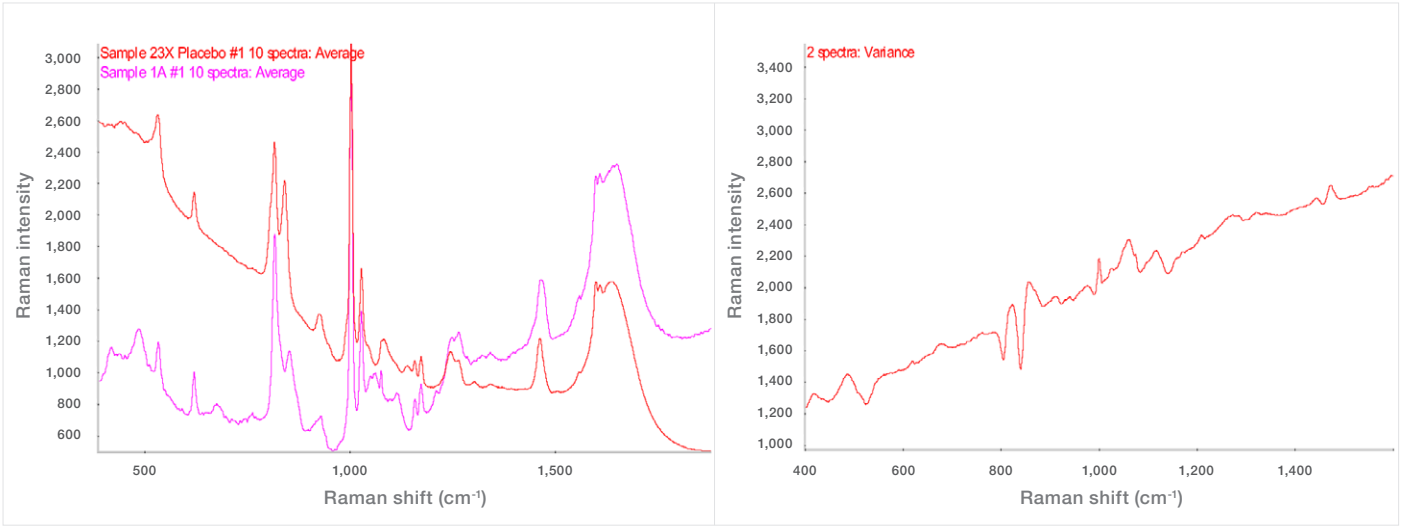


Figure 3. Raman spectra of drug product and its placebo and variance spectrum.

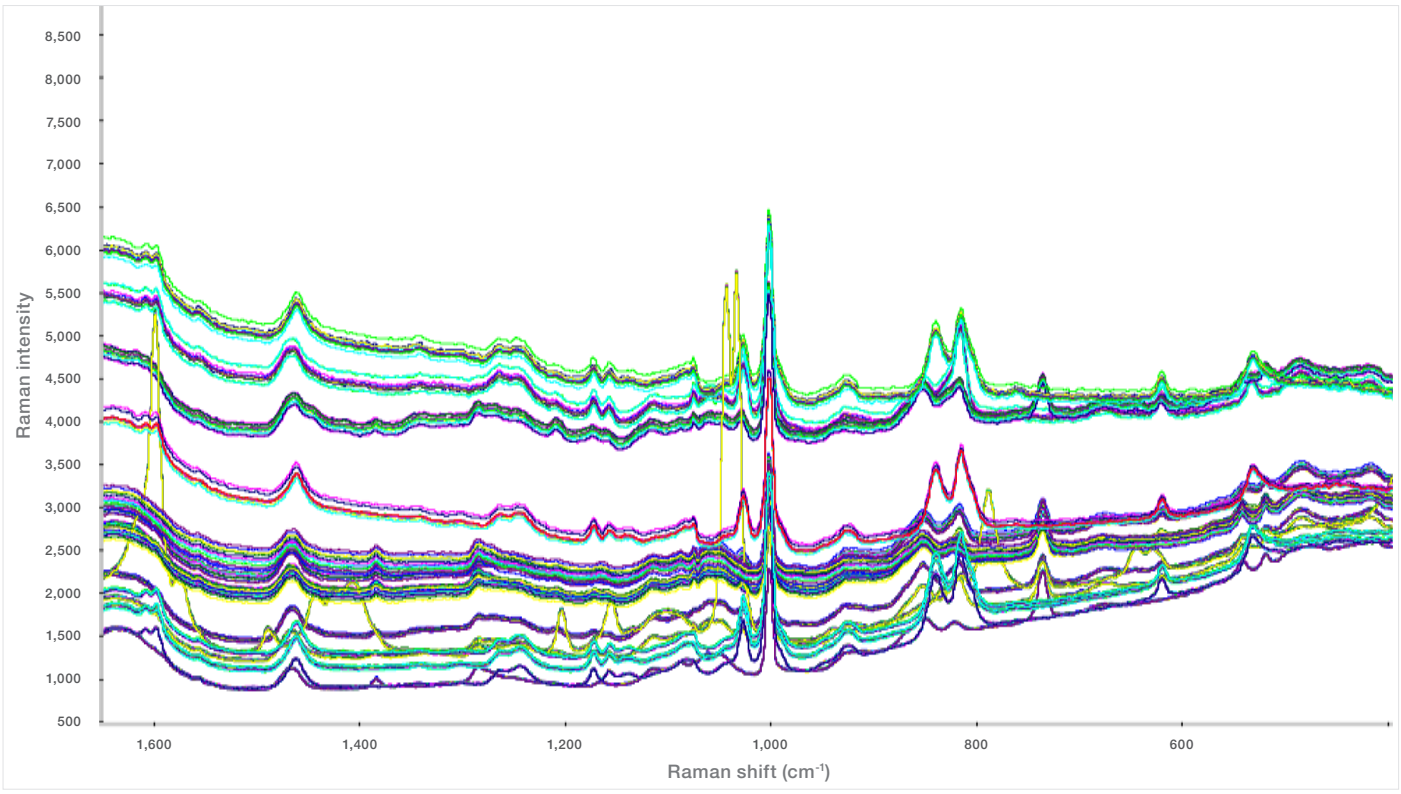


Figure 4. Raman spectra of different classes of drug products.

What does discriminant analysis do?

A discriminant analysis method applies the spectral information in the specified region or regions of an unknown sample spectrum to a stored calibration model to determine which class of standards is most similar to the unknown.

When the method is used to analyze an unknown sample or a class, the software performs a principal component analysis on the spectra of the standards and uses those results to determine score values for the unknown sample spectrum. The score plots are used to produce Mahalanobis distance values, which in turn are used to rank the classes.

The result of a discriminant analysis is the name of the class or classes that are most similar to the spectrum of the unknown sample. The Mahalanobis distance between the unknown sample and each reported class can also be reported. The closer each distance value is to zero, the better is the match.

After cross-validation, principal component scores plot revealed the class differentiation and the report indicated that all the classes of the different products were correctly identified with no mismatches to indicate false positives.

Quantitative analysis of biologics for preservative A and preservative B

As part of this feasibility study, our client also wanted to determine if the DXR3 SmartRaman Spectrometer test could be utilized to replace the HPLC test for measuring the concentration of two preservatives in their drug products. The level of preservative A was 0.85 mg/mL to 3.07 mg/mL and that of preservative B was 0.32 mg/mL to 2.57 mg/mL.

Pure samples of preservatives A and B were acquired as references, and to ascertain their presence in the final drug formulation.

Actual class	Mismatch	Calculated class	Calculated distance	Next class	Next distance
Product D		Product D	0.5809	C	4.5556
Product A		Product A	1.9869	I	12.9617
Product B		Product B	1.3796	E	25.1324
Product C		Product C	0.5417	D	3.8568
Product D		Product D	0.8466	M	9.0495
Product I		Product I	1.7709	A	13.9064
Product M		Product M	0.5284	S	3.3881
Product O		Product O	0.2244	X	17.3044
Product R		Product R	0.5419	C	4.4691
Product T		Product T	0.5944	X	2.3213
Product X		Product X	0.79	T	3.1646
Product S		Product S	1.1837	M	3.0829
Product N		Product N	1.0954	U	15.1798
Product U		Product U	0.1603	T	9.1738
Product S		Product S	1.8544	N	22.1624

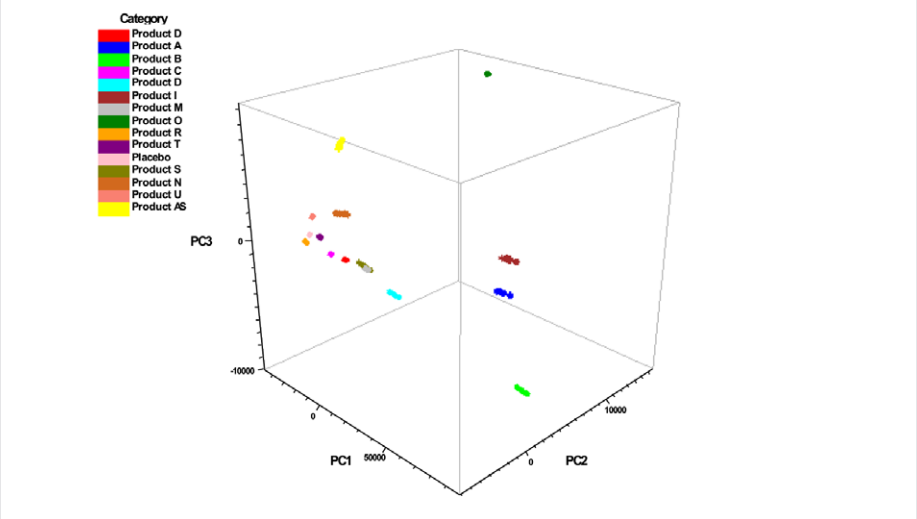


Figure 5. Analysis of preservative A and preservative B.

Samples of varying concentrations as per table 1 were acquired using the same parameters as of spectra acquired for identity test through 3 mL vial. Figure 6 is showing the spectra of the drug product with the two preservatives.

Four standards with the reference values were supplied in 3 mL and 10 mL vials and a validation sample to test the model for 3 mL and 10 mL vials.

Four spectra per standard were acquired and used to build the chemometric method. The final drug product samples were scanned with a DXR3 SmartRaman Spectrometer to acquire spectra in the range of 3500 to 50 cm-1 and captured with a single exposure of the CCD, avoiding stitching artifacts. The sample time took approximately 1 minute. Three spectra were collected per sample. The sample spectra were loaded into TQ Analyst Software for chemometric analysis using a partial least squares (PLS) method.

	Preservative A (mg/mL)	Preservative B (mg/mL)
Standard 1 3 mL and 10 mL	0.85	0.42
Standard 2 3 mL and 10 mL	1.27	1.12
Standard 3 3 mL and 10 mL	1.57	1.75
Standard 4 3 mL and 10 mL	3.07	2.57
Validation – 3 mL	1.57	1.75

Table 2. Calibration and validation sample.

	PLS results for 3 mL Cartridge	
	Preservative A (mg/mL)	Preservative B (mg/mL)
Validation sample: 3 mL	1.58 actual 1.57	1.71 actual 1.75
Real Sample in solution: 3 mL	1.56 actual 1.55	1.69 actual 1.77
Real sample in suspension: 3 mL	0.72 actual 0.69	1.23 actual 1.58

Table 3. Validation result for 3 mL sample.

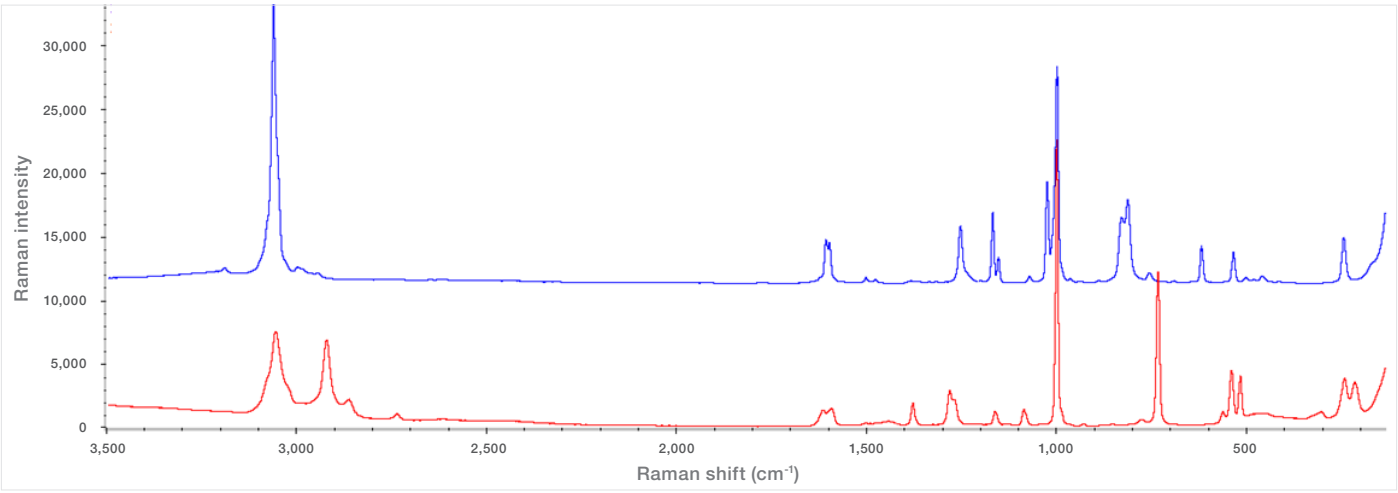


Figure 6. Spectrum in blue is from pure preservative A and spectrum in red is from pure preservative B.

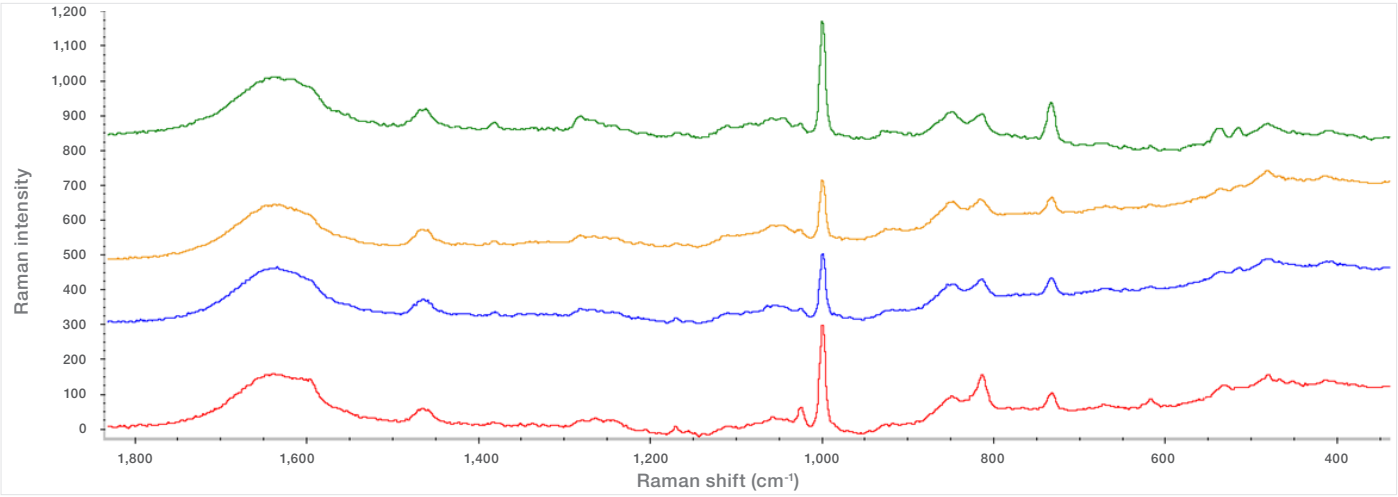


Figure 7. Spectra showing varying concentration of preservatives in final drug product.

Results

PLS analysis of the final drug product samples revealed excellent predictive capabilities within the range of materials tested. The spectra used to develop the PLS method for 3 mL cartridge are shown on calibration plots (Figure 8 and Figure 9) that compare the calculated preservative concentrations versus the actual concentrations. The calibration plot can be used to determine how well the method predicts the actual preservative concentrations in the samples. The plot developed by the chemometric method resulted in a correlation coefficient of 0.998 for preservative A. Root mean square error of calibration (RMSEC) was 0.0425 mg/mL, and the Root mean square error of prediction (RMSEP) calculated was 0.0372 for preservative A. The additional method for preservative B resulted in in a correlation coefficient of 0.999. The RMSEC was 0.0316 mg/mL, and the calculated RMSEP was 0.0496. The method was able to accurately predict the 3 mL validation sample and a real sample in solution (Table 3). The prediction can be improved when suspensions are allowed to settle and liquid phase is analyzed.

When 10 mL vial calibration samples were added to the above PLS method, method performance remained the same and was able to accurately predict the validation samples (Table 4).

Conclusions

A multi-attribute test to establish Final product identification and predicting concentrations of preservatives was done with the DXR3 SmartRaman Spectrometer by developing a discriminant analysis method and partial least square method. The final drug product identification test is part of release testing and current methods used are time-consuming and laborious. This Raman technique successfully demonstrates the ability to measure and monitor preservative concentrations either in the lab environment or at the line. The method developed shows excellent correlation with actual preservative concentrations with errors comparable to the reference analysis method. This application demonstrates the continued capability of the DXR3 Raman Spectrometer to be successfully used in bioprocess environments for implementing multi-attribute final product testing of biologics. Apart from the examples shown here, DXR3 SmartRaman Spectrometer can be used to implement at-line control strategies to monitor protein concentration, excipients concentration, and critical quality attributes like osmolality and pH. Many such examples are cited in the literature for Raman applications in biopharma manufacturing.

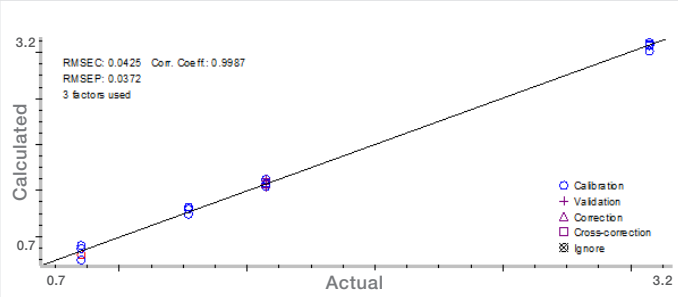


Figure 8. PLS model for preservative A —3 mL cartridge.

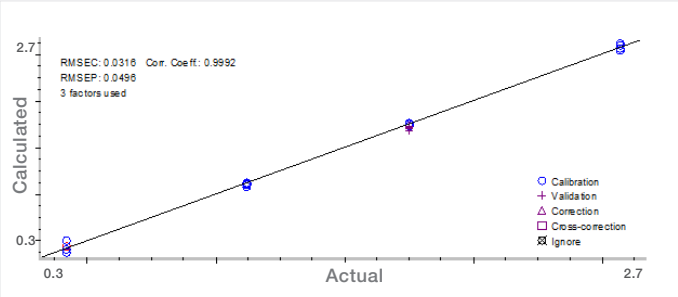


Figure 9. PLS model for preservative B —3 mL cartridge.

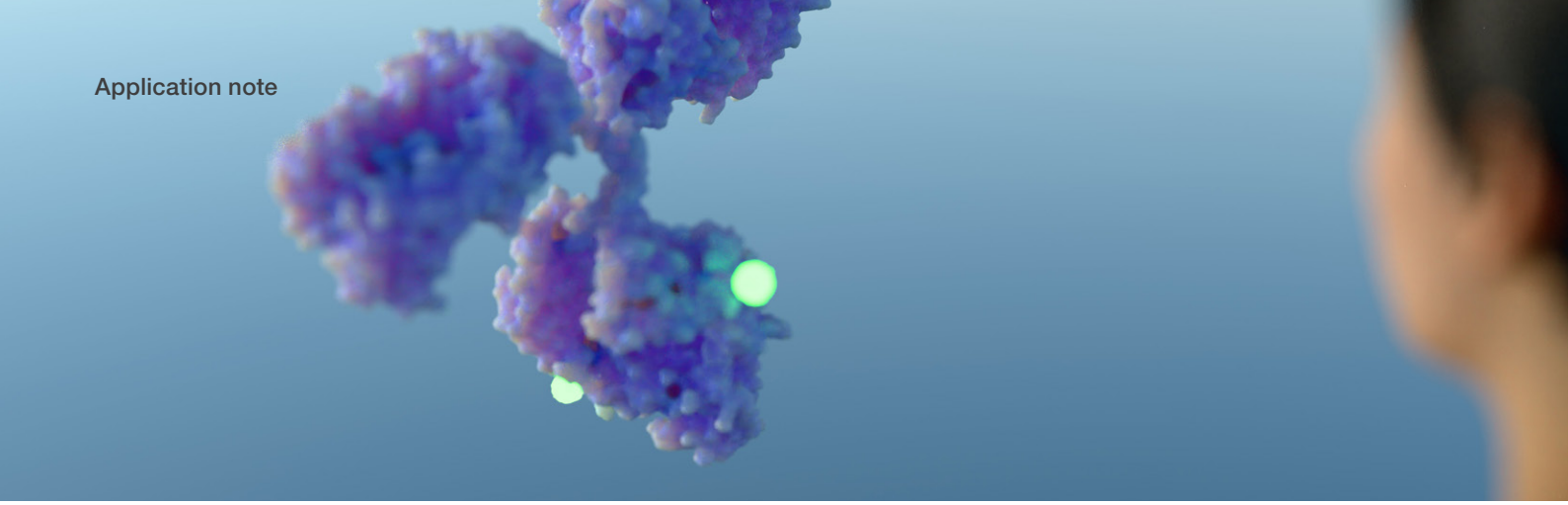
	PLS 3 mL cart and 10 mL vials	
	Preservative A (mg/mL)	Preservative B (mg/mL)
Validation sample: 3 mL	1.58 actual 1,57	1.71 actual 1,75
Real sample in solution: 3 mL	1.56 actual 1.55	1.65 actual 1.77
Real sample in suspension: 3 mL	0.80 actual 0.69	1.21 actual 1.58
Real sample in suspension: 10 mL	0.73 actual 0.68	1.32 actual 1.57

Table 4. Validation results for 3 mL 10 mL vials.

References:

1. European Medicines Agency. Guideline on Real Time Release Testing. https://www.ema.europa.eu/en/documents/scientific-guideline/guideline-real-time-release-testing-formerly-guideline-parametric-release-revision-1_en.pdf (accessed June 15, 2021).
2. Buckley K, Ryder AG. (2017). Applications of Raman spectroscopy in biopharmaceutical manufacturing: a short review. *Applied Spectroscopy*, 71(6): 1085-1116. https://aran.library.nuigalway.ie/handle/10379/7177 (accessed June 15, 2021).

Application note



Process Raman as a comprehensive solution for downstream buffer workflow

Authors

Michelle Nolasco[1]
Andrew Siemers[1]
Kristina Pleitt[1], Ph.D.
Nimesh Khadka[2], Ph.D.

[1]BioProduction, Thermo Fisher Scientific, St. Louis, Missouri USA

[2]Analytical Instrument, Thermo Fisher Scientific, Tewksbury, Massachusetts USA

Industry/Application:
Biopharma PAT / Downstream

Products used:
Thermo Scientific™ MarqMetrix™ All-In-One Process Raman Analyzer, Thermo Scientific™ MarqMetrix™ FlowCell Sampling Optic, Thermo Scientific™ MarqMetrix™ BallProbe™ Sampling Optic

Goals:
Enabling real-time excipient quantification and quality assessment using process Raman

Key analytes:
Arginine, Histidine, Sucrose

Key benefits:

- Enables real-time excipient quantification with cost and time benefits from eliminating the need for laboratory analytics
- Provides a platform to take actionable decision using real time data rather than theoretical estimation.
- Demonstrates potential of process Raman as a PAT tool for automation of Ultrafiltration/Diafiltration (UF/DF) and other downstream processes through simultaneous and real-time monitoring, quality assessment, and allowing multimodal feedback controls.

Introduction

Raman technology is rapidly gaining interest as a promising Process Analytical Technology (PAT) solution for real-time, non-invasive monitoring and control of downstream biopharma processes, especially for therapeutics like monoclonal antibodies (mAbs) and nucleic acids. Raman measurement, based on the vibration of molecular bonds, is highly specific for identification and quantification, even in complex or interfering matrices.

As an in-line PAT tool, Raman spectroscopy offers direct and rapid measurement in aqueous phases without sample preparation. These features make it ideal for monitoring and controlling dynamic processes such as downstream processing.

This study demonstrates a real-time methodology for accurately quantifying formulation excipients in the dynamic ultrafiltration/diafiltration (UF/DF) process using the Thermo Scientific™ MarqMetrix™ All-In-One Process Raman Analyzer (Figure 1). In addition, this study also illustrates a case where process Raman was able to provide real-time information on buffer quality.

Experimental details

Excipient quantification models

Calibration samples with defined concentrations of L-histidine, L-arginine, and sucrose were prepared using a design of experiments (DoE) approach called Uniform Design (UD) derived from number theory.[1] UD significantly reduces the total number of experiments while optimally spans the whole process space for model building and validation. These excipients were chosen due to their relevance in high-concentration monoclonal antibody (mAb) formulations. The analyte concentrations in the mixtures were designed with UD, with ranges of L-histidine (0-15 mg/mL), L-arginine (0-40 mg/mL), and sucrose (0-200 mg/mL) to develop calibration models. Each sample was passed in randomized order through a Thermo Scientific™ MarqMetrix™ FlowCell™ Probe integrated with the MarqMetrix All-In-One Process Raman Analyzer at a flow rate of 100 mL/min. The acquisition parameters were set to a laser power of 450 mW, an integration time of 3000 ms, and an average of 3 spectra, resulting in an 18 second total collection time per spectrum.

A Partial Least Squares (PLS) chemometric model was developed using the spectral range of 800 to 3235 cm^{-1} Raman shift. The spectra were normalized using infinity norm calculated in the spectral region of 2900 to 3230 cm^{-1} and preprocessed with a Savitzky-Golay (SavGol) filter (1st derivative, polynomial order = 2, window width = 13) and mean-centered. Overfitting was minimized by selecting appropriate latent variables using a leave-one-out cross-validation (LOOCV) strategy. To initially validate the model performance, seven different validation samples were collected in three different instruments using the same acquisition parameters as used in training data acquisition.

To allow the generalization of these excipient models to their use in ultrafiltration/diafiltration (UF/DF) process for IgG1 mAb, Raman spectra for IgG1 mAb (5 to 150 mg/mL in various matrices) were added to the training dataset, and new models were developed. The addition of these protein spectra allowed the resulting PLS model to better distinguish Raman signals among L-arginine, L-histidine, and protein.

Ultrafiltration/Diafiltration process

The Raman Process Analyzer with FlowCell Probe was integrated in-line to monitor an UF/DF process (Figure 2). A PES membrane was equilibrated with tris buffer pH 7.0 prior to feeding a purified IgG1 mAb at 10 g/L to a target loading of 500 g/m^2. In the first ultrafiltration (UF) step, the mAb was concentrated at feed rate of 300 $L/m^2/hr$, and TMP was maintained between 10-15 psi via manual flow restrictor. The mAb was then buffer exchanged into final formulation matrix containing L-histidine, L-arginine, and sucrose by manually feeding in the diafiltration (DF) buffer to the recirculation tank to maintain constant volume. After buffer exchange, the mAb was further concentrated to the desired final concentration in the second UF step.

Figure 1. Thermo Scientific MarqMetrix All-In-One Process Raman Analyzer, Thermo Scientific MarqMetrix FlowCell Sampling Optic.

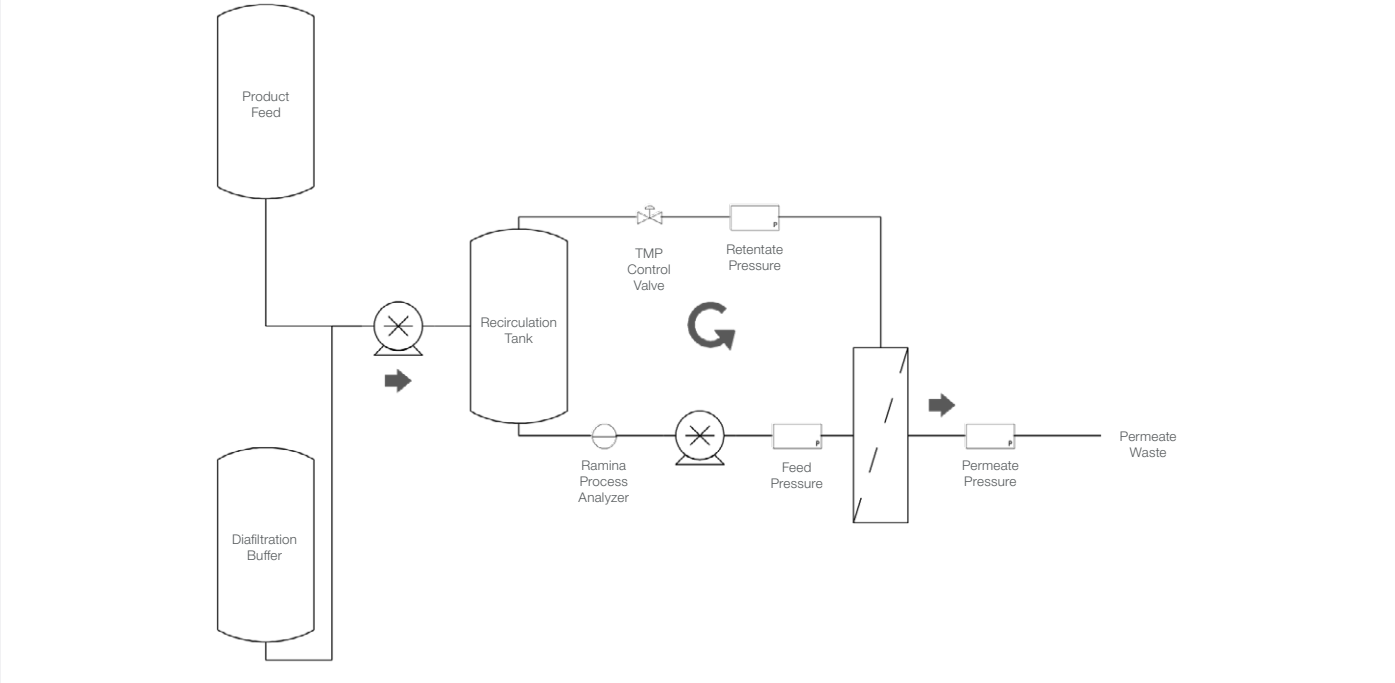


Figure 2. Ultrafiltration/Diafiltration (UF/DF) Process Diagram.

Buffer stability

Buffers in downstream processes are often prepared in advance and typically used based on prior knowledge of their stability, rather than confirming stability through analytical analyses before each UF/DF run. In two of our downstream runs, we noticed that the process Raman spectroscopy predicted a lower-than-expected sucrose concentration in the excipient buffer. To mitigate potential risks, we discarded the previously prepared buffer and made a fresh batch. The predictions from process Raman analyzer on the new buffer were much closer to the reference values obtained from HPLC (high performance liquid chromatography). To further validate this capability in a controlled experiment, we measured the Raman spectra of the excipient buffer, which contains L-histidine, L-arginine, and sucrose, at room temperature for 15 days. The buffer was monitored through a Thermo Scientific MarqMetrix BallProbe Sampling Optic integrated with the MarqMetrix All-In-One Process Raman Analyzer. The acquisition parameters were set to a laser power of 450 mW, an integration time of 3000 ms, and an average of 3 spectra, resulting in an 18 second total collection time per spectrum. Additionally, we collected data on pH, osmolarity, and performed HPLC analysis at various time intervals.

Results

The Partial Least Squares (PLS) models for L-histidine, L-arginine, and sucrose were initially tested using seven independent samples collected on three different process Raman analyzers. Data were mathematically processed to standardize spectra across instruments before applying the models. All the spectra were interpolated to have a common x-axis by equally spacing the 2048 pixels across 60 to 3250 cm^{-1} Raman shift, followed by relative y-axis standardization using the SRM fluorescence data as described in the NIST standardization protocol.[2]

Figure 3 shows the correlation plot of predicted versus reference values for L-histidine, L-arginine, and sucrose for the validation samples. A correlation coefficient of over 95% and a root means square error (RMSE) of less than 5% of the reference value for calibration, cross-validation, and prediction across three instruments demonstrate the reliability of process Raman to accurately predict the concentrations of these excipients, as well as easy model transferability.

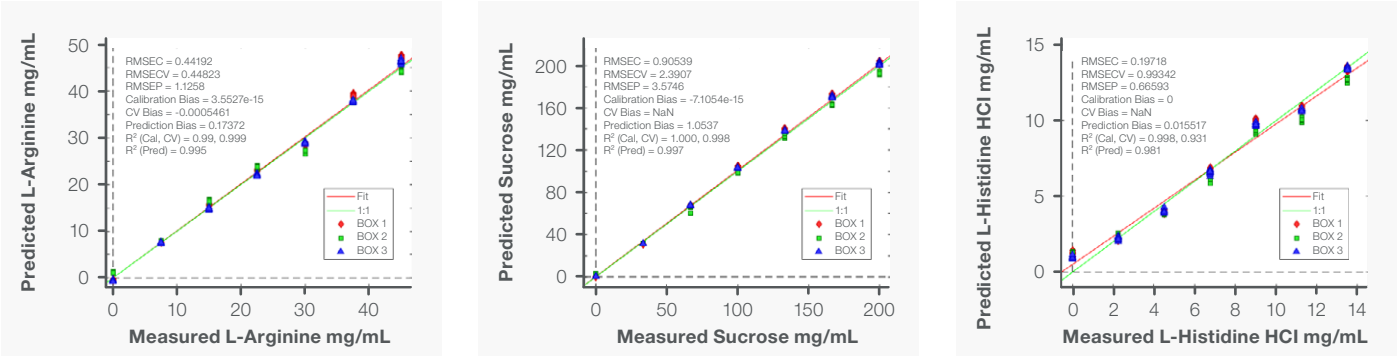


Figure 3. Correlation plot of predicted vs reference values for L-histidine, L-arginine, and sucrose across three different instruments.

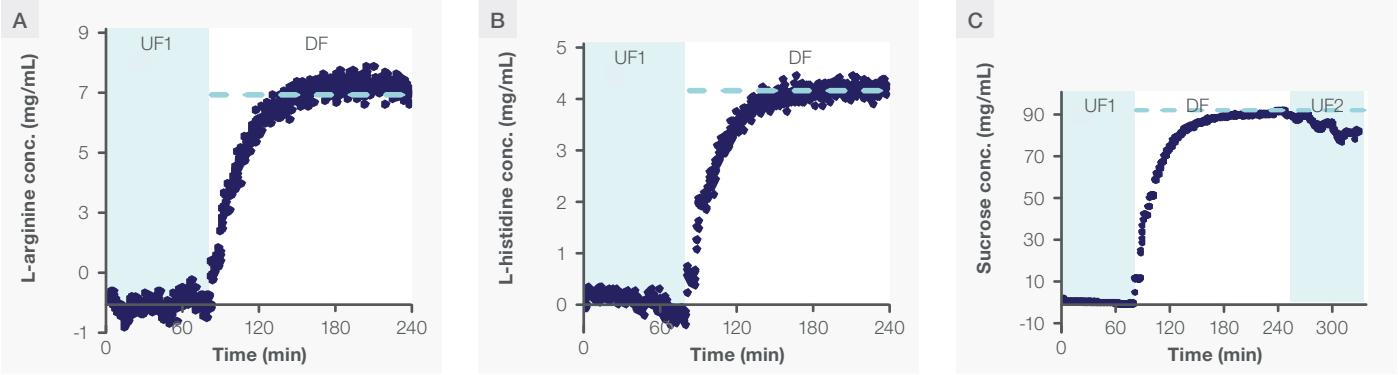


Figure 4. Raman concentration predictions during bench-scale UF/DF of IgG1 mAb for L-arginine (A), L-histidine (B), and sucrose (C); target end conc. shown with light blue dashed line.

The L-histidine, L-arginine, and sucrose models were then applied to the data acquired during the ultrafiltration/diafiltration (UF/DF) process. The predicted values for L-histidine, L-arginine, and sucrose are shown in Figures 4A, 4B, and 4C, respectively. The average predicted concentrations of L-histidine, L-arginine, and sucrose at the end of buffer exchange were compared to the reference values, resulting in a prediction error of less than 5% of the reference values (Table 3). This clearly illustrates the capability of process Raman analyzers to monitor and quantify excipients in real-time.

The real-time prediction of sucrose concentrations over 15 days at room temperature in a briefly air-exposed formulation buffer, containing L-histidine, L-arginine, and sucrose, is shown in Figure 5A. This scenario mimics what may occur during the storage of the formulation buffer. Initially, the predicted sucrose concentration was 86 mg/mL and remained stable for 5 days, but then steadily decreased to 57 mg/mL by day 15. Predictions from the arginine and histidine PLS models behaved similarly to those from the sucrose model, showing accurate and stable values up to day 5, but then steadily increasing until day 15 (data not shown). Raman spectral analysis revealed that the decrease in the sucrose peak was accompanied by the appearance of glucose and fructose Raman peaks (Figures 5B and 5C). HPLC analysis confirmed the intactness of arginine and histidine for all 15 days, and the hydrolysis of sucrose into glucose and fructose (Figures 5D and 5E).

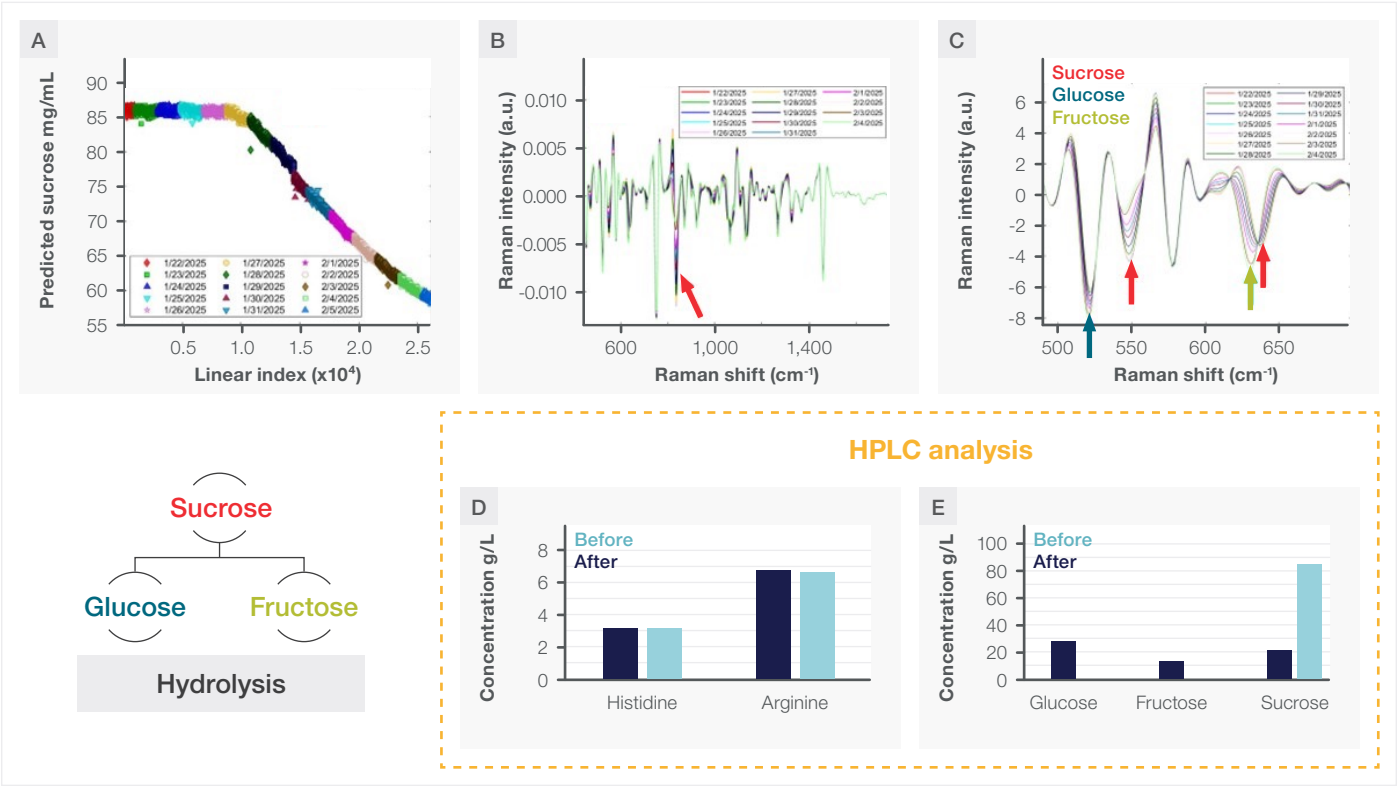


Figure 5. Showing the sucrose prediction over 14 days (A). Initially, the sucrose concentration was 86 mg/mL that lowered to about 57 mg/mL over the span of 14 days. The decrease in sucrose prediction were evident by decrease in sucrose specific intensity of 835 cm^{-1} that is assigned to the twisting (τ (CH_2)) with some contribution form symmetric stretching (ν(CC)) vibrational mode (B). Showing in figure C, the decrease in sucrose specific band (red arrow; ~550 cm^{-1} assigned as in-plane bending (β(OCO)) is followed by increase in glucose (blue; ~525 cm^{-1}) and fructose (green; ~ 640 cm^{-1}) specific Raman band that are assigned mainly to the deformation of CCC, CCO, and OCO bands. In figure D and E the result of day 1 (Before; blue) and day 15 (After; cyan) are compared where arginine and histidine remained unchanged while sucrose hydrolyzed to glucose and fructose.

Excipient	Reference concentration (g/L)	Predicted concentration (g/L)	% absolute error
L-histidine	4.2	4.1	1.0
L-arginine	7.0	7.1	1.4
Sucrose	92.4	95.6	3.4

Table 3. Prediction Error Calculation.

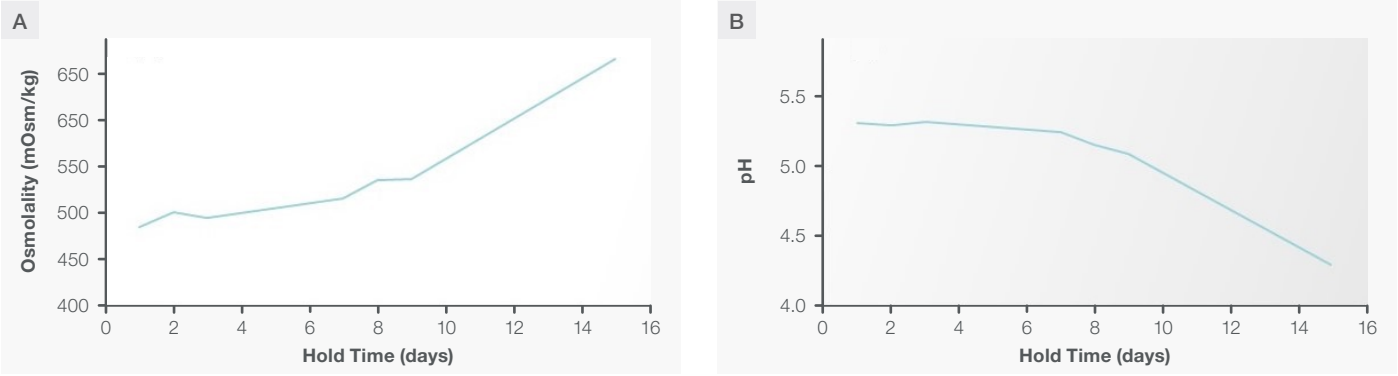


Figure 6. Showing increase in osmolarity by 40% (A) and decrease in pH by 1 unit during the hold period (B).

Acidic hydrolysis of sucrose is well documented in the literature.[3] To investigate if the root cause had pH-association, we examined the osmolarity and pH profiles during the experiment (Figures 6A and 6B). The onset of sucrose hydrolysis coincided with an increase in osmolarity and a decrease in pH, suggesting that the hydrolysis of sucrose into glucose and fructose is most likely driven by the lower pH. Since no acid was added to the system, the pH decrease was likely due to external factors. Although we did not identify the exact cause of the pH decrease, potential factors in practice could include bacterial growth, improper pH adjustment during buffer preparation, or dissolution of carbon dioxide or other acid-producing gases, among others.

Note that the sucrose PLS model was developed using a mixture of L-histidine, L-arginine, sucrose, and protein, and lacks any spectral information from glucose and fructose. Since glucose and fructose have significant overlaps in a wide spectral region, this explains why the Raman predictions were higher compared to the reference HPLC values.[4] The same reasoning applies to the discrepancies observed in Raman predictions for L-arginine and L-histidine (data not shown) when compared to the HPLC values (Figure 5D). In both cases, predictions can be improved by augmenting the model with additional training data that includes glucose and fructose spectral information. However, this was beyond the scope of the current work.

Not including glucose and fructose spectral information in the models is advantageous for monitoring buffer quality. As glucose and fructose are produced by sucrose hydrolysis, new spectral features appear that were not present in the training dataset. The Q residual is one of the model statistics calculated using the residual spectra remaining after projecting the original spectra into the model space.[5] As the spectral information of glucose and fructose increases with the progress of sucrose hydrolysis, the magnitude of the Q residual increases over time, as shown in Figure 7. Users can leverage this information to design quality control measures based on the reduced Q vs. T^2 plot to assess buffer quality.[5] For instance, in this study, a mean value of 0.15 for the reduced Hotelling T^2 and 0.5 for the reduced Q residual, with 95% confidence intervals for upper and lower limits (red dotted oval in Figure 7), can be used as quality control thresholds. Any buffer with reduced Q and T^2 values beyond these limits is deemed to fail quality control. All spectra after day 5 had low reduced T^2 and high reduced Q values, thus failing the quality control.

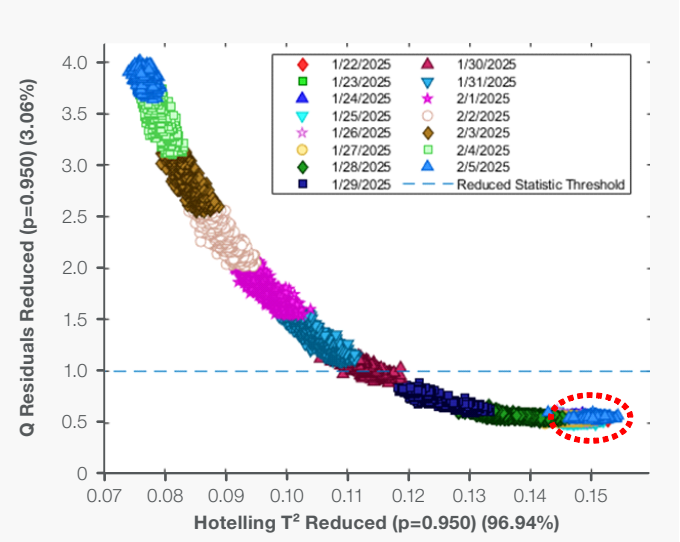


Figure 7. Showing increase in Q residual with sucrose hydrolysis, calculated by projecting the Raman data into the PLS sucrose model. The red dotted oval showing one of the possible control limits for quality assessment.

Conclusion

1. Real-time quantification of L-arginine, L-histidine, and sucrose were demonstrated with absolute error of < 5% in a UF/DF process using process Raman analyzer configured with FlowCell probe. In absence of in-line PAT tools to quantify these excipients, the volume needed for diafiltration (V_{df}) is carried out based on following mathematical expression:

$$V_{df} = V_0 * D = V_0 * ln(C_0 / C_f) / ln(1–R)$$

Equation 1.

where:

- *V_0 is the initial volume of the solution.*
- *D is the number of diavolumes required.*
- *R is retention factor*
- *C_0 is initial concentration of solute*
- *C_f is the final concentration of solute*

In practice, the retention factor for excipients is typically assumed to be 0, as the pore size of the diafiltration membrane is significantly larger than the hydrodynamic size of excipients. However, charge buildup across the membrane results in electrochemical potential which in turn prevents the free mobility of charged excipients, thereby increasing their retention factor above 0. This effect is known as the Gibbs-Donnan effect.[6] In such scenarios, the empirically calculated volume needed for diafiltration (V_{df}) can result in incomplete buffer exchange, which may affect the functionality and stability of monoclonal antibodies (mAbs). An in-line process Raman analyzer provides a reliable solution to this issue by offering real-time monitoring of excipient concentrations. This enables tighter process control and ensures product quality by allowing for immediate adjustments to the diafiltration process, thereby preventing incomplete buffer exchange and maintaining the stability and functionality of the therapeutic product.

Similarly, in Figure 4C, the sucrose concentration in the retentate decreased during UF2, as confirmed by offline HPLC analysis. Given the hydrodynamic size of sucrose relative to the pore size of the membrane, sucrose should theoretically exchange freely between the retentate and filtrate, resulting in equal concentrations in both. However, as the protein concentration increases, the osmotic pressure also rises, making the exclusion of water thermodynamically unfavorable. To balance the osmotic pressure difference, *sucrose is excluded along with water.*[6] This effect was accurately captured in the real-time predicted data from the process Raman analyzer. Thus, process Raman provides unique capabilities to ensure product quality by offering real-time data, rather than relying on empirical hypotheses.

2. This work, combined with our previous demonstrations of accurate in-line protein quantification during UF/DF processes,[7,8] clearly highlights the value of process Raman for downstream process monitoring. Raman spectroscopy allows for the simultaneous measurement of multiple critical process parameters (CPPs) with a single scan. These findings establish process Raman as a PAT tool with unparalleled benefits compared to other analytical methods.
3. Simultaneous measurement of protein and excipient concentrations not only allows tighter process control but also opens opportunities for automating downstream processing.
4. The ability of process Raman to provide real-time insights into buffer quality before its use in UF/DF runs offers substantial value by preventing batch failures. This capability enhances quality control, making Raman spectroscopy an essential tool for integration as an in-line sensor to improve downstream process monitoring, control, and automation.

References

1. Zhang, L.; Liang, Y.-Z.; Jiang, J.-H.; Yu, R.-Q.; Fang, K.-T. Uniform Design Applied to Nonlinear Multivariate Calibration by ANN. *Analytica Chimica Acta* **1998**, *370* (1), 65–77. https://doi.org/10.1016/S0003-2670(98)00256-6.
2. Choquette, S. J.; Etz, E. S.; Hurst, W. S.; Blackburn, D. H.; Leigh, S. D. Relative Intensity Correction of Raman Spectrometers: NIST SRMs 2241 through 2243 for 785 Nm, 532 Nm, and 488 Nm/514.5 Nm Excitation. *Appl Spectrosc* **2007**, *61* (2), 117–129. https://doi.org/10.1366/000370207779947585.
3. Torres, A. P.; Oliveira, F. a. r.; Silva, C. I. m.; Fortuna, S. p. THE INFLUENCE of pH ON the KINETICS of ACID HYDROLYSIS of SUCROSE. *Journal of Food Process Engineering* **1994**, *17* (2), 191–208. https://doi.org/10.1111/j.1745-4530.1994.tb00335.x.
4. Wiercigroch, E.; Szafraniec, E.; Czamara, K.; Pacia, M. Z.; Majzner, K.; Kochan, K.; Kaczor, A.; Baranska, M.; Malek, K. Raman and Infrared Spectroscopy of Carbohydrates: A Review. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2017**, *185*, 317–335. https://doi.org/10.1016/j.saa.2017.05.045.
5. Kumar, S.; Martin, E. B.; Morris, J. DETECTION OF PROCESS MODEL CHANGE IN PLS BASED PERFORMANCE MONITORING. *IFAC Proceedings Volumes* **2002**, *35* (1), 125–130. https://doi.org/10.3182/20020721-6-ES-1901.00752.
6. Agrawal, P.; Wilkstein, K.; Guinn, E.; Mason, M.; Serrano Martinez, C. I.; Saylae, J. A Review of Tangential Flow Filtration: Process Development and Applications in the Pharmaceutical Industry. *Org. Process Res. Dev.* **2023**, *27* (4), 571–591. https://doi.org/10.1021/acs.oprd.2c00291.
7. Nolasco, M.; Pleitt, K.; Khadka, N. Using a Process Raman Analyzer as an In-Line Tool for Accurate Protein Quantification in Downstream Processes.
8. Nolasco, M.; Pleitt, K.; Khadka, N. Raman-Based Accurate Protein Quantification in a Matrix That Interferes with UV-Vis Measurement.

ThermoFisher SCIENTIFIC

Notes

Learn more at **thermofisher.com**

thermo scientific