Application note

Using a Process Raman Analyzer as an In-line Tool for Accurate Protein Quantification in Downstream Processes

Authors

Michelle Nolasco[1], Kristina Pleitt[1], Nimesh Khadka[2]

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

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

Significance

In-line and real-time measurement of accurate protein (monoclonal antibody, or mAb) concentration has been demonstrated over a wide dynamic range (0-135 g/L) using the Raman process analyzer for the ultrafiltration/ diafiltration (UF/DF) as shown in Figure 1.

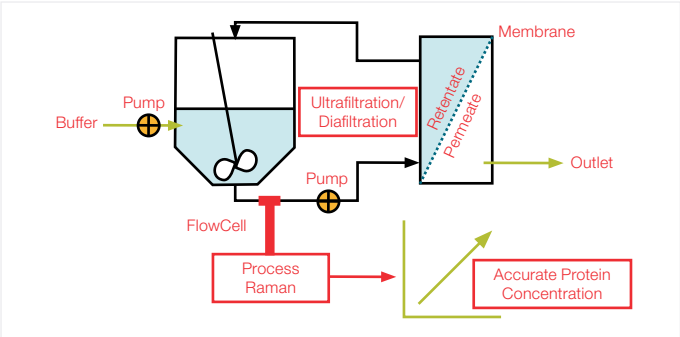

Figure 1. Diagram showing integration of process Raman in a UF/DF process using a flowcell probe for in-line estimation of accurate protein concentration.

The methodology described in the paper offers actionable results for downstream processing, providing valuable insights for monitoring and controlling protein concentration. Additionally, we have demonstrated that the Raman models developed for measuring mAb concentration are transferrable to a similar class of monoclonal antibodies.

Introduction

Therapeutic proteins (e.g., monoclonal antibodies, insulin, Fc-fusion protein, antibody-drugs conjugates, hormones) are used to treat a variety of conditions and diseases such as cancer, immunologic diseases, communicable diseases like COVID-19, and many others.[1] The efficacy and functionality of a therapeutic protein are dependent on dosage and its physical state and structure. Abrupt changes in the environment of the protein, such as shifts in pH, temperature, shear force, or chemical modification, can induce conformational changes within the protein. The structural changes may lead to protein denaturation, aggregation, and even degradation—all of which can lower the efficacy of the therapeutic protein or trigger negative health consequences for the patient. Thus, ensuring the quantity and quality of therapeutic proteins throughout the manufacturing process—from production, through the fill and finish stage, and during storage— is of paramount importance.[2]

Therapeutic proteins like monoclonal antibodies are produced in a bioreactor where cells are maintained at optimum conditions by controlling nutrition and physical factors like pH, temperature, and dissolved oxygen. The bioreactor's controlled environment leads to improved monoclonal antibody production. Following the upstream process, operations like centrifugation, filtration, chromatography, viral inactivation, concentration, buffer exchange, formulation, and fill and finish are performed to purify the monoclonal antibodies and formulate them into drug products. These operations are collectively categorized as downstream processes (Figure 2).

In this study, we placed a Thermo Scientific™ MarqMetrix™ All-In-One Process Raman Analyzer in a UF/DF run, as described in Figure 1, for use as an in-line process analytical technology (PAT) solution for measuring mAb concentration. In-line PAT enables real-time monitoring and control of critical process parameters that are key to reducing batch-to-batch process variability and ensuring uniformity of products. A well-controlled process also implies high efficiency, product quality, and minimized manufacturing costs.[3] Currently, ultraviolet- visible (UV-Vis) spectrometry is the standard technology for in-line quantification of protein concentration during downstream purification.[4] However, in recent years, process Raman has gained popularity as a complementary technology for in-line monitoring of protein concentration with its additional benefits of being able to measure excipient concentrations, buffer components, and critical quality attributes (CQAs) of products.[5] Here, we demonstrate the use of process Raman as a viable PAT solution for accurate and real-time quantification of mAb concentration during downstream processing. In addition, we also show the model's transferability to a different monoclonal antibody.

Experimental design and data analysis

Calibration model development

Partial least square (PLS) calibration models for mAb quantification were developed using the calibration samples of product A ranging in concentration from 0 to 135 mg/mL. The samples were passed through the FlowCell probe integrated with process Raman at a 100 mL/min flow rate. The Raman spectra were acquired using a 785 nm laser with the following acquisition parameters: laser power 450 mW; integration time 3000 ms; average 3 (i.e., a single spectrum per 18 s); and ten replicates per concentration.

PLS models were developed for mAb quantification using two spectral regions. The "Amide I PLS model" utilized the spectral region of approximately 1550 to 1850 cm^{-1}, while the "Extended Region PLS model" utilized a broader spectral region from approximately 850 to 1850 cm^{-1}. The terminal region of the spectra (~ 3098 to 3230 cm^{-1}) that corresponds to water band vibrations was also included in the normalization model. The models were built after following the data preprocessing steps: 1) normalization to the water band in every spectrum using infinity norm; 2) SavGol filter 1st derivative (smoothing window 13 and polynomial order 2); and 3) mean centering. Both models were internally validated using a leave-one-out cross-validation (LOOCV) strategy (contiguous block of 10). Initially, both models were built using the latent variables ranging from 1 through 20. The root mean square error of calibration (RMSEC), and the root mean square error of cross-validation (RMSECV) was calculated for each model. Finally, the optimum latent variable PLS model was selected using the following criteria: 1) adding more latent variables did not significantly improve the RMSECV; and 2) the values of RMSEC and RMSECV were similar. To evaluate model specificity, the variable importance in the projection (VIP) scores was calculated for the above two models. All chemometric works were performed using Solo 9.3 (2024) Eigenvector Research, Inc, Manson, WA USA 98831 software.

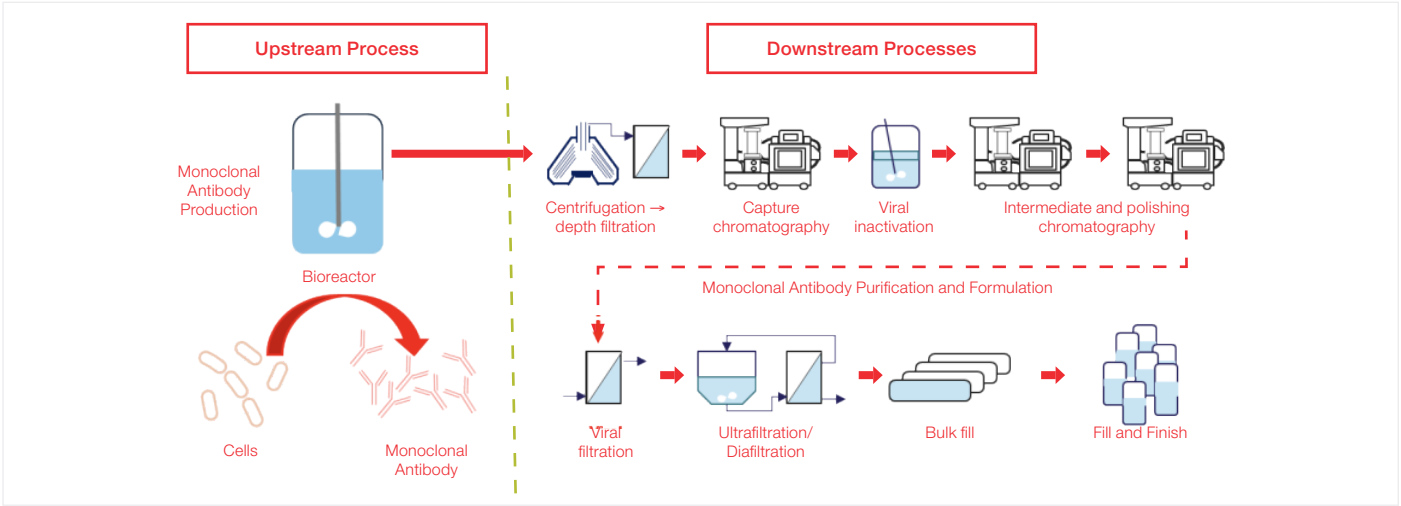

Figure 2. Upstream and downstream workflow for production of monoclonal antibodies.

Validation of model performance

The Raman spectra for the test samples were collected for two different monoclonal antibodies (products A and B) using the procedure described for the training data. The formulation buffers for products A and B were different during the diafiltration steps. These acquired spectra were fed into the models to get the real-time prediction of protein concentration. The predicted concentrations were then compared with the reference in-line and offline UV-Vis values to estimate the prediction error.

Results

The correlation plot between measured and predicted (during cross-validation) protein concentrations for the Amide I PLS models is shown in Figure 3A. The Amide I PLS model was developed using four latent variables as the RMSECV did not improve by adding more latent variables (*data not shown*). The RMSEC and RMSECV for the Amide I PLS model were 0.526 mg/mL and 0.607 mg/mL respectively. The ratio of RMSEC and RMSECV being close to 1 suggests that the model is not overly fitted. Similarly, the R^2 of CV is close to 1, and negligible CV bias indicates that the model fits well with the training data. The model statistics are summarized in Table 1.

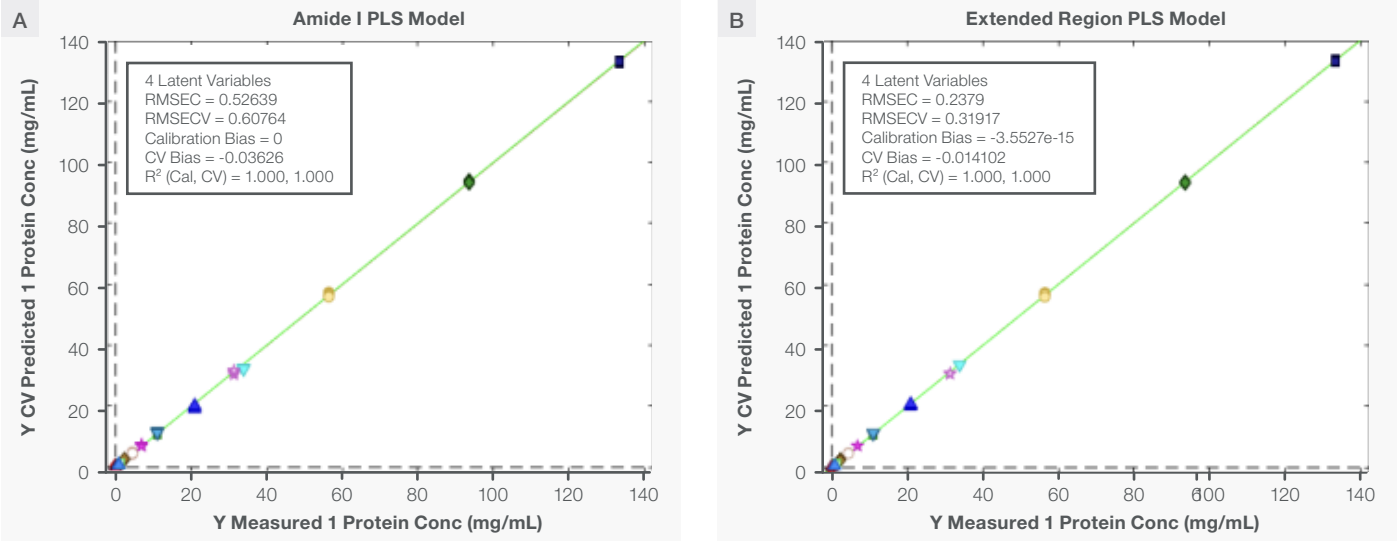

Figure 3. The correlation plot for measured and cross validation predicted protein concentration for mAb using Amide I and Extended Region PLS model.

PLS Model	Latent Variables	RMSEC (mg/mL)	RMSECV (mg/mL)	CV Bias	R^2 CV
Amide I PLS model	4	0.526	0.607	-0.036	~1
Extended Region PLS model	4	0.237	0.319	0.014	~1

Table 1.

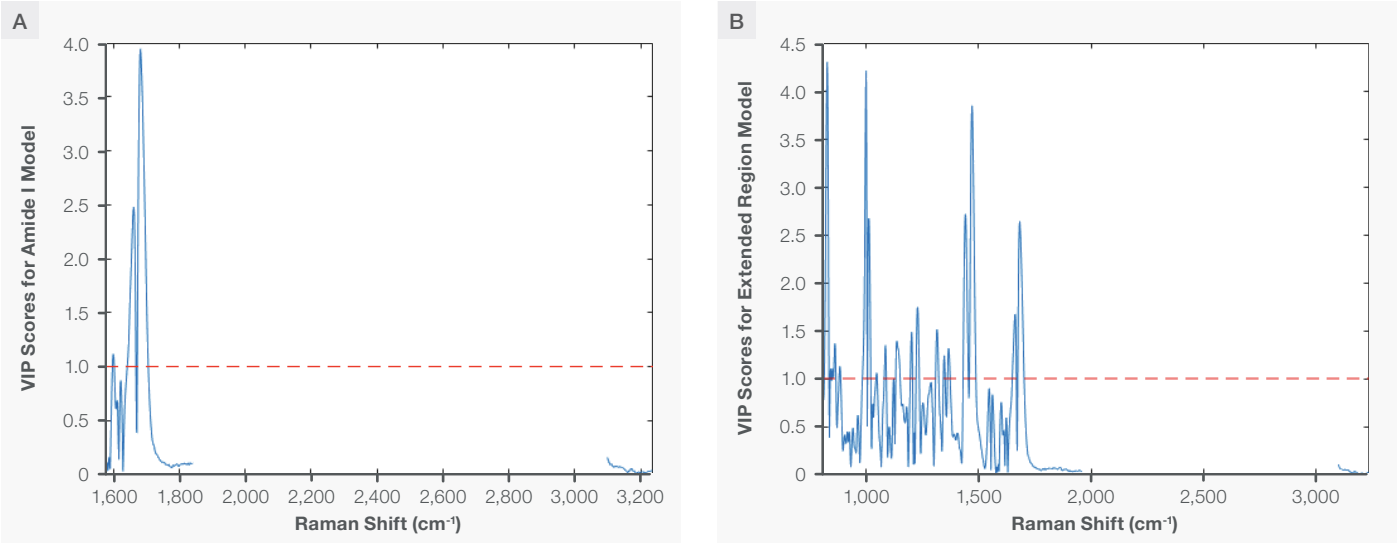

Figure 4. The VIP scores plot for Amide I PLS model (plot A) and Extended Region PLS model (plot B).

The Amide I model is primarily based on the Raman signature of the carbonyl group (–C=O) of the peptide bond (–CO–NH–) of mAb. The carbonyl group on different secondary structures has different Raman shifts ranging from 1600 to 1750 cm^{-1}. The mAb secondary structure is primarily a β-sheet structure. The carbonyl in the β-sheet secondary structure has a Raman peak at ~ 1670 cm^{-1}. Thus, it can be hypothesized that the Raman shift at ~1670 cm^{-1} should influence the Amide I PLS model. The VIP scores were calculated to assess the influence of each Raman shift on the model. The Raman shifts with VIP scores over one are considered significant for the model. Figure 4A shows the VIP scores plot for the Amide I PLS model. The red dotted line represents the threshold of a VIP score equal to 1. The region from 1640 to 1700 cm^{-1} have VIP scores of more than one and are influential to the model. The region around 1670 cm^{-1} has the highest VIP score in the Amide I PLS model. This indicates the carbonyl Raman signature in the β-sheet secondary structure highly dominates the Amide I PLS model. In other words, the VIP scores plot for the Amide I PLS model indicates that the model has high specificity for mAb.

Based on the published work, the Amide I region of mAb is unique. It has a distinct Raman signature compared to other molecules (excipients and buffers) commonly used in the downstream processes. No spectral interference in the Amide I region suggests that the Amide I PLS model is across different matrixes. In addition, most of the mAb have similar mass (a similar number of carbonyl residues) and secondary structure. Thus, the Amide I PLS model is transferable across multiple mAb within the same classes. This was the basis of our intent to develop the Amide I PLS model.

The correlation plot for the four latent variables of the Extended Region PLS model is shown in Figure 3, and the model statistics are summarized in Table 1. Like the Amide I PLS model, the Extended Region PLS model's excellent statistics suggest that it adequately captures the spectral information and correlates it with the measured concentration. The VIP scores plot for the Extended Region PLS model is shown in Figure 4B. Besides the Amide I region, the Raman shift that corresponds to the CH deformation (~1450 cm^{-1}), the breading mode of the phenylalanine ring (~1005 cm^{-1}), and the tyrosine vibration mode (~850 cm^{-1}) are influential in this model. Thus, the VIP scores plot for the Extended Region PLS model indicates the specificity of the model for mAb.

The RMSEC and RMSECV for the Amide I PLS model are higher than the Extended Region PLS model, as shown in Table 1, which suggests that the former model is less accurate than the latter. The Extended Region PLS model leverages additional variables (Raman shifts) in explaining the information of the training dataset. Although the Extended Region PLS model provides more accuracy, in some cases, the transferability of this model across different matrices may result in higher prediction error due to spectral overlap. Thus, both models have pros and cons, and depending on the need, one model might be a better choice.

Performance of models on the UF/DF run using product A (mAb)

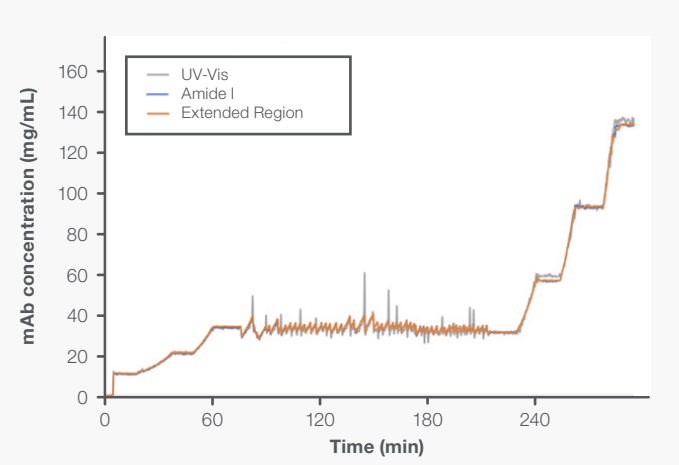

Figure 5. Showing excellent correlation of the real-time monitoring of UF/DF of product A.

The real-time prediction of protein concentration during the process of UF/DF using data from in-line UV-Vis (grey), Amide I PLS model (blue trace), and Extended Region PLS model (orange) is demonstrated in Figure 5. As shown in the figure, all three predictions overlayed each other, demonstrating excellent agreement. Minute data discrepancies between in-line UV-Vis and Raman were observed. The Raman data was acquired every 18 sec, while the in-line UV-Vis data was acquired every 12 sec. Since the UF/DF process is highly dynamic, the difference in the acquisition times for the in-line UV-Vis and Raman instruments explains the discrepancies. Such discrepancies can be easily overcome with proper control of the process dynamics or acquisition settings. Finally, when the Raman prediction from both models was compared with the offline UV-Vis reference values, the absolute prediction errors were below 5% throughout the process.

Transferability of models on product B (mAb) UF/DF run

The Amide I and Extended Region PLS models were developed using the Product A data set. A UF/DF run was carried out with different protein (product B) in a different formulation matrix to validate the model performance with other monoclonal antibodies. Figure 6A shows the real-time monitoring and prediction of Product B concentration during the UF/DF process using Amide I and Extended Region PLS models. The prediction from the Amide I model (blue trace) and Extended Region Model (orange trace) excellently overlay each other. To the right in Figure 6B, the absolute prediction error between the predicted (Raman) and reference (offline UV-Vis) is shown as a function of protein concentration. The blue and orange bars represent the prediction errors from the Amide I PLS and Extended Region PLS models, respectively. The overall prediction errors were below 3% across the concentration range tested.

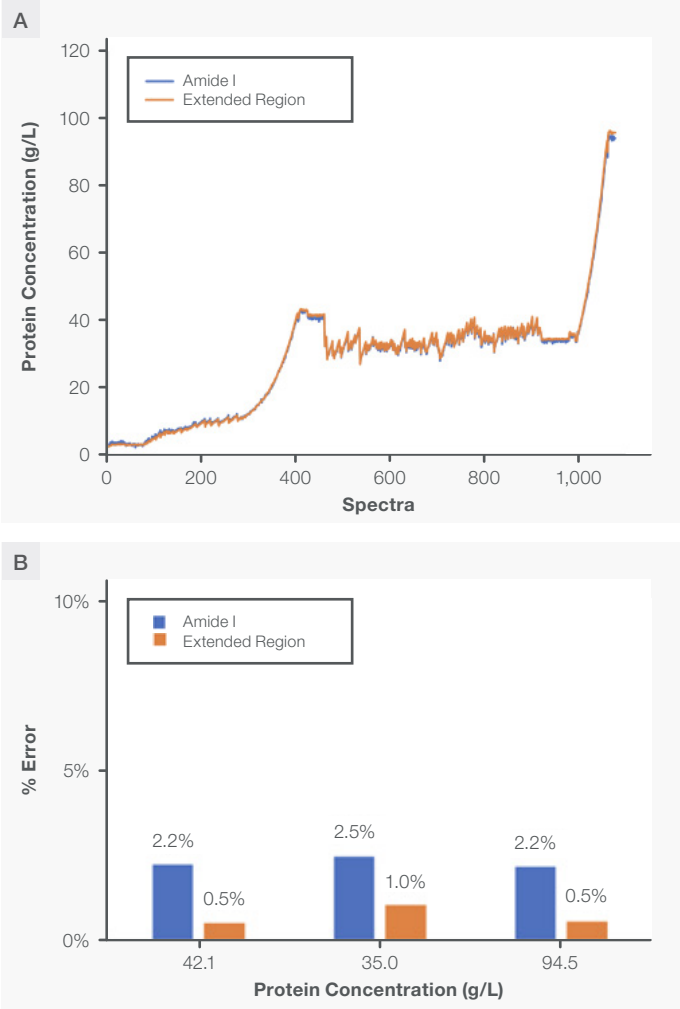

Figure 6. Data demonstrate excellent model transferability between proteins with different formulation matrices. The calibration models for predicting protein concentration were built with Product A and applied to Product B that has a different formulation buffer. The absolute prediction error was < 3 %.

Conclusions

- ✓ We demonstrated process Raman to be a rapid, reliable, and easy-to-use PAT solution for in-line monitoring of protein concentration during downstream processes, with an accuracy comparable to the reference values from in-line and offline UV-Vis instrumentation.
- ✓ In this study, we outlined two strategies for developing the chemometric models for the quantification of mAb-based on specific Raman signatures. Because of the high specificity of both models for mAb, we also demonstrated their excellent transferability to a different mAb in a UF/DF run, regardless of differences in buffer formulations.
- ✓ Based on works published in the literature, process Raman is a broader PAT solution for downstream applications because of the additional benefits it brings regarding measurement of buffer components, critical quality attributes of proteins (aggregation, secondary structure, disulfide region), protein modification (antibody-drugs conjugates), formulation components, and more.6–8

References

1. Leader, B.; Baca, Q. J.; Golan, D. E. Protein Therapeutics: A Summary and Pharmacological Classification. *Nat Rev Drug Discov* **2008**, 7 (1), 21–39. https://doi.org/10.1038/nrd2399.
2. Alt, N.; Zhang, T. Y.; Motchnik, P.; Taticek, R.; Quarmby, V.; Schlothauer, T.; Beck, H.; Emrich, T.; Harris, R. J. Determination of Critical Quality Attributes for Monoclonal Antibodies Using Quality by Design Principles. *Biologicals* **2016**, *44* (5), 291–305. https://doi.org/10.1016/j.biologicals.2016.06.005.
3. Research, C. for D. E. and. PAT — *A Framework for Innovative Pharmaceutical Development, Manufacturing, and Quality Assurance*. U.S. Food and Drug Administration. https://www.fda.gov/regulatory-information/search-fda-guidance-documents/pat-framework-innovative-pharmaceutical-development-manufacturing-and-quality-assurance (accessed 2023-03-21).
4. Rolinger, L.; Rüdt, M.; Hubbuch, J. Comparison of U.V.- and Raman-Based Monitoring of the Protein A Load Phase and Evaluation of Data Fusion by PLS Models and CNNs. *Biotechnol Bioeng* **2021**, *118* (11), 4255–4268. https://doi.org/10.1002/bit.27894.
5. Rolinger, L.; Rüdt, M.; Hubbuch, J. A Critical Review of Recent Trends, and a Future Perspective of Optical Spectroscopy as PAT in Biopharmaceutical Downstream Processing. *Anal Bioanal Chem* **2020**, *412* (9), 2047–2064. https://doi.org/10.1007/s00216-020-02407-z.
6. Sukumaran, S. Protein Secondary Structure Elucidation Using FTIR Spectroscopy. 4.
7. Dolui, S.; Mondal, A.; Roy, A.; Pal, U.; Das, S.; Saha, A.; Maiti, N. C. Order, Disorder, and Reorder State of Lysozyme: Aggregation Mechanism by Raman Spectroscopy. *J. Phys. Chem. B* **2020**, 124 (1), 50–60. https://doi.org/10.1021/acs.jpcb.9b09139.
8. Hernández, B.; Pflüger, F.; López-Tobar, E.; Kruglik, S. G.; Garcia-Ramos, J. V.; Sanchez-Cortes, S.; Ghomi, M. Disulfide Linkage Raman Markers: A Reconsideration Attempt: Disulfide Raman Markers. *J. Raman Spectrosc.* **2014**, *45* (8), 657–664. https://doi.org/10.1002/jrs.4521.

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