Use of Raman Spectroscopy and PCA for quality evaluation and out-of-specification identification in biopharmaceutical products

Use of Raman Spectroscopy and PCA for quality evaluation and out-of-specification identification in biopharmaceutical products

Ambra Massei[1,2], Nunzia Falco[2], Davide Fissore[1]

[1] Dipartimento di Scienza Applicata e Tecnologia, Politecnico di Torino, Corso Duca degli Abruzzi 25, 10129 Torino (Torino), Italy

[2] Global Drug Product Development, Merck Serono SpA, Via Luigi Einaudi 11, 00012 Guidonia Montecelio (Roma), Italy

Keywords

Raman spectroscopy, Principal Component Analysis, Monoclonal Antibody, Biopharmaceutical Stress Identification, Spectral Discrimination

Abbreviations

CQA	Critical Quality Attributes
FDA	Food and Drug Administration
HMW	High Molecular Weight
LMW	Low Molecular Weight
mAb	Monoclonal Antibody
PAT	Process Analytical Technology
PCA	Principal Component Analysis
PC	Principal Component
PLS	Partial Least Squares Regression
PLS-DA	Partial Least Squares – Discriminant Analysis
PQA	Product Quality Attributes
ROA	Raman Optical Activity

List of Figures

Figure 8: Scores plot of PCA model developed on the thermally stressed samples in the Raman shift range 150-3425 cm^{-1}. The scores are displayed in the plane described by the 1^{st} and 2^{nd} PC and are grouped according to the exposure duration to the stress (the arrows indicate the increasing exposure to the stress).

Figure 9: Loadings plot obtained by the PCA model developed only on thermally stressed samples according to the type of stress. In particular samples subjected to: (a) lowest stress time; (b) middle stress time and (c) highest stress time. Two zones are highlighted by means of frames: the black one refers to the amino acid signals and the grey one to the amide region.

Figure 10: Workflow of the method used in this paper to distinguish among different stress conditions or dilutions occurred into the samples.

List of Tables

Table 1: Number of samples for each case study considered in the study.

Abstract

Over the past three decades, there was a remarkable growth in the approval of antibody-based biopharmaceutical products. These molecules are notably susceptible to the stresses occurring during drug manufacturing, often leading to structural alterations. A key concern is thus the ability to detect and comprehend these alterations caused by processes, such as aggregation, fragmentation, oxidation levels, as well as the change in protein concentration throughout the process steps, potentially resulting in out-of-spec products.

In the present study, Raman spectroscopy, coupled with Principal Component Analysis (PCA), has proven to be an excellent tool for characterizing protein-based products. Notably, it offers the advantages of being minimally invasive, rapid and relatively insensitive to water. Therefore, it was successfully employed to discriminate between various stresses impacting a monoclonal antibody (mAb). The molecule used in this study is a fully human IgG1 fusion protein. Thermal stress was induced by incubating the samples at 50°C for one month, while oxidative stress was induced by introducing hydrogen peroxide. Additionally, dilutions were performed to explore a broader range of protein concentrations.

Specific key bands were identified in the Raman spectra, which facilitated the PCA classification and allowed for their association with distinct changes in the secondary and tertiary structures of the protein. Notably, it was observed that signals corresponding to amino acids exhibited a decrease in intensity with increasing levels of thermal stress, while other alterations were noted in the amide bands. It was shown that changes in the range 2800-3000 cm^{-1} pertains to the dilution process, while specific peaks of C-H stretching were essential for the discrimination between the oxidative-stressed samples and the thermal and diluted counterparts. Furthermore, the model calibrated on the mAb demonstrated remarkable performance when used to evaluate a different product, e.g. a hormone.

1. INTRODUCTION

1.1 General background

Biopharmaceutical products, including monoclonal antibodies (mAb), stand as leader in the market owing to their highly targeted and efficacious treatments for chronic diseases. [1,2,3] However, the main challenge with these molecules lies in the potential pathways of degradation, which must be identified and monitored. [2] Additionally, the process development of these molecules involves time-consuming and cost-ineffective analyses. [3]

Advanced solutions are required to effectively monitor structural changes in the mAb that could lead to immunogenicity and reduced drug efficiency. [2] Although Raman spectroscopy was widely employed in the identification of small molecule raw materials [4], its application in characterizing protein therapeutics has only recently gained attention. [4,5,6] This surge of interest is attributed to the US Food and Administration (FDA) initiative promoting the adoption of Process Analytical Technology (PAT) for a more innovative approach to pharmaceutical development and manufacturing. [2,3]

Raman spectroscopy is a vibrational spectroscopic technique providing molecular-level insights into the chemical and physical processes that occur during pharmaceutical unit operations. Its notable advantage lies in its compatibility with aqueous solutions, as water exhibits weak signals in Raman spectra. Consequently, Raman spectroscopy has found numerous applications in both liquid and freeze-dried formulations, as documented in the literature, across various stages of pharmaceutical product development/manufacturing. [7,8,9,10,11,12,13,14] Being a spectroscopic technique, multivariate data analysis is needed to extract meaningful insights from the spectra, allowing for reducing laboratory efforts, costs and saving time.

Different unit operations have been investigated, including blending, granulation [15,16,17,18], as well as tableting and coating processes [19,20,21]. For instance, Vergote et al. and De Beer et al. illustrated the utility of Raman spectroscopy in overseeing the blending process, effectively identifying the mixing endpoint. [22,23] In the field of tableting, Kontoyannis et al., Niemczyk et al. and Johansson et al. used Raman spectroscopy for quantifying components in tablets and capsules. [24,25,26] Also, other unit operations strictly related to biopharmaceuticals, like cell culture, chromatography, filtration and lyophilization were deeply studied. Bhatia et al. used Raman spectroscopy to quantify amino acids in aqueous culture solutions, showing prominent results for the quantification of tyrosine, tryptophan and phenylalanine. [27] Nitika et al. coupled Raman

spectroscopy with artificial intelligence for real-time monitoring and control of charge variants during the cation exchange chromatography in mAb production processes. [28] Feidl et al. implemented Raman spectroscopy using an in-line flow cell for in-line monitoring the downstream chromatographic processing of biotherapeutics. [29] Lamsal et al. applied the surface-enhanced Raman spectroscopy to examine fouling during nanofiltration. They demonstrated the possibility to identify functional groups of organics involved in membrane fouling. [30] Krishnan et al. designed a small-scale model setup using a freeze-drying Raman imaging system for monitoring the mannitol crystallization. [31]

Several studies were published on the prediction and discrimination of protein. Pieters et al. used Raman spectroscopy as a potential tool to detect real-time protein conformational changes during freeze-drying process, using specific bands associated with the secondary structure of the protein. This enabled the discrimination between native-like and non-native conformational states of proteins with 95% accuracy. [32] Nitika et al. and Ettah et al. demonstrated the feasibility of Raman spectroscopy in providing information about the secondary and tertiary structure as well as the aggregation mechanisms of antibody. Raman features were used to identify changes in conformation, molecular interactions and hydrogen bonding. [33,34] Wei et al. developed a rapid, multi-attribute and non-invasive method targeting various product quality attributes (PQAs). They showed that from a single spectroscopic scan it is possible to quantitatively measure different PQAs, such as protein concentration, osmolality, pH, polysorbate 20 and methionine concentration. [35]

All the referenced studies primarily focused on the quantitative application of Raman spectroscopy. There is a limited body of work exploring the potential of Raman spectroscopy to discern the specific type of stress affecting a given molecule. When a protein is subjected to a particular stress, it undergoes observable changes in secondary and tertiary structure. Raman spectra were shown to provide insights into the protein's backbone, side-chain groups and can serve to monitor and regulate the higher-order structure (α-helix, β-sheet and random coil conformations). Bethan et al. used Raman spectroscopy to detect post-translational modifications and degradation in mAb following exposure to pH stress, thermal stress, agitation and oxidative stress. They successfully demonstrated the capacity of Raman spectroscopy to differentiate between force-degraded samples based on distinctive features in spectra. [2] Also, Thiagarajan et al. investigated the applicability of Raman and Raman optical activity (ROA) to detect instabilities due to heat stress on an antibody. They found that ROA provided differences in the tertiary structure of the monoclonal antibody. [36]

1.2 Multivariate Data Analysis – Principal Component Analysis (PCA)

Exploratory data analysis is often the first step of any chemometric processing, providing an immediate, direct and visually intuitive portrayal of the data information. Moreover, it allows for the identification of similarities and distinctions among the samples, the presence of groups, clusters or trends. Also, outlier samples can be detected through this analysis. [37]

The data needs to be arranged in the form of a matrix denoted as **X,** as shown in Fig. S1. Each row corresponds to a sample (*N*), labeled by the "Sample ID", while the columns represent the variables (*J*), signifying the Raman intensity values at specific Raman shifts. In Fig. S1a, the column labeled "type of samples" designates the category indicating the stress condition each sample was subjected to. The Raman shift range explored encompassed 150-3425 cm^{-1}. In other words, each row in the matrix encapsulates the entire spectrum of a particular sample, as demonstrated in the close-up view of a generic row in Fig. S1b. All the spectra acquired by Raman were processed using SIMCA (Version 17.0.0.24543, Sartorius, Germany).

Principal Component Analysis (PCA) reduces a complex dataset into a simpler structure leading to a minimum the loss of information. It is based on a bilinear decomposition of the initial dataset **X** that can be described by Eq. 1: [37,38]

$$\boldsymbol{X} = \boldsymbol{T} \cdot \boldsymbol{P}^T + \boldsymbol{E} \quad \textbf{[1]}$$

The loadings matrix **P** represents the principal components (PC), indicating the directions in which the data were projected. It quantifies the influence of each variable in the new reduced system. The scores matrix **T** denotes the samples coordinates within the new (reduced) space. Any residual noise is captured in the matrix of the residuals **E.** [37] These scores and loading vectors in SIMCA resulted in two graphs (score plot and loading plot) which serve to elucidate the outcomes. Specifically, the scores plot accounts for the variance among the data, whereas the loading plot elucidates the factors contributing to this variance. In the context of this paper, the variance arises from various stresses acting on the molecule (as depicted in the score plot), while the loading plot shows the main peaks involved. [5]

A multivariate control chart built according to the statistic Hotelling T^2(T2) was used to identify the outliers. It is calculated by Eq. 2:

$$T2_i = \sum_{a=1}^{A} \frac{t_a^2}{\lambda_a} \quad \textbf{[2]}$$

where t_a is the a-th score and λ_a the corresponding variance. Hotelling's T^2 provides a single statistic that summarizes the separation between groups in a multivariate space. The larger the value of Hotelling's T^2, the greater the difference between the samples.

1.3 Motivation of the work

In the present paper, multivariate data analysis was combined with Raman spectroscopy to investigate their powerful ability in discerning the type of stress impacting a mAb. In fact, Raman spectroscopy exhibited sensitivity to changes in secondary structure induced by factors such as temperature or denaturants. During the pharmaceutical development and manufacturing of therapeutic proteins, oxidation and thermal-induced degradation are frequently encountered. The vulnerability to these stress factors can modify the secondary and tertiary structures of the protein, affecting its stability and biological activity, and potentially impacting the product efficacy and patient safety. Therefore, it is necessary to closely monitor these stresses (and the variations they induce in different CQAs) during the development of therapeutic proteins. [39,40]. The primary aim of this paper was to conduct a feasibility study on the applicability of Raman spectroscopy, using a monoclonal antibody, for differentiating among the stresses and identifying the good samples, i.e., those that were either not subjected to a specific stress or met the specifications set by the company. Therefore, a forced thermal and oxidative degradation study was carried out on the molecule by subjecting it for 34 days at 50°C (with various intermediate sampling points) and by oxidating it with hydrogen peroxide for up to 30 minutes (with distinct intermediate sampling every minute). Also, dilutions were performed on the samples to discriminate different protein concentrations. Furthermore, a focus on the thermal stress phenomenon was carried out to highlight the spectral regions mainly involved by gradually increasing the exposure time to a specific stress. The robustness of the developed model was tested in a different product, a hormone, characterized by a different formulation composition. The decision on which products to use was based on the relevance of the product for this case study. In fact, processing of this molecule appears to be particularly critical and various forced degradation studies were conducted in the past using reference analytical techniques. Additionally, having access to a molecule with a different structure than the antibody, namely a hormone, it was considered for use to observe whether the model would also work with a different product.

2. MATERIALS AND METHODS

2.1 Sample preparation

A monoclonal antibody (formulation A), provided by Merck S.p.A. Rome - Italy (an affiliate of Merck KGaA, Darmstadt, Germany), was used as model protein. Additionally, a hormone (labeled as "formulation B") was employed as an external test set to assess the robustness of the developed model and its applicability to products not included in the initial calibration. Table S1 enumerates the excipients used in the two distinct formulations, emphasizing the disparities between them.

Thermal and oxidative stresses were applied to the protein to assess their impact on various critical quality attributes, including aggregation, fragmentation and oxidation levels. Additionally, dilutions of the drug product with placebo from 40 mg/ml to 2 mg/ml were prepared to cover a wider range of protein concentrations. Each sample was intentionally subjected to extreme stress conditions, resulting in values (e.g., high HMW% and LMW% values) that might not typically occur in real-world processes. Table 1 provides an overview of the number of samples generated for each type of stress. A total of fifty-five spectra were acquired for formulation B, which had not been subjected to any stresses. The generation of the following stresses were performed according to Standard Operating Procedure (SOP) and internal protocols of the company.

Table 1: Number of samples for each case study considered in the work.

Case study	Number of samples
No stress (formulation A)	13
No stress (formulation B)	55
Thermal stress (out-of-spec)	38
Thermal stress (in-spec)	12
Dilutions	29
Oxidative stress	21

The samples were labelled based on the type of stress they were subjected to. Not all samples were stressed: Raman spectra of pure samples were also acquired and labelled as "no stress" for both formulations under study (A and B), while all the other case study were applied only to formulation

A. Moreover, thermal stress resulted in two distinct sample groups: those out-of-specifications, characterized by aggregates values exceeding 5%; and those meeting product specifications, characterized by aggregates values lower than 5% (in accordance with the company's standard operating procedure).

a. Thermal stress

To induce accelerated product degradation, thermal stress was performed by incubating all formulation A based products at 50°C for thirty-four days, exceeding the normal storage condition. To simulate different exposure times, the stress was stopped at different time instants by storing the specific sample at -80°C. According to previous knowledge on the molecule, this thermal stress led to a notable rise in aggregation and fragmentation levels, measured by chromatography and capillary electrophoresis methods. [41,42]

b. Oxidative stress

The oxidative stress was applied to the formulation A based product for a duration up to 20 minutes, by adding 200 µL of hydrogen peroxide (1%). A number of 20 samples was prepared and stressed at the same time. Then, the stress was interrupted every minute (up to 20 minutes) by introducing 120 mg of methionine into the sample. This process yielded 21 samples, each representing a distinct level of oxidation corresponding to a specific time point. [43]

c. Dilution of samples

Samples of formulation A based product were progressively diluted with a placebo, starting from a protein concentration of 40 mg/ml up to 2 mg/ml with a concentration interval of 1 mg/ml.

2.2 Raman spectroscopy acquisition

All Raman spectroscopic acquisitions were performed using a RamanRXN2 instrument (Kaiser Optics, Ann Arbor, MI, USA). The laser has an excitation wavelength of 785 nm, and the power was at its maximum output, 400 mW. Each sample (6 mL) was loaded into 10R vials and the distance between the probe and the bottom of the vial was optimized to maximize the signal-to-noise ratio. An immersible contact probe was used for the measurements. Spectra in the range of 150-3425 cm^{-1}

were acquired with a 4 cm^{-1} resolution. Each spectrum was collected for 2.25 minutes (3 scans of 45 second exposure time per scan). Aluminum foils were used to protect samples from light.

2.3 Multivariate Data Analysis

Spectroscopic techniques are extensively employed in pharmaceutical product development due to their non-invasive and non-destructive nature, applicable to both off-line and in-line analyses. They provide a huge amount of information that can be explored using multivariate data analysis. [37]

Preprocessing is an essential step for every spectrum, as it enables the elimination of signals unrelated to the target property or sample differentiation. This process effectively diminishes chemically irrelevant variations, thus accentuating the pertinent ones. [44]

All the data were pretreated using the first-derivative tool as a baseline correction tool. Moreover, the smoothing step was useful to remove random noise, preserving useful spectra information. In the present case, Savitzky-Golay algorithm was used setting up a second polynomial order with a 15-window size, leading to noise reduction. This filter is based on a polynomial equation fitted in a least squares sense with a predefined interval of spectral points. [45] Then, the pre-treated data were used to develop the PCA model.

3. RESULTS AND DISCUSSION

To evaluate the suitability of Raman spectroscopy for identifying stress factors affecting a monoclonal antibody, various PCA models were developed. The initial model considered the entire dataset, while a stress-dependent model was dedicated to the thermally stressed samples, aiming to elucidate the physical mechanisms involved by changing the magnitude of the stress. Moreover, to assess the robustness of the model, a validation test was conducted using a different product, a hormone, as prediction dataset in the previous developed model (specific for the antibody).

3.1 Peak assignments and physical explanation

Raman spectroscopy is used in the context of proteins to characterize the secondary structure, made up of hydrogen bonding between amino acids in the polypeptide chain. Being a vibrational technique, it lies on the vibrations of chemical bonds, associated to the backbone C-N, C-C and C-O bonds in proteins. Moreover, the amide bands are particularly important in Raman spectroscopy for studying the secondary structure in proteins. Specifically, the amide I band arises primarily from C=O stretching vibrations and is the most informative for secondary structure analysis, where the specific signals for alpha helixes, beta sheets and random coils are located.

In Fig. 1, the pretreated spectra are reported for each case-study. A zoom of the range of acquisition was reported up to 1800 cm^{-1}, corresponding to the fingerprint region of the antibody in the Raman spectra.

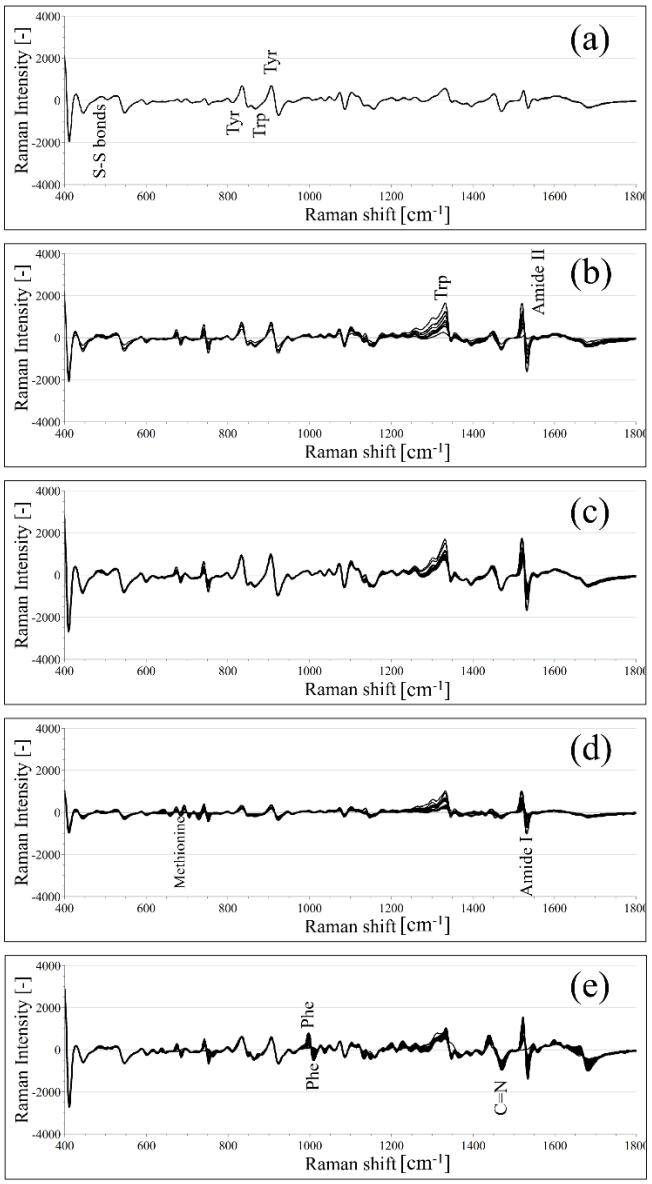


Figure 1: Pre-processed Raman spectra of formulation A samples in the fingerprint region for the different types of samples: (a) no stress, (b) thermal in-spec, (c) thermal out-of-spec, (d) oxidative and (e) diluted samples.

In a Raman spectrum, each molecule is characterized by specific peaks linked to the vibrational frequency of different functional groups. It is difficult to assign each peak in Raman spectra, but it is more common to talk about different regions or zones. The analyzed stresses might generate modifications on the secondary and tertiary structures of the protein, since aggregation, fragmentation and oxidation phenomena are involved. A mAb formulation is made up of protein and excipients in aqueous buffer. Therefore, the amide I ($1600 - 1700\ \mathrm{cm}^{-1}$), II ($1500 - 1600\ \mathrm{cm}^{-1}$) and III ($1200 - 1300\ \mathrm{cm}^{-1}$) regions are used for characterizing changes in the secondary structures of the protein, since the carbonyl (C=O) and the amine (N-H) groups of the protein contribute to hydrogen bonding in α-helix, β-sheet and random coil. Moreover, to stabilize the tertiary structure of the

protein, disulfide bridges and aromatic amino acids in the side chains are present, leading peaks in the spectra between 500 cm^{-1} and 700 cm^{-1}. The aromatic residues of amino acids cause many vibrational bands in the fingerprint region, such as those at 702 cm^{-1}(methionine), 835 and 850 cm^{-1} (tyrosine), 870 cm^{-1} (tryptophan), 1000 cm^{-1} and 1027 cm^{-1} (phenylalanine). The C-N stretching vibrations in combination with N-H bending are attributed to the secondary structure of the protein in the region of the amide-II bond (1520 cm^{-1}). Moreover, the CH_2 bending and CH deformation are related to the primary structure of the protein. The peak at 1121 cm^{-1} is related to the C-N bond and the one at 1454 cm^{-1} at the C-H deformation. The peaks in the amide-I region around 1680 cm^{-1} are attributed to the β-sheets. [2, 34, 35, 46, 47,48]

By inspecting also Figure S2, it is possible to have a look on the raw Raman spectra. In Fig. S3 the pretreated spectra in the entire range of acquisition for each case-study are reported, highlighting the specific peak of O-H stretch at 3148 cm^{-1} and the C-H stretching region at around 2920 cm^{-1}.

3.2 PCA on the entire dataset

The discussion of the PCA results is based on the interpretation of the scores and loadings and is focused on the groupings related to the kind of stress the sample was subjected to. Any outlier samples were identified and removed from the dataset. These exhibited anomalous Raman spectra, attributed to erroneous spectrum acquisition. Therefore, PCA allowed to identify outlier samples, different from the majority present in the analyzed dataset.

One single PCA model was developed including all the stress, but for a clearer and less confusing visualization, the results have been presented in two separate score plots for PC1 vs PC2 (Figure 2a and 2b). The first three PCs resulted to be the most meaningful and are reported in Figure 2 and Figure 3. As a matter of fact, they explained the 99% of the variance, so most of the information contained in the raw data. In Figure 2a and 2b, PCA provides a clear grouping trend related to the kind of stress subjected by the samples (refer to the legend for interpreting the groups). In fact, by inspecting PC1 vs PC2, a distinct separation into two regions becomes evident: the oxidative-stressed samples are situated in the positive PC1 range, while both the thermal-stressed (out-of-spec) and diluted samples are found in the negative range. The samples labeled as “thermal in-spec”, and the “no stress” ones are overlapping in the central region of the plot. This is a successfully proof of concept of this technology for the aim. In fact, "thermal in-spec" samples are characterized by a lower magnitude of stress, yielding HMW% and LMW% values that conform to the company's specified standards, as it was previously reported. Consequently, from the standpoint of statistical quality control, it is favorable that PCA accurately identifies these samples as being in good agreement with the non-

stressed ones. A close inspection of Fig. 3, relating PC1 vs PC3, provides information about a trend related to the protein concentration. By moving from negative to positive values of PC3, the protein concentration increases in the formulation from 2 mg/ml up to 40 mg/ml. A link with the loading plot, reported in Figure 4, can be found.

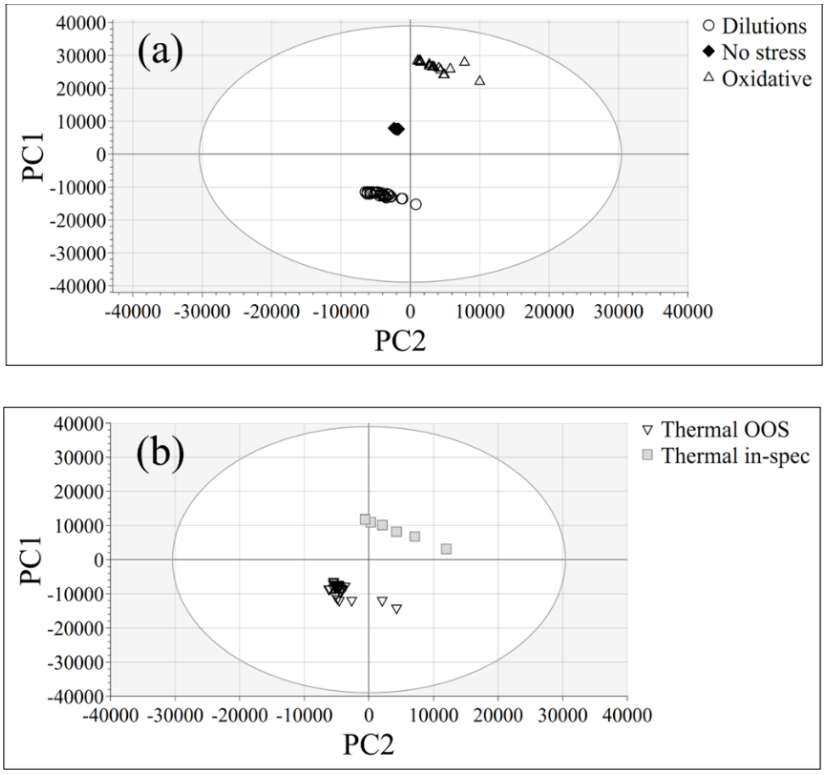


Figure 2: Scores Plot of the entire dataset using a single PCA model in the Raman shift range 150-3425 cm^{-1}. The scores are displayed in the plane described by the 1st and 2nd PC and are grouped according to the type of samples analyzed. To facilitate the visualization, diluted, no stress and oxidized samples are shown in (a) and the thermal-stressed samples in (b). In plot (a), three distinct clusters are observed for the unstressed samples, subjected to dilution or oxidative stress. In plot (b), however, two clusters are observed for the thermally stressed samples, showing distinctions between those within and outside the specification. Specifically, it is noted that the samples within specification roughly fall within the same area as the unstressed samples in plot (a).

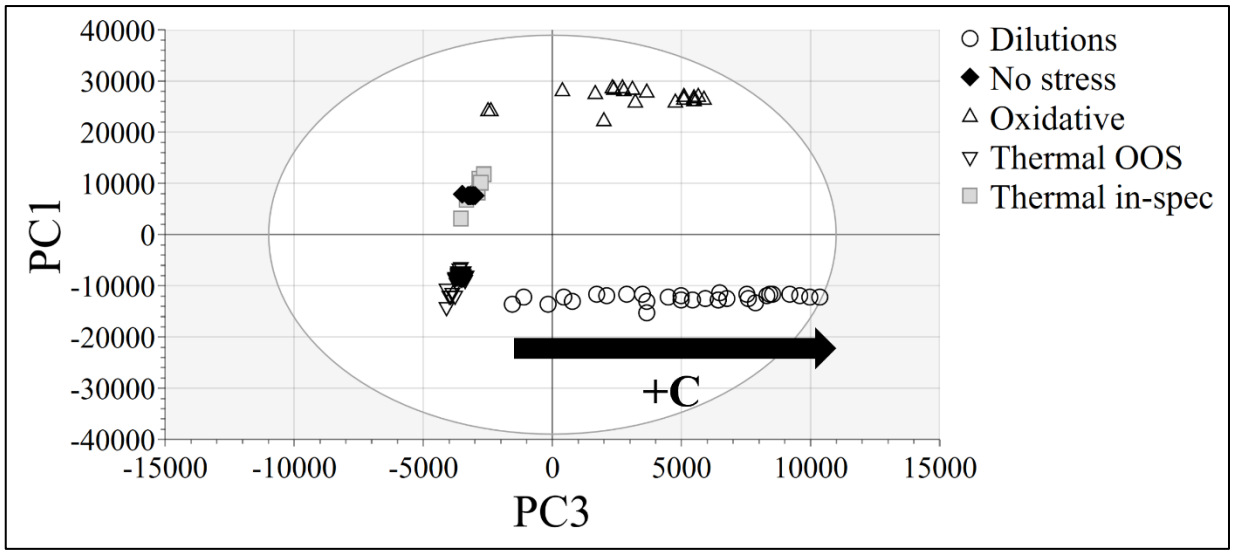


Figure 3: Scores Plot of the entire dataset in the Raman shift range 150-3425 cm^{-1}. The sores are displayed in the plane described by the 1st and 3rd PC and are grouped according to the type of samples analyzed. Moreover, the diluted samples are ordered according to the protein content.

Figure 4 depicts the loadings plots, leading to understand which are the most significant variables for each type of stress, characterized by higher peaks in the trend of the loadings. The loadings plots are used as a link with the original spectral data, facilitating the interpretation of the scores in accordance with actual signals. A close inspection of Fig. 4 reveals the specific peaks of C-H stretching to be essential for the discrimination between the oxidative-stressed samples and the thermal and diluted counterparts. Additionally, variations in the fingerprint regions are found, likely attributable to structure rearrangements at the secondary and tertiary levels. In Fig. 4a, PC1 shows an increase in the peak intensity at around 2800 cm^{-1} for the thermal-stressed (out of specification) and diluted samples, being located in the negative PC1 values in the score plot reported in Figure 2a and 2b. For these two stress factors also another peak around 3000 cm^{-1} acquire notably importance, mainly related to the O-H- and C-H stretching, as deeply analyzed in Section 3.1. Figure 4b also highlights the peak at 2800 cm^{-1} is relevant for the samples located in the region of positive PC2 values in the scores plot (see Figure 2a and 2b). In other words, the oxidative stressed samples and the thermally one (in specification) are positively correlated with PC2: an increase in the signal at 2800 cm^{-1} leads to higher score values. On the other hand, for the samples located in the negative part of Figure 2a and 2b (diluted samples, no-stressed and thermally stressed out of specification) the opposite occurs, meaning that they are anti-correlated. Regarding protein content, Fig. 4c provides the loading plot for PC3. The range between 2800 and 3000 cm^{-1} predominantly pertains to the dilution process. It stands to reason that as the protein content in the formulation decreases, the number of bonds between the protein itself and the excipients is likely to diminish, resulting in variations in peak intensity.

Notably, the peak at around 2870 cm^{-1} exclusively appears in the diluted samples, absent in the non-stressed ones, indicating its association with the alteration in protein concentration. Furthermore, variations in the amide-I bands are observed, playing a role in the dilution process.

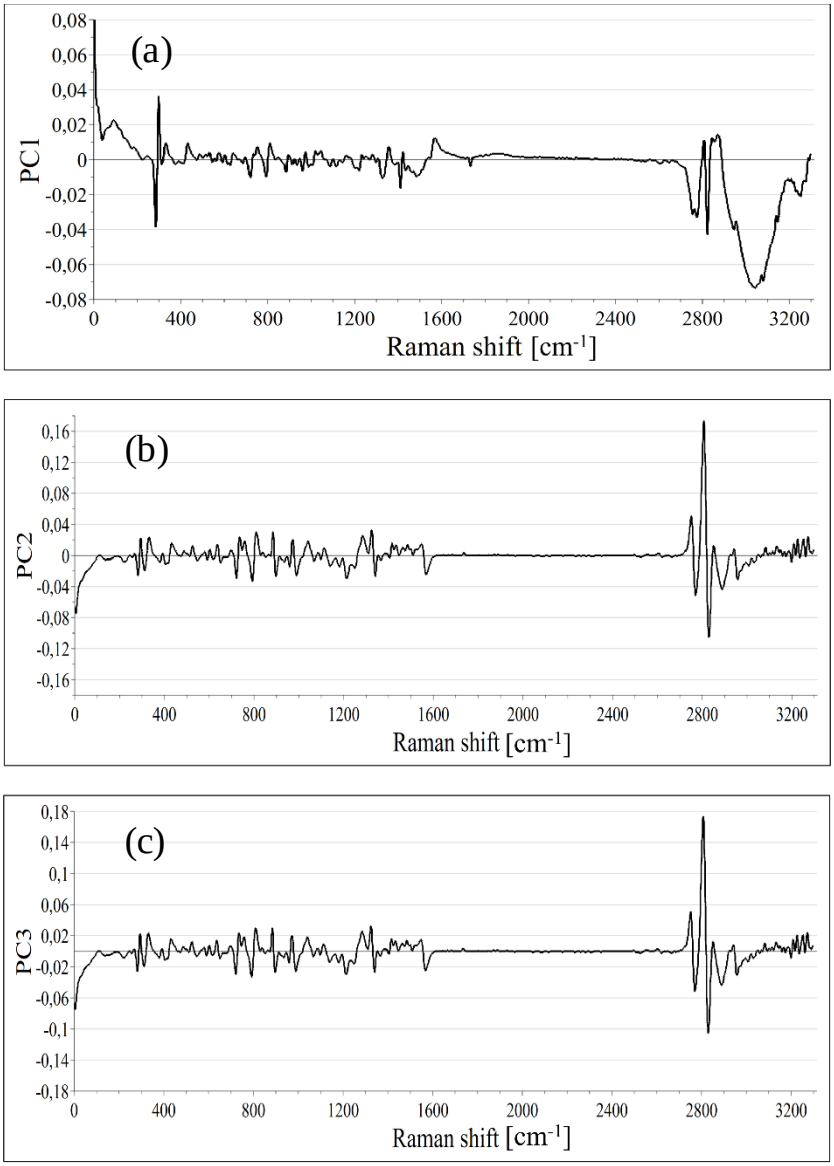


Figure 4: Loading plots for the first three PCs showing (a) PC1, (b) PC2 and (c) PC3 obtained by the single PCA model developed on the entire Raman shift range by considering the whole dataset.

The developed model encompassed the entire dataset during the calibration phase. This allowed to evaluate the potential of Raman spectroscopy in discerning between the different stresses the protein experienced. However, it did not shed light on whether the model could be applied to different samples not included in the calibration process. To address this, the dataset was split into two portions: for each stress, 2/3 of the data served as the calibration set, while the remaining 1/3 was

allocated for prediction. The projection scores plot, shown in Fig. 5b, demonstrates the model capacity to project outcomes for new samples, based solely on the calibration dataset (calibration scores plot in Fig. 5a). Thus, if a new sample is generated, it can be assessed by the model to determine its position on the graph. PCA is an unsupervised model that does not allow for prediction in its usual understanding. In this case, the new unseen samples were projected on the PCA space defined by the training samples. Furthermore, as validation, the graph depicting Hotelling's T2 value in relation to the sample is reported in Figure 6 for the prediction datasets. Notably, in the calibration dataset only three points (90, 95 and 98) were identified as outliers; whereas, in the prediction dataset, there was only one outlier. Upon investigation, it was found that these points resulted from an issue occurred during the acquisition of Raman spectra, confirming the anomalous behaviour in the PCA model.

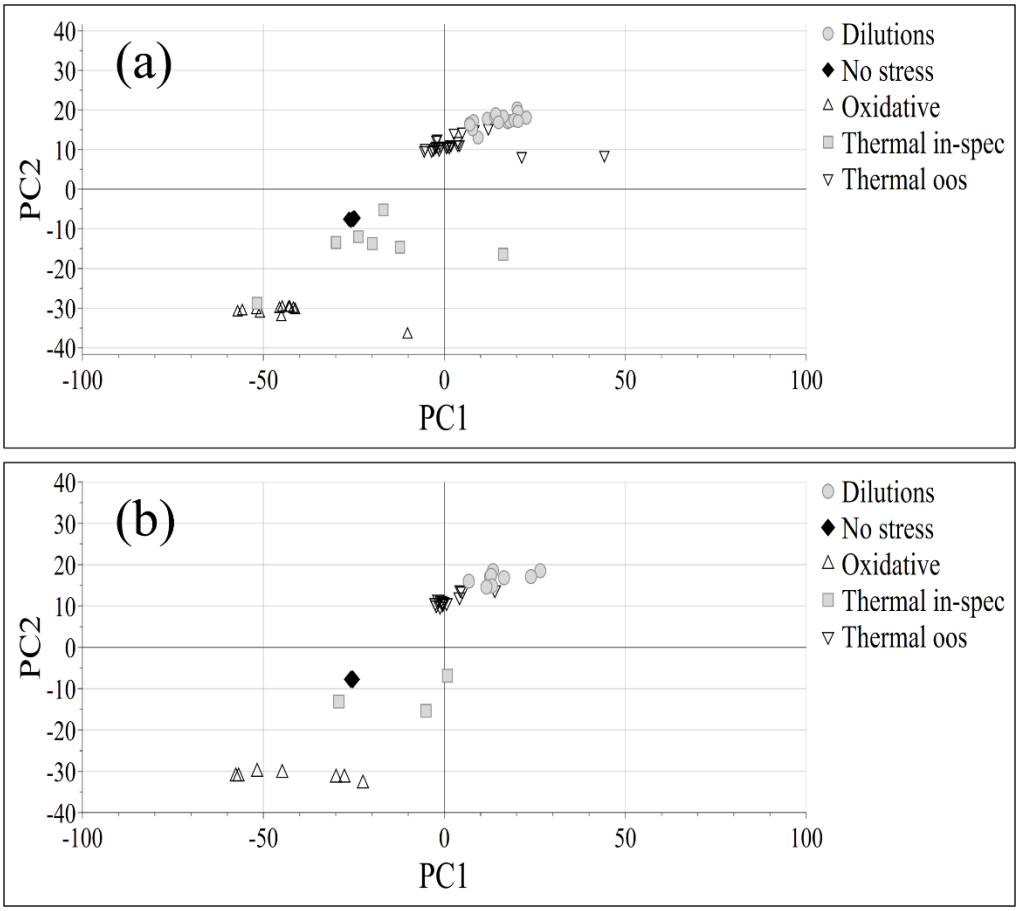


Figure 5: Scores plot of PCA model developed on the entire range for (a) calibration dataset and (b) prediction dataset obtained using only 1/3 of the data for each type of category.

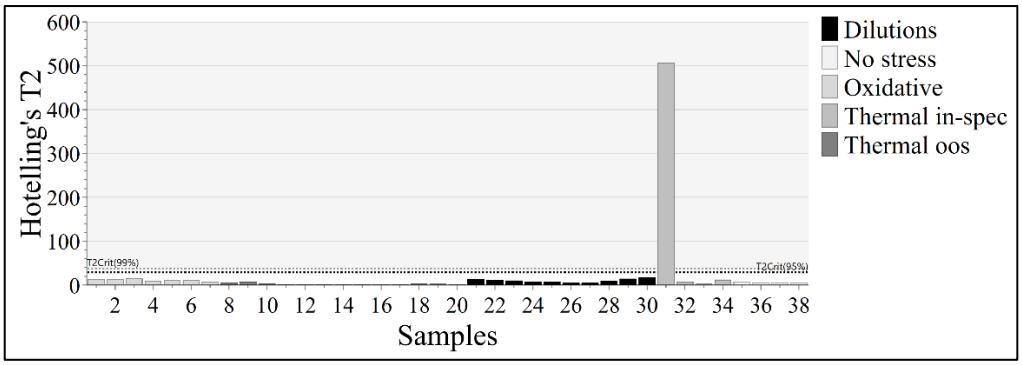


Figure 6: Hotelling's T2 graph for the prediction dataset

An external validation was conducted by applying the model to project in the previous reduced space a different product than the one used in the calibration phase. Specifically, the model, initially developed using an antibody molecule, was tested using a hormone. The two molecules have different structures, but the bonds involved in the Raman spectrum are quite similar. Only the projection score plot and Hotelling's T2 are presented in Fig. 7a to emphasize the robustness of the developed model. It is worth noting that there was only one outlier, which was characterized by a notably high T2 value, as shown in Fig. 7b.

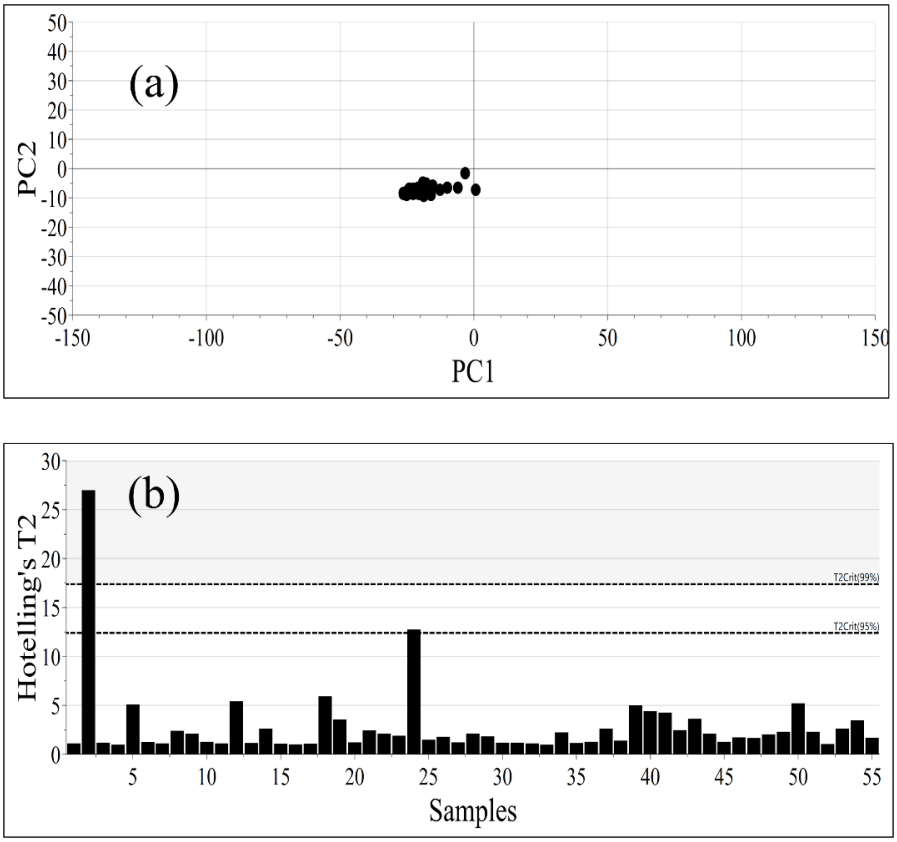


Figure 7: (a) Projection score plot for formulation B product and (b) corresponding Hotelling's T2 graph

3.3 PCA only on thermal stressed samples

To gain a deeper understanding of what occurs during thermal stress in the protein and which bonds are primarily involved in this phenomenon, a PCA was conducted focusing exclusively on the thermally stressed samples.

Initially, the attention was focused on the entire dataset of thermally stressed samples, aiming to elucidate the alterations taking place in the samples during the thermal stress. It is expected that different bonds are involved depending on the duration of exposure to thermal stress, leading to changes in the spatial arrangement of the molecules.

Figure 8 displays the scores plot relating PC1 vs PC2. In Fig. 8 the PCA model discerns three distinct groups. Samples on the right portion of the graph, characterized by the lowest stress exposure (up to 7 days), are identified in the region with positive values of PC2 and negative values of PC1. In the central portion of the graph, moderately stressed samples are found, further dividing into three subgroups. Conversely, samples exposed to the highest levels of stress are identified by negative values of both PC1 and PC2. These two principal components, accounting for approximately 94% of the variance, successfully categorize the thermally stressed samples into three major groups, offering a clear distinction. Furthermore, an increasing trend is observed when moving diagonally for the lowest stressed samples (from positive values of PC2 to 0) and for the highest stressed ones (from negative values of PC2 to 0). On the contrary, in the central region, a more stochastic behavior is noted.

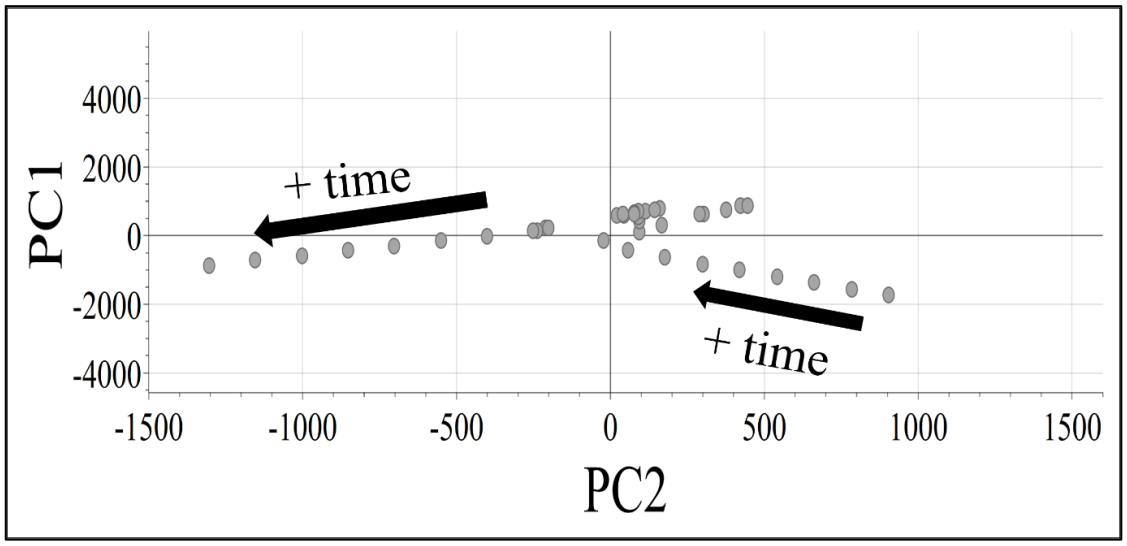


Figure 8: Scores plot of PCA model developed on the thermally stressed samples in the Raman shift range 150 -3425 cm^{-1}. The scores are displayed in the plane described by the 1st and 2nd PC and are grouped according to the exposure duration to the stress (the arrows indicate the increasing exposure to the stress).

Upon examining PC3, as illustrated in Fig. S4, it becomes evident that a discernible pattern emerges within the middle-stressed samples, allowing for a clear differentiation among the three subgroups. The green samples exhibit an ordered arrangement, with stress exposure increasing along the horizontal axis, progressing from negative to positive values of PC3. This indicates that PC3 effectively characterizes the escalating stress levels of the middle-stressed samples. Notably, the yellow samples are located alongside the highest-stressed ones, while the purple ones form a distinct and isolated group.

According to the entity of the stress, different regions in the spectra are affected due to physical changes in the structure of the molecules. As a matter of fact, a distinct PCA was performed for each individual cluster, and the resulting loading plots were examined to determine the corresponding influential variables based on the intensity of stress experienced by each sample. Fig. 9 illustrates the loading plots for the narrowed range, providing insight into how the spectral signals change with prolonged exposure to thermal stress. Only PC1, which captures the maximum variance and thus most of the information in the data, was presented.

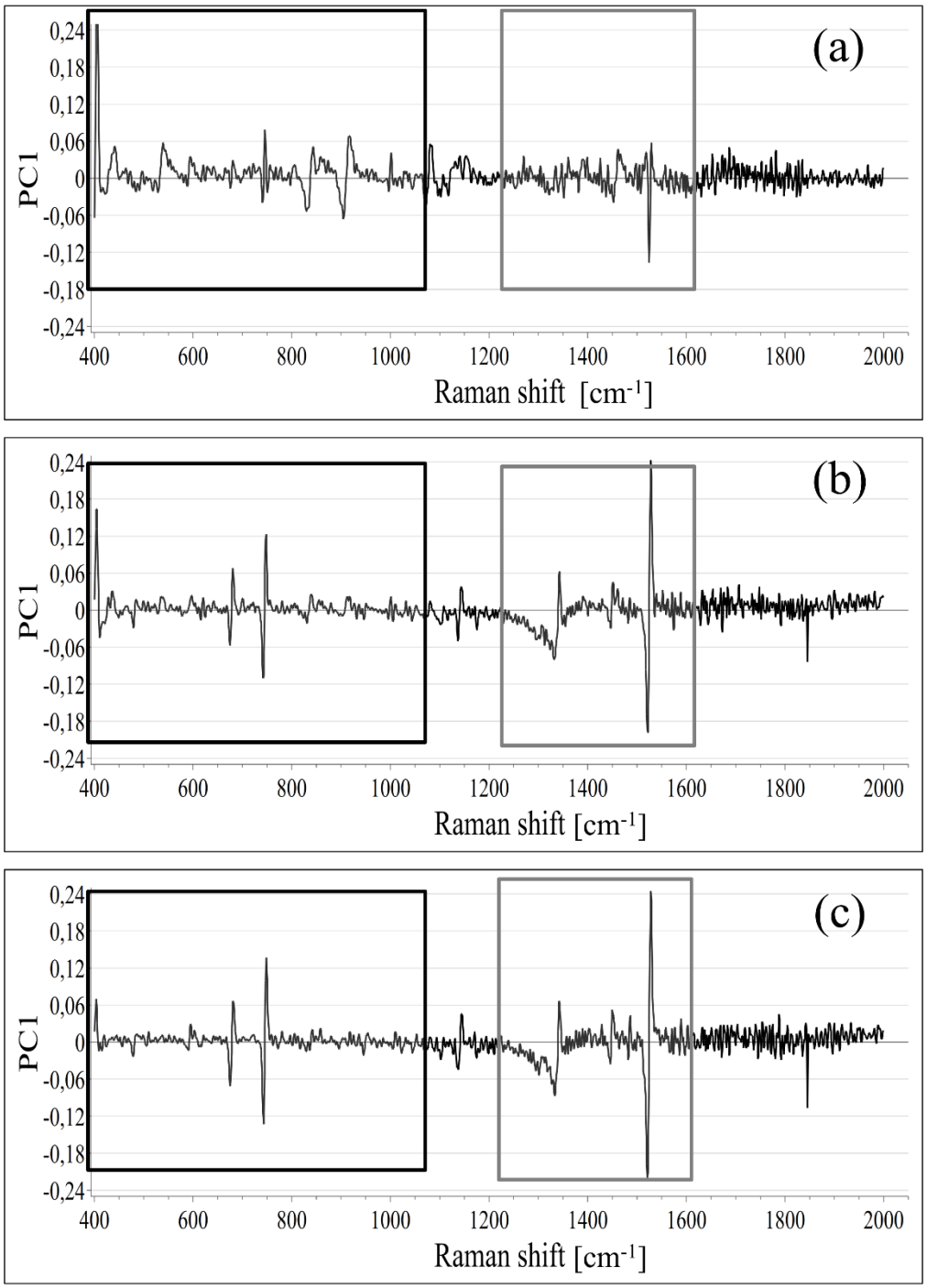


Figure 9: Loadings plot obtained by the PCA model developed only on thermally stressed samples according to the type of stress. In particular, samples subjected to: (a) lowest stress time; (b) middle stress time and (c) highest stress time. Two zones are highlighted by means of frames: the black one refers to the amino acid signals and the grey one to the amide region.

As evident by the highlighted frames, the spectral regions corresponding to amino acid bands (Raman shift lower than 1000 cm^{-1}) show signals diminishing in intensity with prolonged exposure to thermal stress. This suggests that certain crucial bonds may be disrupted due to the aggregation and

fragmentation phenomena. Additionally, with extended exposure to thermal stress over time, it is expected that all associated phenomena progress, resulting in more pronounced alterations in the secondary and tertiary structure. This accounts for the observed increase in Raman signals intensity of the amide bands with escalating stress levels.

It has to be highlighted that the same analysis was conducted also on the oxidative stressed samples, but no valuable insights were obtained, while in the case of the thermal stressed ones different behaviors according to the entity of the stress was pointed out.

3.4 Lessons learned

To sum up, the workflow reported in Figure 10 summarizes the pipeline for the development of classification models, particularly useful during stability studies or in process developments of drug products in pharmaceutical company.

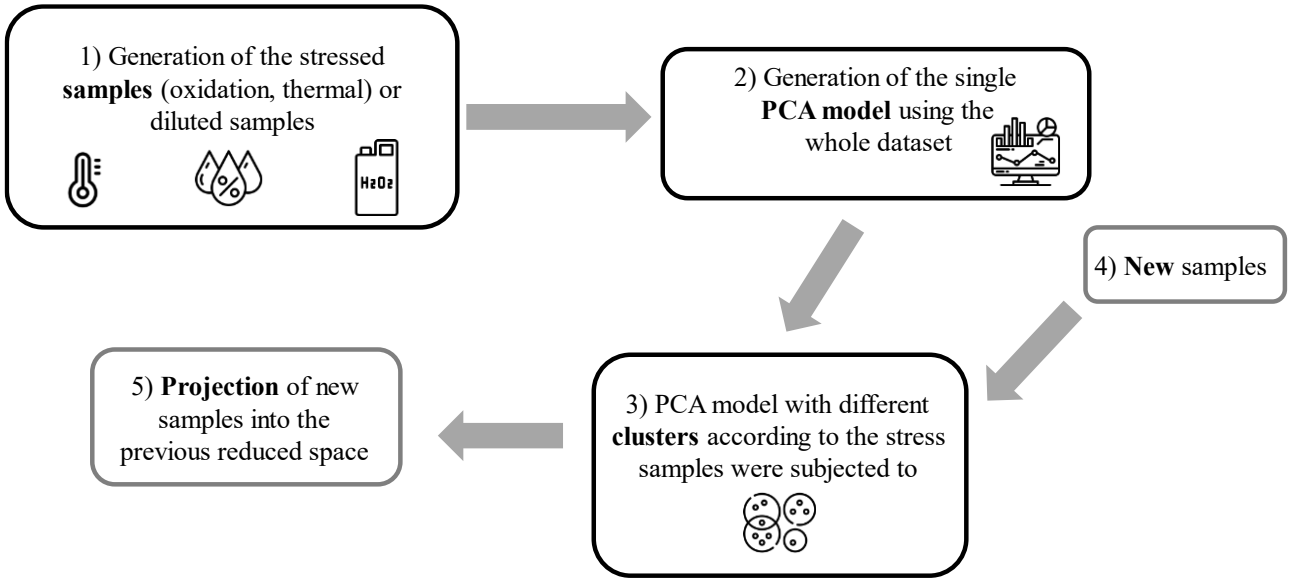


Figure 10: Workflow of the method used in this paper to distinguish among different stress conditions or dilutions occurred into the samples.

The steps to be followed are thus summarized in the followings:

- Samples must be generated, being aware to explore different stress conditions or dilution levels.
- A single data matrix is obtained and used to develop a PCA model. In this step, pay attention to the pre-processing of the dataset.

- Once the model has been obtained, check if clusters are present in the scores plot. If not, it may be that the analytical technique you are using do not see any differences in the samples. If yes, it means that the analytical technique is suitable for the purpose and is able to see differences in the samples.
- New samples, different from the ones used for the development of the PCA models, can be used for robustness purposes.
- The new samples will be projected into the previous reduced spaces, using the loadings previously found.

The limitation of the present model is that it cannot allow for prediction, but only for projection of samples in the reduced space. Therefore, a classification model is needed to gain the predictivity feature.

4. CONCLUSIONS

The feasibility study shown in this paper successfully employed Raman spectroscopy to discern various types of stress affecting the protein.

The technique demonstrated to be powerful in distinguishing between thermally stressed, oxidized, and diluted samples. Additionally, non-stressed samples were distinctly classified into a separate cluster, together with the in-specification samples, by the PCA model.

Furthermore, the method facilitated an in-depth analysis of samples subjected to thermal stress, enabling the identification of three distinct groups based on the magnitude of the stress they experienced. This underscores the robustness and versatility of Raman spectroscopy in characterizing protein stress responses.

The conclusions presented here can be viewed as a significant aid for quantitative analysis. When a new sample undergoes various processing steps, it will inevitably undergo changes in its Critical Quality Attributes (CQAs). By incorporating it into the PCA model, one can discern which CQAs were impacted by the process, as well as the specific bonds and spectral regions involved. Consequently, you can determine the necessary PLS model to quantify the extent of the stress and pinpoint the region that necessitates concentrated model refinement.

However, it has to be highlighted that PCA is an unsupervised model, not allowing for a prediction of new samples in its usual understanding. Therefore, a classifier model like Partial Least Squares – Discriminant Analysis (PLS-DA) could be developed as future steps to create a predictive model based on the preliminary results obtained by the reported exploratory data analysis.

REFERENCES

[1] Biologics Market by Types (Monoclonal antibodies, therapeutic proteins and vaccines) by trends, by regions and by key players – Global forecast to 2021. The business research company.

[2] B.S. McAvan, L.A. Bowsher, T. Powell, J.F. O'Hara, M. Spitali, R. Goodacre and A. J. Doig. *Raman Spectroscopy to monitor post-translational modifications and degradation in monoclonal antibody therapeutics.* Anal. Chem. 92 (2020) pp. 10381-10389. https://doi.org/10.1021/acs.analchem.0c00627

[3] M.K. Shukla, P. Wilkes, N. Bargary, K. Meagher, D. Khamar, D. Bailey and S.P. Hudson. *Identification of monoclonal antibody drug substances using non-destructive Raman spectroscopy*. Spectrochim. Acta A Mol. Biomol. 299 (2023) pp. 1386-1425. https://doi.org/10.1016/j.saa.2023.122872

[4] R.J. Soto, D.S. Meriage, R.L. Wagner, T. Wang and D.J. Semin. *Enabling rapid identification of biotherapeutic protein products using handheld Raman spectrometers and principal component analysis.* J. Raman Spectrosc. 52 (2021) pp. 1281–1293. https://doi.org/10.1002/jrs.6130

[5] C.L. Morais, K.M. Lima, M. Singh and F.L. Martin. *Tutorial: multivariate classification for vibrational spectroscopy in biological samples*. Nat. Protoc. 15 (2020) pp. 2143–2162. https://doi.org/10.1038/s41596-020-0322-8

[6] A.A. Makki, V. Massot, H.J. Byrne, R. Respaud, D. Bertrand, E. Mohammed, I. Chourpa and F. Bonnier. *Understanding the discrimination and quantification of monoclonal antibodies preparations using Raman spectroscopy*. J. Pharm. Biomed. Anal. 194 (2021), 113734. https://doi.org/10.1016/j.jpba.2020.113734

[7] B. Nagy, A. Farkas, E. Borbas, P. Vass, Z.K. Nagy and G. Marosi. *Raman spectroscopy for PAT of pharmaceutical secondary manufacturing.* AAPS PharmSciTech. 20 (2019). https://doi.org/10.1208/s12249-018-1201-2.

[8] A. Paudel, D. Raihada and J. Rantanen. *Raman spectroscopy in pharmaceutical product design.* Adv. Drug Deliv. Rev. 89 (2015), pp. 3-20. https://doi.org/10.1016/j.addr.2015.04.003

[9] D. Bhumika, D. Patel and P.J. Mehta. *An overview: application of Raman spectroscopy in pharmaceutical field.* Bentham Science. 6 (2010), pp. 131-141. https://doi.org/10.2174/157341210791202654

[10] J. Rantanen. *Process analytical applications of Raman spectroscopy.* J. Pharm. Pharmacol. 59 (2007), pp. 171-177. https://doi.org/10.1211/jpp.59.2.0004

[11] T. R. M. De Beer, M. Alleso, F. Goethals, A. Coppens, Y. Vander Heyden, H. Lopez De Diego, J. Rantanen, F. Verpoort, C. Vervaet, J. P. Remon and W. R. G. Baeyens. *Implementation of a Process Analytical Technology System in a Freeze-Drying Process Using Raman Spectroscopy for In-Line Process Monitoring.* Anal. Chem. 79 (2007), pp. 7992-8003. https://doi.org/10.1021/ac070549h

[12] S. Romero-Torres, H. Wikstrom, E. R. Grant and L. S. Taylor. *Monitoring of mannitol phase behaviour during freeze-drying using non-invasive Raman spectroscopy.* J. Pharm. Sci. Technol. 61 (2007), pp. 131-145.

[13] T. R. M. De Beer, M. Wiggenhorn, R. Veillon, C. Debacq, Y. Mayeresse, B. Moreau, A. Burggreaeve, T. Quinten, W. Friess, G. Winter, C. Varvaet, J. P. Remon and W. R. G. Baeyens. *Importance of Using Complementary Process Analyzers for the Process Monitoring, Analysis, and Understanding of Freeze Drying.* Anal. Chem. 81 (2009), pp. 7639-7649. https://doi.org/10.1021/ac9010414

[14] H. de Waard, T. De Beer, W. L. J. Hinrichs, C. Vervaet, J. P. Remon and H. W. Frijlink, "Controlled Crystallization of the Lipophilic Drug Fenofibrate During Freeze-Drying: Elucidation of the Mechanism by In-Line Raman Spectroscopy," American Association of Pharmaceutical Scientists. AAPS PharmSciTech. 12 (2010), pp. 569-575. https://doi.org/10.1208/s12248-010-9215-z

[15] H. Wikström, W. J. Carroll and L. S. Taylor. *Manipulating theophylline mono-hydrate formation during high-shear wet granulation through improved understanding of the role of pharmaceutical excipients.* Pharm. Res. 25 (2008), pp. 923-935. https://doi.org/10.1007/s11095-007-9450-x

[16] H. Wikström, P. J. Marsac and L. S. Taylor. *In-line monitoring of hydrate formation during wet granulation using Raman spectroscopy*. J. Pharm. Sci. 94 (2005), pp. 209-219. https://doi.org/10.1002/jps.20241

[17] G. Walker, S. E. J. Bell, M. Vann, D. S. Jones and G. Andrews. *Characterization of fluidized bed granulation processes using in-situ Raman spectroscopy*. Chem. Eng. Sci. 64 (2009), pp. 91-98. https://doi.org/10.1016/j.ces.2008.09.011

[18] G. Walker, S. E. J. Bell, M. Vann, D. S. Jones and G. Andrews. *Fluidized bed characterization using Raman spectroscopy: applications to pharmaceutical processing*. Chem. Eng. Sci. 62 (2007), pp. 3832-3838. https://doi.org/10.1016/j.ces.2007.04.017

[19] S. Romero-Torres, J. D. Perez-Ramos, K. R. Morris and E. R. Grant. *Raman spectroscopic measurements of tablet-to-tablet coating variability.* J. Pharm. Biomed. Anal. 38 (2005), pp. 270-274. https://doi.org/10.1016/j.jpba.2005.01.007

[20] S. Romero-Torres, J. D. Perez-Ramos, K. R. Morris and E. R. Grant. *Raman spectroscopy for tablet coating thickness quantification and coating characterization in the presence of strong fluorescent interface*. J. Pharm. Biomed. Anal. 41 (2006), pp. 811-819. 10.1016/j.jpba.2006.01.033

[21] J. Muller, K. Knop, J. Thies, C. Uerpmann and P. Kleinebudde. *Feasability of Raman spectroscopy as PAT tool in active coating.* Drug Dev. Ind. Pharm. 36 (2010), pp. 234-243. https://doi.org/10.3109/03639040903225109

[22] G. J. Vergote, T. R. M. De Beer, C. Varvaet, J. P. Remon, W. R. G. Baeyens, N. Diericx and F. Verpoort, *In-line monitoring of a pharmaceutical blending process using FT-Raman spectroscopy*. Eur. J. Pharm. Sci. 21 (2004), pp. 479-485. https://doi.org/10.1016/j.ejps.2003.11.005

[23] T. R. M. De Beer, C. Bodson, B. Dejaegher, B. Walczak, P. Vercruysse, A. Burggraeve, A. Lemos, L. Delattre, Y. Vander Heyden, J. P. Remon, C. Vervaet and W. R. G. Baeyens. *Raman spectroscopy as a process analytical technology (PAT) tool for the in-line monitoring and understanding of a powder blending process*. J. Pharm. Biomed. Anal. 48 (2008), pp. 772-779. https://doi.org/10.1016/j.ejps.2003.11.005

[24] C. Kontoyannis. *Quantitative determination of $CaCO_3$ and glycine in antacid tablets by laser Raman spectroscopy.* J. Pharm. Biomed. Anal. 13 (1995), pp. 73-76. https://doi.org/10.1016/0731-7085(94)00119-m

[25] T. Niemczyk, M. Delgado-Lopez and F. Allen. *Quantitative determination of bucindolol concentration in intact gel capsules using Raman spectroscopy*. Anal. Chem. 70 (1998), pp. 2762-2765. https://doi.org/10.1021/ac971252u

[26] J. Johansson and S. Folestad. *Raman spectroscopy opening the PAT toolbox*. European Pharmaceutical Review, pp. 36-42, 2003.

[27] Bhatia, H., Mehdizadeh, H., Drapeau, D., and Yoon S. *In-line monitoring of amino acids in mammalian cell cultures using raman spectroscopy and multivariate chemometrics models.* Eng. Life Sci. 18 (2018), pp. 55-61. https://doi.org/10.1002/elsc.201700084

[28] Nitika, N., Keerthiveena, B., Thakur, G., and Rathore, A.S. *Convolutional Neural Networks Guided Raman Spectroscopy as a Process Analytical Technology (PAT) Tool for Monitoring and Simultaneous Prediction of Monoclonal Antibody Charge Variants.* Pharm. Res. 41 (2024), pp. 463-479. https://doi.org/10.1007/s11095-024-03663-9

[29] Feidl, F., Garbellini, S., Vogg, S., Sokolov, M., Souquet, J., Broly, H., Butté, A., and Morbidelli, M. *A new flow cell and chemometric protocol for implementing in-line Raman spectroscopy in chromatography.* Biotechnol. Prog. 35 (2019), e2847. https://doi.org/10.1002/btpr.2847

[30] Lamsal, R., Harroun, S.G., Brosseau, C.L., and Gagnon, G.A. *Use of surface enhanced Raman spectroscopy for studying fouling on nanofiltration membrane.* Sep. Purif. Technol. 96 (2012), pp. 7-11. https://doi.org/10.1016/j.seppur.2012.05.019

[31] Krishnan, S., Cao, W., and Phillips, J. *Applying Raman spectroscopy to design lyophilization cycles for protein formulation development.* Am. Pharm. Rev. 12 (2009). https://www.researchgate.net/publication/228911839_Applying_Raman_Spectroscopy_to_Design_of_Lyophilization_Cycles_for_Protein_Formulation_Development

[32] S. Pieters, Y. Vender Heyden, J.M. Roger, M. d'Hondt, L. Hansen, et al. *Raman spectroscopy and multivariate analysis for the rapid discrimination between native-like and non-native states in freeze-dried protein formulations.* Eur. J. Pharm. Biopharm. 85 (2013) pp. 263-271. https://doi.org/10.1016/j.ejpb.2013.03.035

[33] N. Nitika, H. Chhabra and A. S. Rathore. *Raman spectroscopy for in situ, real time monitoring of protein aggregation in lyophilized biotherapeutic products*. Int. J. Biol. Macromol. 15 (2021), pp. 309-313. https://doi.org/10.1016/j.ijbiomac.2021.02.214

[34] I. Ettah and L. Ashton. *Engaging with Raman Spectroscopy to Investigate Antibody Aggregation.* Antibodies. 7 (2018), pp. 24. https://doi.org/10.3390/antib7030024

[35] B. Wei, N. Woon, L. Dai, R. Fish, M. Tai, W. Handagama, A. Yin, J. Sun, A. Maier, D. McDaniel, E. Kadaub, J. Yang, M. Saggu, A. Woys, O. Pester, D. Lambert, A. Pell, Z. Hao, G. Magill, J. Yim, J. Chan, L. Yang, F. 25 Macchi, C. Bell, G. Deperalta and Y. Chen. *Multi-attribute Raman spectroscopy (MARS) for monitoring product quality attributes in formulated monoclonal antibody therapeutics.* mAbs, 14 (2022). https://doi.org/10.1080/19420862.2021.2007564

[36] Thiagarajan, G., Widjaja, E., Heo, J. H., Cheung, J. K., Wabuyele, B., Mou, X., & Shameem, M. *Use of Raman and Raman optical activity for the structural characterization of a therapeutic monoclonal antibody formulation subjected to heat stress*. J. Raman Spectrosc., 46 (2015), pp. 531-536.

[37] A. Biancolillo and F. Marini. *Chemometric Methods for Spectroscopy-Based Pharmaceutical Analysis.* Front. Chem. 6 (2018). https://doi.org/10.3389/fchem.2018.00576

[38] Shlens J. *A tutorial on principal component analysis*. 2014. Available at: https://arxiv.org/pdf/1404.1100.pdf. Accessed November 12, 2020.

[39] Malani, H., Shrivastava, A., Nupur, N., Rathore, A.S. *LC-MS characterization and stability assessment elucidate correlation between charge variant composition and degradation of monoclonal antibody therapeutics.* AAPS J. 26 (2024), pp. 42. https://doi.org/10.1208/s12248-024-

00915-9

[40] Zheng, K., Ren, D., John Wang, Y., Lilyestrom, W., Scherer, T., Hong, Justin K.Y. and Ji, Junyan A. *Monoclonal antibody aggregation associated with free radical induced oxidation.* Int. J. Mol. Sci. 22 (2021), pp. 39-52. https://doi.org/10.3390/ijms22083952

[41] Hawe, A., Kasper, J.C., Friess, W., & Jiskoot, W. *Structural properties of monoclonal antibody aggregates induced by freeze-thawing and thermal stress.* Eur. J. Pharm. Sci. 10 (2009), pp. 79-87. https://doi.org/10.1016/j.ejps.2009.06.001

[42] Fincke, A., Winter, J., Bunte, T., & Olbrich, C. *Thermally induced degradation pathways of three different antibody-based drug development candidates.* Eur. J. Pharm. Sci. 62 (2014), PP. 148-160. https://doi.org/10.1016/j.ejps.2014.05.014

[43] Nowak, C., Cheung, J.K., Dellatore, S.M., Katiyar A., Bhat R., Sun, J., Ponniah, G., Neill, A., Mason, B., Beck, A., & Liu, H. *Forced degradation of recombinant monoclonal antibodies: A practical guide.* mAbs 9 (2017), pp. 1217-1230. https://doi.org/10.1080%2F19420862.2017.1368602

[44] C.L.M. Morais, K.M.G. Lima, M. Singh and F.L. Martin. *Tutorial: multivariate classification for vibrational spectroscopy in biological samples.* Nat. Protoc. 15 (2020), pp. 2143-2162. https://doi.org/10.1038/s41596-020-0322-8

[45] Savitzky, A. and Golay, M. J. *Smoothing and differentiation of data by simplified least squares procedures*. Anal. Chem. 36 (1964), pp. 1627–1639. https://doi.org/10.1021/ac60214a047

[46] A. Hauptmann, G. Hoelzl, M. Mueller, K. Bechtold-Peters and T. Lorting. *Raman marker bands for secondary structure changes of frozen therapeutic monoclonal antibody formulations during thawing.* Pharm. Biotecnol. 112 (2023), PP. 51-60. https://doi.org/10.1016/j.xphs.2022.10.015

[47] Zhang, C., Springall J.S., Wang X. and Barman I. *Rapid quantitative determination of aggregation and particle formation for antibody drug conjugate therapeutics with label-free Raman spectroscopy*. Anal. Chem. 1081 (2019), pp. 138-145. https://doi.org/10.1016/j.aca.2019.07.007

[48] L. Zhang, M. J. Henson and S.S. Sekulic, 2005. *Multivariate data analysis for Raman imaging of a model pharmaceutical tablet*. Anal. Chim. Acta. 545 (2005) pp. 262–278. https://doi.org/10.1016/j.aca.2005.04.080

Table S1: Composition of the two formulations analyzed (A and B) in this study.

Ingredient	***Formulation A***	***Formulation B***
Sucrose	*No*	*Yes*
L-methionine	*Yes*	*Yes*
Poloxamer 188	*No*	*Yes*
Meta cresol	*No*	*Yes*
Trehalose dihydrate	*Yes*	*No*
Polysorbate 20	*Yes*	*No*
L-Histidine	*Yes*	*No*
Sodium chloride	*Yes*	*No*
Disodium hydrogen phosphate di-hydrate	*No*	*Yes*
Sodium di-hydrogen phosphate mono-hydrate	*No*	*Yes*

Sample ID	Type of samples	Raman Shift [cm^{-1}]
1	Oxidative	
2	Thermal out-of-spec	
3	Thermal in-spec	Raman Intensity (100-3425 cm^{-1})
.	No stress	
.	Dilution	
.		
168		

(a)

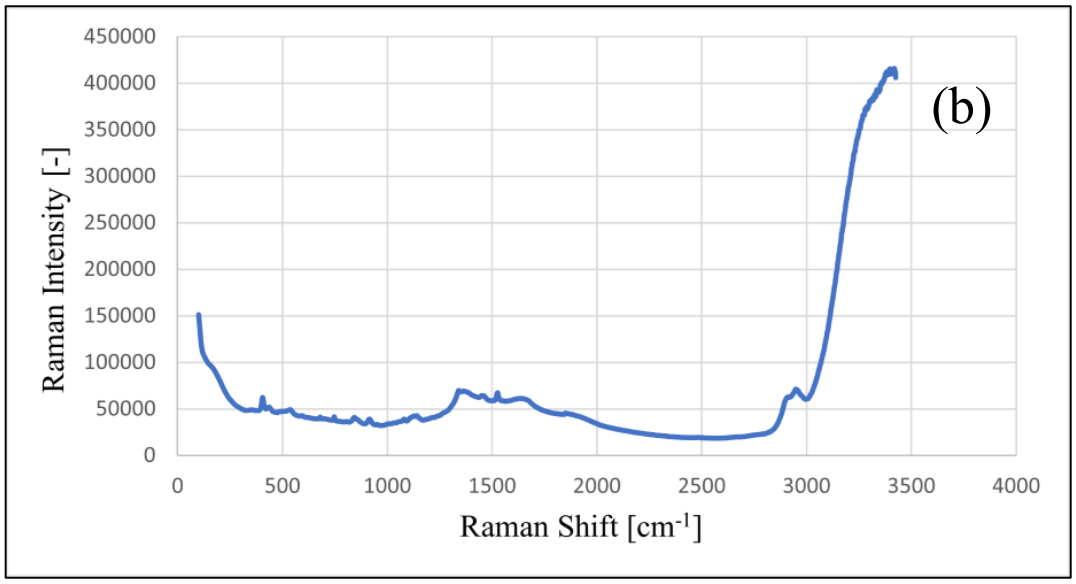
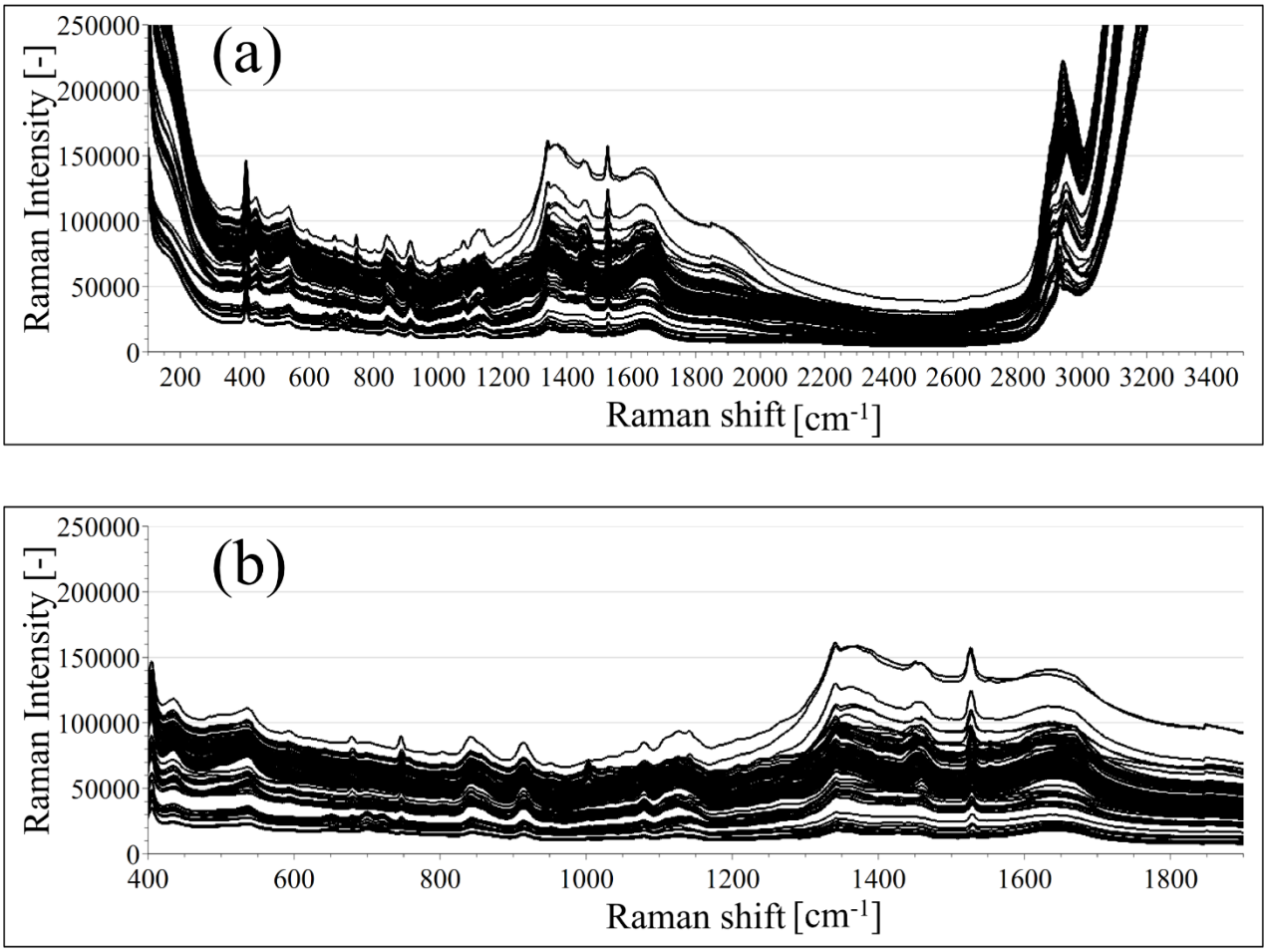


Figure S1: Dataset arrangement visualization: detail of matrix X structure. Each row in the matrix represents the spectrum of a specific sample, as shown in (b).

Figure S2: Raw Raman spectra of samples of formulation A (a) in the entire range of acquisition and (b) in the fingerprint region.



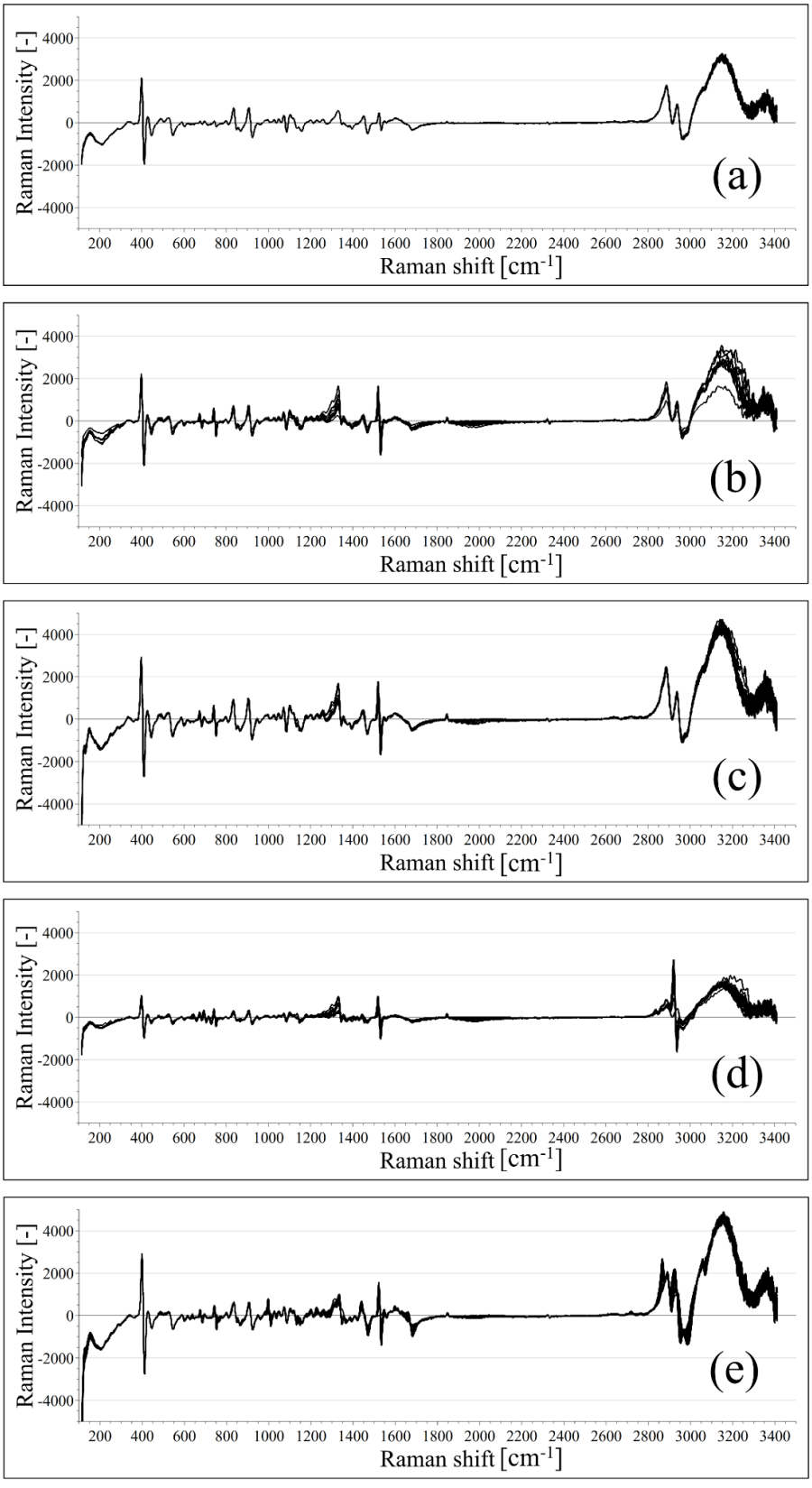


Figure S3: Preprocessed Raman spectra of samples of formulation A according to the type of samples: (a) no stress, (b) thermal in-spec, (c) thermal out-of-spec, (d) oxidative stress and (e) diluted samples in the entire range of acquisition.

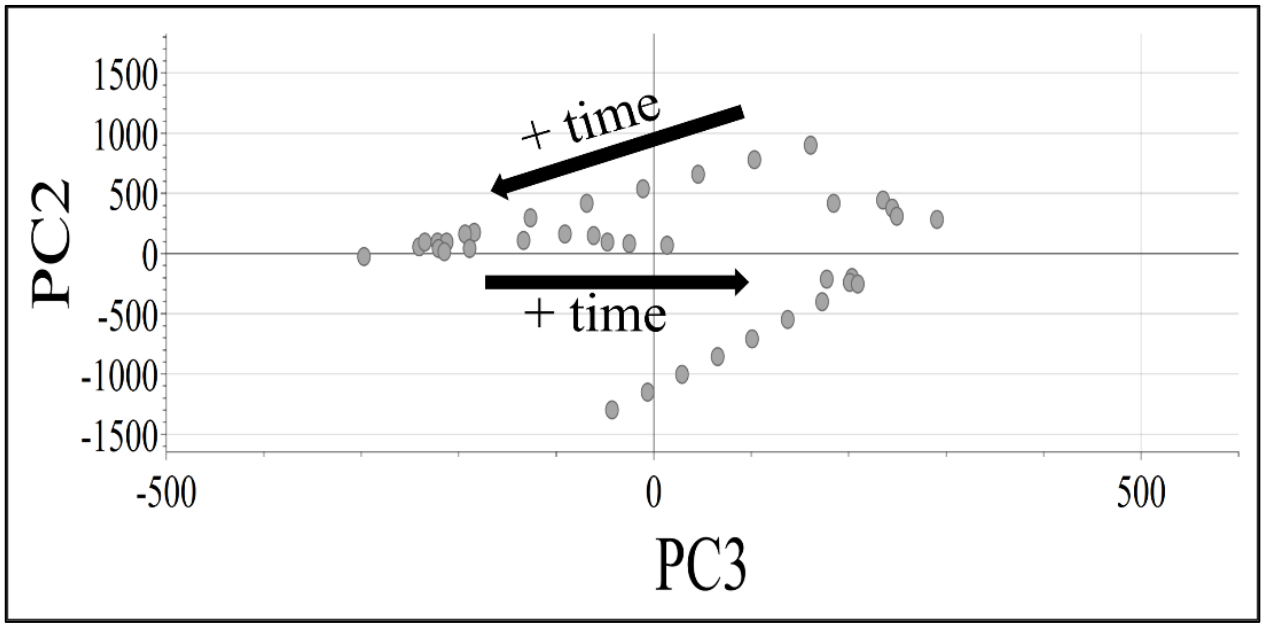


Figure S4: Scores plot of PCA model developed on the thermally stressed samples in the range 150-3425 cm-1. The scores are displayed in the plane described by the 2nd and 3rd PC. The arrows indicate the increasing exposure duration to the thermal stress.