Published version at: **DOI:** 10.1016/j.aca.2019.12.020

Multi-Attribute Quality Screening of Immunoglobulin G using polarized Excitation Emission Matrix Spectroscopy.

Ana Luiza de Faria e Silva, Saioa Elcoroaristizabal, and Alan G. Ryder.*

Nanoscale BioPhotonics Laboratory, School of Chemistry, National University of Ireland, Galway, Galway, H91 CF50, Ireland.

* To whom all correspondence should be addressed.

Tel: +353 91 49 2943 **Fax:** +353 91 49 4596 **Email:** alan.ryder@nuigalway.ie (A.G.R)

Citation: Multi-Attribute Quality Screening of Immunoglobulin G using polarized Excitation Emission Matrix Spectroscopy. A.-L. de Faria e Silva, S. Elcoroaristizabal, and A.G. Ryder. Analytica Chimica Acta, 1101, 99-110, (**2020**).
DOI: 10.1016/j.aca.2019.12.020

Abstract

Immunoglobulin G (IgG) is often used as a starting material for the production of functionalised antibodies, like Antibody Drug Conjugates (ADCs), PEGlyated-conjugates, or radioimmunoconjugates. The gross structural quality of the protein starting material is, therefore, an important factor in determining final product composition, purity, and quality. In terms of structural quality, one needs to know both the aggregation content and the tertiary structure of the protein. The measurement of structural quality in solution can thus be difficult, but the use of intrinsic fluorescence measurements might offer a solution because of its high sensitivity, ease of use, and when implemented in via multi-dimensional techniques like polarized Excitation Emission Matrix (pEEM) spectroscopy, its high information content.

Here we demonstrate how pEEM measurements can be used as a multi-attribute screening method for protein quality using a polyclonal rabbit immunoglobulin (rIgG) model system. By using both Rayleigh scatter and fluorescence emission in combination with simple chemometric data analysis methods like Principal Component analysis (PCA) and unfolded partial least squares (U-PLS) one can simultaneously measure protein concentration, structural variance, and particle/aggregate composition.

Furthermore, one can generate quantitative prediction models for non-reversible aggregation content as described by size exclusion chromatography (SEC) and obtain qualitative information about reversible aggregate content, which cannot be obtained from SEC measurements. In conclusion, the pEEM measurement approach is a potentially useful Process Analytical Technology (PAT) method for downstream processing operations in biopharmaceutical manufacturing.

Keywords: Protein, Immunoglobulin G, Fluorescence, Multidimensional, Polarization, Aggregation.
stdev, standard deviation;

1. Introduction

The chemical modification of monoclonal antibodies (mAb) or their fragments to produce new functional entities such as PEGylated proteins [1] or Anti-body Drug Conjugates (ADCs) [2] is becoming more widespread as the use of mAb based therapies continues to increase. Immunoglobulin G (IgG) is probably the most widely used mAb, and the Y-shaped protein is composed of two antigen binding sites (Fab) and one domain responsible for the effector function (Fc). The structure is stabilized by covalent (intra and interchain di-sulphide bridges) and non-covalent, hydrogen bonds, ionic bonds, van der Waals interactions [3]. The functional modification of proteins can have a variety of objectives including modulating pharmacokinetic properties by attachment of polyethylene glycol (PEG) [4], or providing new therapeutic modes of action such as ADCs, [5] and radioimmunoconjugates [6] for cancer therapy, or for diagnostics by the attachment of fluorophores [7].

Protein conjugates can be prepared using a variety of different chemical strategies such as the modification of lysine or disulphide bonds. However, irrespective of the attachment method, the quality of the starting and reagent materials will be critical to the quality of the final product. The small molecule raw materials (buffers, conjugates, linkers, solvents) can all be easily assessed using standard chemical analysis techniques like HPLC. However, the protein substrate, is an altogether more complex challenge because purity is not the only consideration, one much also assess the structural status of the protein in solution. The key considerations here are the tertiary and quaternary structures in solution, because these are influenced by a wide variety of factors (e.g. pH, temperature, ionic strength, mechanical forces, etc), and changes in these structures will influence the course of chemical reactions. Another area in which IgG quality (concentration and structural assessment) needs rapid assessment involves the rehydration of lyophilized proteins. Lyophilisation offers many benefits for the transport and storage of therapeutic

proteins [8]; however, the reconstitution step introduces a potential source of concentration and structural variability.

Therefore, for protein starting materials and reconstituted lyophilised proteins, there is a need to verify both the concentration and tertiary/quaternary structural integrity before use, as an element of the Quality by Design (QbD) principles being adopted by the (bio)pharmaceutical sector [9]. There are a variety of tools which can be used for concentration or structural analysis for Quality Assurance (QA) or Quality Control (QC) characterisation of proteins like IgG. Other quality attributes such as charge and glycosylation require the use of ion exchange chromatography and mass spectrometry [10] for example, but these are not considered here. UV-visible absorption spectroscopy is widely used for total protein concentration, size exclusion chromatography (SEC), and dynamic light scattering (DLS) can be used for quaternary (i.e. aggregation) analysis, and field-flow fractionation (FFF) gel electrophoresis (e.g. Gel Electrophoresis, SDS-PAGE) are generally used to assess purity, integrity, and aggregation content [11]. Assessing conformational changes can be achieved by spectroscopic methods such as Circular Dichroism, Fourier transform infrared (FTIR), and by the use of intrinsic fluorescence measurements. The pros and cons of each technique are well known, but issues like single attribute measurements (UV-vis), relatively time-consuming sample preparation/analysis (SDS-page, SEC), modifying of environment (SEC, DLS, SDS-page), can all cause issues with allowing for rapid online measurements. This is particularly significant for QA measurements in the working solutions of proteins, as these may have incompatible buffers and concentrations for some techniques. Finally, no single technique can provide all the information required to properly assess protein quality in terms of tertiary/quaternary structure variance, and thus a combination of methods is usually recommended. This, however, adds to the analysis time, and may be impractical in some situations.

Using intrinsic fluorescence spectroscopy (IFS) of proteins has potential because it is sensitive (down to nanomolar) can be implemented in a wide variety of buffers, and at varying concentrations thus avoiding issues with upsetting quaternary structural equilibria. However, IFS is normally implemented using single excitation wavelengths and has thus inherently low information content. The information content per measurement can be dramatically increased via the use of Excitation Emission Matrix (EEM) measurements. An EEM better explores the protein emission space providing a molecular fingerprint of the various fluorophores present in the mixture. The shape of the EEM plot is representative of the fluorophores present and their photophysical interactions via Foster resonance Energy transfer (FRET) because of their close proximity in a protein. For rIgG there are 50 tyrosine (Tyr) and 24 tryptophan (Trp) residues contributing to the overall intrinsic fluorescence. The combination of these effects with the environmental-sensitive fluorescence emission, are responsible for the shape observed in the matrix providing a very diagnostic tool for assessing protein changes and stability. EEM can be implemented

robustly for routine analytical applications [12], is widely used in water analysis [13] and for cell culture media analysis [14-16]. EEM and PARAFAC analysis has also been used to study structure changes in BSA [17]. For proteins, where more information is needed one can use polarized EEM (pEEM) measurements and chemometric data analysis. When factor-based analysis can be implemented, this anisotropy resolved multidimensional emission spectroscopy [18, 19] can be used to explore individual fluorophore emission contributions. .

For IgG type proteins however, PARAFAC analysis of either pEEM [20] or polarized TSFS data [21] is rather complicated and does not, as yet offer a robust approach to routine protein variance analysis. The key reasons are that the large numbers of fluorophores in IgG result in extensive FRET and the presence of residual light scatter both hinder component resolution.

The chemical modification of mAb's or their fragments to produce new functional entities will be sensitive to a variety of factors related to the protein substrate. Three of the critical variables that need to be measured, and thus controlled are the real concentration of native (active) rIgG forms, protein conformation (tertiary), and the aggregation profile (quaternary). As these factors can be affected by a wide variety of processes including storage conditions and times, one needs a quick, non-destructive, and effective method for assessing protein variability in solution. Here we used a polyclonal rabbit Immunoglobulin (rIgG) model, due to its relative low cost, and because it is widely used for a variety of applications [22-24]. The rIgG pAb offers a challenge as it is comprised of a mixture of closely related structures, with different epitope recognition abilities, produced by a large number of B cell clones [22]. One might then expect that the rIgG should show significant batch-to-batch variability because they have been produced in different animals at different times [22, 25].

Here we wanted to assess the efficacy of pEEM measurements for rapid lot-to-lot quality assessment of protein. In this case we used rIgG which was being used as a model starting material for conjugation reactions. This is also a convenient initial model for downstream product characterisation in biopharmaceutical manufacturing as the sample set comprised of three types of samples: storage compromised, mechanically stressed, and unstressed samples of rIgG. pEEM should be able to provide information about concentration (to replace absorbance measurements), tertiary structure variance, and quaternary structure/aggregation (to replace and/or complement SEC) in a single measurement. To achieve this, we use both the fluorescence and Rayleigh scatter components of pEEM measurements, in combination with simple data analysis procedures.

2. Materials and methods

2.1 Materials: IgG from rabbit serum (≥95%, essentially salt-free, lyophilized powder), L-Tryptophan (Trp), ethylenediaminetetraacetic disodium (EDTA), sodium phosphate monobasic, sodium phosphate dibasic hepta-hydrate, and sodium chloride were purchased from Sigma-Aldrich and used as received. High purity water, HPW, (Honeywell, chromatography grade) was used for preparing a 0.01 M Phosphate buffer with 0.150 M saline and 10mM EDTA (pH 7.0±0.1) in HPW which were membrane filtered (0.2 μm) using PES Captiva Premium Syringe filter purchased from Agilent. Four different rIgG lots (Table S1, SI) from Sigma were used: SLBP7449V (Lot 1), SLBW8687 (Lot 2), SLBM2617V (Lot 3) and SLBZ5214 (Lot4). Unstressed samples were prepared using Lot 1 (n=15), 2 (n=5), and 4 (n=6) in PBS/EDTA buffer pH 7.0 (approximately 1 g/L). The different stock solutions for each lot were prepared on different days using the same rIgG bulk lots but different source vials and were then filtered before being split into aliquots for analysis.

A set of stressed rIgG samples was prepared to simulate aggregate containing solutions that could be formed during manufacturing, or improper storage, shipping, and handling. For this, Lot2 aliquots were transferred to Lobind Eppendorf tubes and mechanically stressed using a vortex shaker under different conditions. For Lot2, soln.6 the shaking was done with 3mL of solution in 5mL tubes whereas for soln.7 the sample volume was 2mL in a 5mL tube, producing a bigger headspace (and thus increased the chance of aggregation at the air-water interface [26]). Lot3 samples were prepared from a batch of the lyophilized powder which had been improperly stored. This has been inadvertently stored at room temperature for one month, opened, and then stored at 2–8°C for 18 months prior to use.

An L-Trp solution with an equivalent concentration to 1mg/mL IgG was prepared in PBS/EDTA buffer pH=7.0 and used here as a negative control. Protein concentration was calculated from absorbance measurements at 280 nm after correction for scatter contribution (Table S1, SI). Samples were transferred to cuvettes for spectroscopic analysis and afterwards they were recovered into Eppendorf tubes and frozen at –70°C for later SEC analysis.

2.2 Instrumentation and data collection: All spectra were collected from solutions held in 1×1 cm pathlength quartz cuvettes (Lightpath Optical, UK) at 25°C. UV-visible spectra (200–600 nm, 2 nm resolution) were measured using a Cary 60 spectrometer (Agilent) using the corresponding PBS solution as a reference. EEM measurements were made using a Cary Eclipse fluorescence spectrophotometer (Agilent Technologies) fitted with a bespoke polarizer accessory [27] and a temperature controlled multi-cell holder. EEM spectra were collected over an excitation/ emission ($\lambda_{ex}/\lambda_{em}$) range of 240–320/260–450 nm with 2 nm step increments and 10 nm excitation/emission slit widths, scan rate of 1200

nm min^{-1}, and with a 600V photomultiplier voltage. All samples were measured using four different polarization configurations: HH (horizontal-horizontal), HV (horizontal-vertical), VH (vertical-horizontal), and VV (vertical-vertical). Spectra were collected in S/R mode but uncorrected for instrument response.

SEC was undertaken using an Agilent 1260 HPLC system equipped with a diode array detector (DAD). The column was a 300×7.8 mm mAb PAC-SEC1 column (ThermoFisher) with 5 µm particle size. Protein samples were stored at –80°C and thawed for 2 hours at 2–8°C before use, then filtered, and injected (10 µL) at 30°C with 50mM Sodium Phosphate + 300 mM NaCl buffer, pH 6.8, as the mobile phase, and a 0.76 mL/min flow rate. SEC analysis was carried out in triplicate, on each stock solution prepared, however, for Lot1, it was only possible to measure each stock solution once (n=5).

2.3 Data analysis and Chemometric tools: SEC data analysis was carried out using Agilent OpenLab CDS Version 2.3. (ChemStation integrator). The resolution factor (R) was used to evaluate the accuracy of chromatographic separation with R> 1.5 indicating a good separation between adjacent peaks:

$$R=2\frac{tR2-tR1}{W1+W2} \qquad \text{Eq.2}$$

2.4 Simple Aggregation Assays: Two univariate spectroscopic indexes were used as qualitative indicators of protein aggregation. The first was the UV-visible aggregation index (UV-AI) and was calculated from absorbance measurements at 280 and 350 nm [28, 29] (see SI for details). The second measure of aggregation (Fl-AI) was similar to that reported by Nominé *et al*. [30] and was the ratio of the scatter at 280 nm to the fluorescence intensity at 340 nm (SI) calculated after blank subtraction. An aqueous Trp solution was used as a negative control. Three different values were calculated using the parallel polarized (Fl-$AI_{\parallel}$), perpendicular polarized (Fl-$AI_{\perp}$), and the total, or unpolarized (Fl-AI_T) emission spectra extracted from the EEM measurements. The Fl-$AI_{\parallel}$ value was used throughout the manuscript as this was the most sensitive to particle content.

2.5 Polarized EEM analysis: EEM_{VH} data was G factor ($G=EEM_{HV}/EEM_{HH}$) corrected to give the perpendicular polarised data, $EEM_{\perp}$ ($=EEM_{VH}\times G$). The VV emission spectra are the parallel polarization ($EEM_{\parallel}$) data. Unpolarized EEM (EEM_T) datasets were generated from the $EEM_{\parallel}$ and $EEM_{\perp}$ measurements using the following equation: $EEM_T=EEM_{\parallel}+2\times EEM_{\perp}$. Raman scatter was removed by using a buffer blank subtraction [31] while the Rayleigh scatter (RS) of each sample was estimated as a separate bilinear component and modelled by PARAFAC, resulting in a structure that was then reshaped and subtracted from the original matrices [32, 33]. After RS removal, the area was replaced with missing values (Figure S2).

The Rayleigh scatter in EEM is normally removed for chemometric analysis, however, this means that potentially useful information concerning particle size and distribution is being discarded [34]. Here, scatter information (as a Rayleigh spectrum, 240–320 nm) extracted by PARAFAC was evaluated separately as a qualitative measurement for identifying changes in protein aggregation, and in particular the formation of reversible, soluble aggregates.

2.6 Multivariate data analysis: Chemometric analysis were performed using PLS_Toolbox8.2.1® working in MATLAB (ver. 9.1.0) environment, and an in-house written program (FluorS). EEM data and RS‖ spectra were smoothed (Savitzky–Golay filter), normalized, and robust centered prior to using robust principal component analysis (ROBPCA). For the RS ‖ spectra normalization was by the total area. ROBPCA is a variation of the classical PCA approach which is less influenced by the presence of outliers and is thus a more suitable method for screening applications [35]. Apart from sample discrimination, ROBPCA was used to identify the significant spectral features which differentiated samples in the scores plots. Quantitative modelling for concentration quantification was implemented using unfolded partial least squares (U-PLS)[35]. U-PLS regression was implemented for quantification of concentration and % of aggregates and the model complexity (number of latent variables) was determined using standard cross-validation methods [36]. For aggregation content U-PLS models, data were pre-processed by blank subtraction, missing data (missing area replaced with zeros), normalized by the maximum peak, and unfolded. Model quality method was accessed using the root mean square error (RMSE) and R^2 parameters.

3. Results & Discussion:

The key goal here to determine whether pEEM measurements could provide a more accurate, simpler, and faster assessment of protein variance than conventional SEC or simple spectroscopic measurements. The first step was to assess the raw data generated by the various techniques and then look at the discrimination capability of each method.

3.1. Comparison of methods:

We first assess the ability of each technique to detect variation between rIgG samples that were known to be different because of stress, different source lots, etc (total sample set, n=34). SEC chromatograms (Figure 1A) showed three peaks with retention times (minutes) of 9.28 (oligomers, peak 1), 10.05 (dimers/trimers, peak 2), and 11.36 (monomer, peak 3). The separation of polyclonal rIgG with its variable composition leads to relatively broad peaks, however, resolution values (R) indicated a good separation between dimer/trimer and monomer peaks (R>2) despite the poor resolution between

dimer/trimer and oligomer peaks (R<1.5). Here we used the total area (rt = 8-14 min), and the areas under peak 3 (monomer) and combined area of peaks 1 and 2 (aggregates), to calculate the respective percentages of monomers and aggregates, which we refer to as the aggregation profile. SEC of rIgG samples (Figure 1A) measured in triplicate were consistent (indicating good reproducibility between injections, Table 1) and overlaid almost perfectly once normalized (normalized data not shown). The small variances in the raw data can be attributed to small concentration variations between injections, stock solutions, and the buffer components.

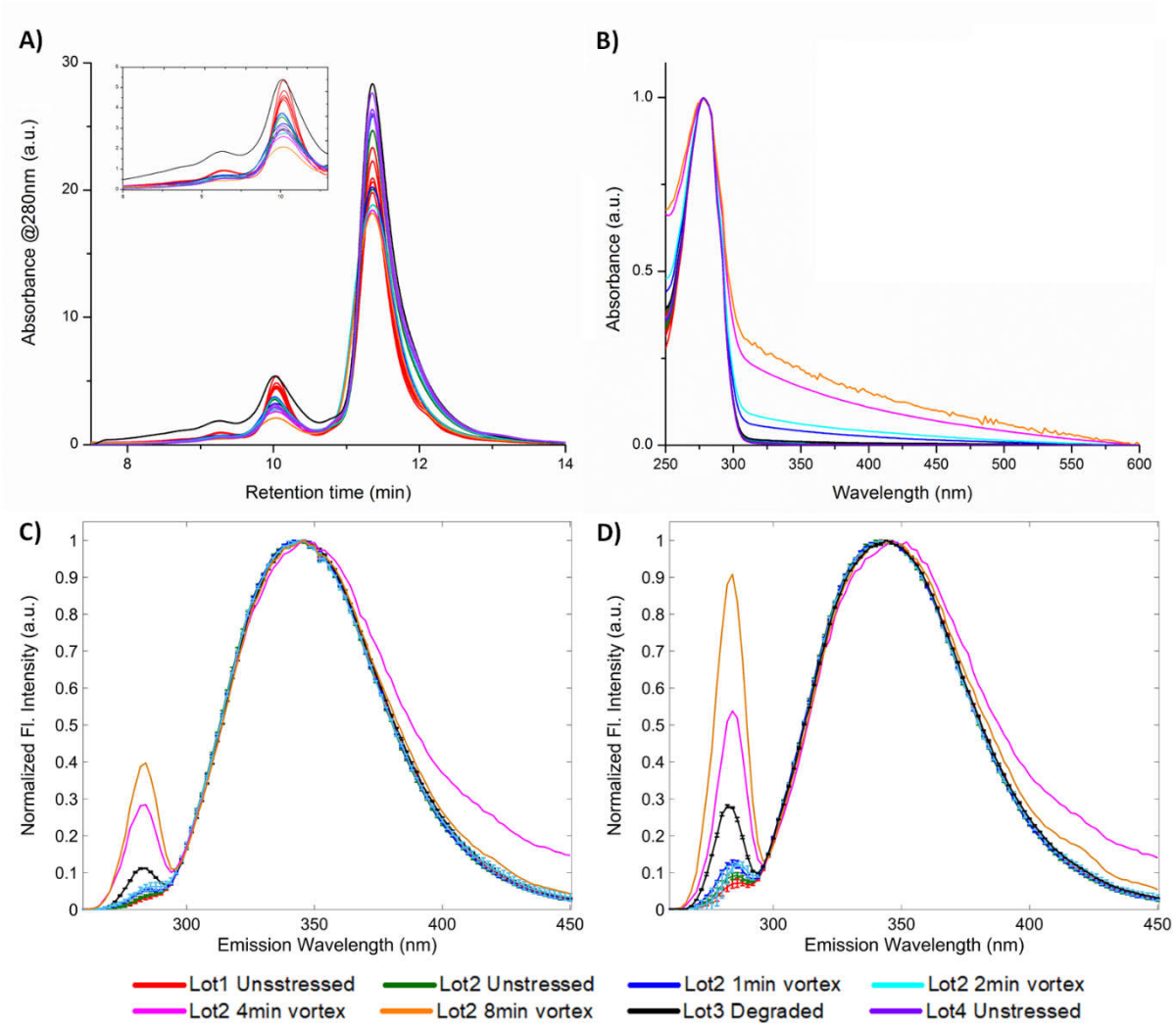

Figure 1: (A) Overlaid size exclusion chromatograms showing the concentration and % of oligomers as the main difference between samples. Labels 1-3 correspond to higher order aggregates, dimers/trimers and monomer, respectively.; (B) Normalized UV-Visible spectra show significant differences in light scattering for stressed samples; Average (± std) 2D-fluorescence spectra of rIgG measured at λ_{ex}=280 nm of normalized unpolarised (C) and parallel polarized (D) spectra.

Absorbance spectroscopy (Figure 1B) is an easy, fast and inexpensive method generally used for protein quantification. The rIgG absorption maxima is at 280 nm due largely to the aromatic amino acids. An

increase in apparent absorbance at longer wavelengths (>300nm) was also observed for some samples which is related to light scatter (vide infra). To increase sensitivity, the obvious approach is to use fluorescence spectroscopy since protein emission is very sensitive to structural and aggregation changes [37]. The simplest way is to use a single excitation wavelength, typically 280 nm. Figure 1 shows the normalized unpolarized and parallel polarized emission spectra (λ_{ex}= 280 nm) with the significant increase in Rayleigh band intensity in the parallel polarization mode.

Despite the higher sensitivity and specificity of intrinsic fluorescence, single wavelength measurements (i.e. 2-D spectra) were not ideal for analysing complex proteins with multiple fluorophores and the EEM measurements were better because they provide more information for identifying subtle structural changes [38]. rIgG had strong fluorescence emission over the $\lambda_{ex}/\lambda_{em}$ range: 270-300 / 300-380 nm, with the fluorescence maxima for all samples at $\lambda_{ex}/\lambda_{em}$ =292 / 344 nm, which is mostly due to Trp emission. To assess and quantify the degree of rIgG spectral variation we calculated the standard deviation at every point in the pEEM map for the different polarizations (***Figure 2***). This shows the spectral regions where the most variation occurs and when taking all 34 samples (***Figure 2***B/D/F) the variability was highest for the directly excited Trp emission and also in the longer wavelength ($\lambda_{ex}/\lambda_{em}$>300/400nm) region associated with "deep blue autofluorescence" [39, 40]. This also showed that $EEM_{\parallel}$ measurements were the most sensitive to sample variance and thus this data used for all further sample analysis (*vide infra*).

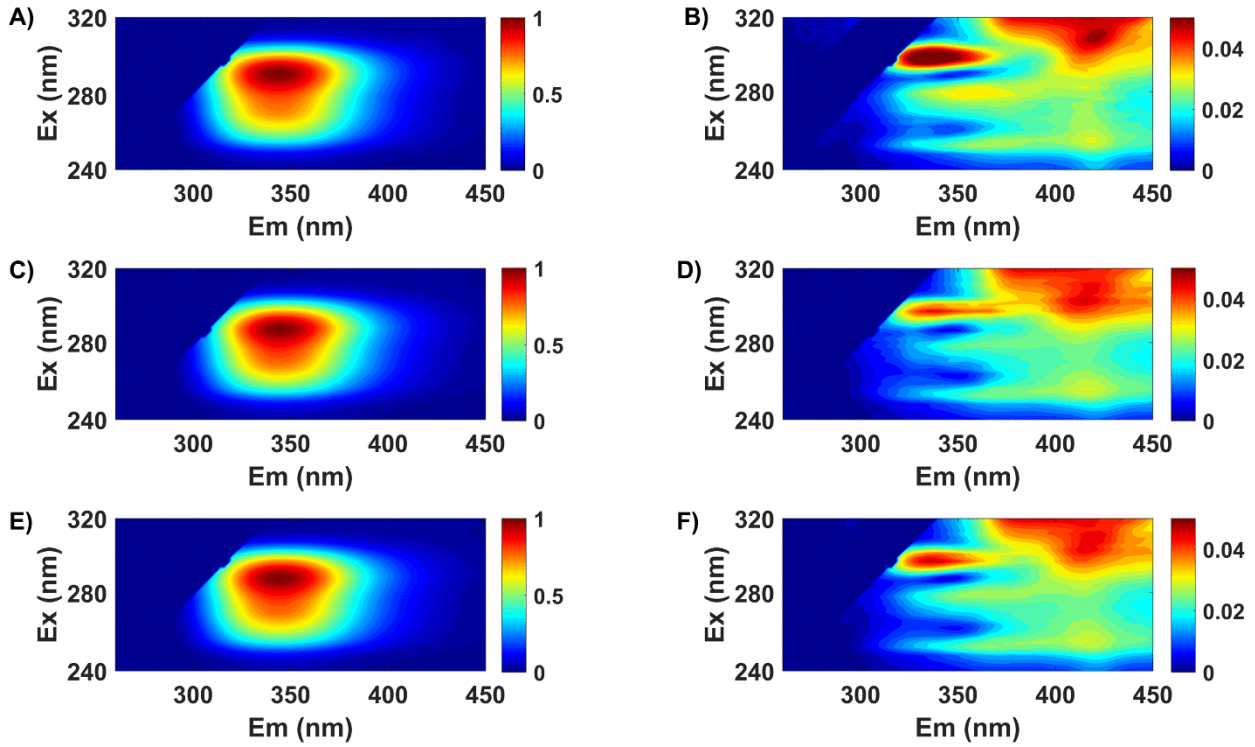

Figure 2: *Mean (left) and stdev (right) plots calculated using* $EEM_{\parallel}$ spectra *(A/B),* $EEM_{\perp}$ *(C/D), and* EEM_{T} *(E/F) data from all rIgG samples (n=34).*

The long wavelength emission plays an import role in discriminating lots with varying aggregation content (*vide infra*). It has been reported in other studies that this deep blue autofluorescence has been associated with protein aggregation and fibrillation, however, others indicate that this can be induced by oxidation of Trp or Tyr, or is due to carbonyl emission [39, 40]. While, the source of this emission is a source of controversy, it is nonetheless a real, but weak signal. A control study (Figures S5/S6, SI) using a Trp solution confirmed that the variance in this long wavelength region was not due to instrumental or measurement factors. It should be noted this emission is very weak, typically 2-3% of the main band for unstressed samples, and also very sensitive to Inner Filter Effects (IFE) [41]. However, for the stressed samples this long wavelength signal increased significantly (Figure S4, SI) to more than 20% for the most mechanically stressed samples. If one needs to rely only on this spectral region, then data collection needs to be optimised to increase signal to noise ratio (SNR).

Another useful source of information from EEM measurements is the Rayleigh scattered light which is normally discarded during fluorescence analysis [42]. However, this can be very useful in providing information about changes in particle concentration and size distribution [34]. There is a complex relationship between Rayleigh (and Mie) scattered light intensity and particle size, concentration and size distribution, which is also affected by refractive index, solution properties, and excitation wavelength. This complex light-material interaction is the basis of DLS [43] and other quantitative

particle sizing technologies [44]. The Rayleigh signal was, as expected, strongest in EEM$_{\parallel}$ because of the polarized nature of scattered light and this was used for analysis. To ensure measurement robustness and because we are only concerned with qualitative analysis, we first used the ratio between Rayleigh scatter and fluorescence intensity maxima (I_{RS}/I_{Fmax}) to identify HMWS formation in solution. We considered two excitation wavelengths 296 nm and 316 nm, the variation in I_{RS}/I_{Fmax} ratio across the full sample set (**Table 1**) was larger at 296 nm (RSD 73.86%) than 316 nm (51.93%), whereas all the unstressed samples the values were 6.53 and 9.18% respectively. This we suggest was due to a combination of factors including changing particle size distribution (increased RSD) and resonant light scatter effects (difference between 296 and 316 nm). The RS$_{\parallel}$ spectrum extracted from the EEM data provides a better measurement as it captures more of the particle/size distribution induced changes than a single wavelength.

3.2. Physical Homogeneity:

Physical homogeneity for a single protein containing sample type (the case here) has to consider two aspects both size and the fold state (i.e. quaternary and tertiary structure). This also should encompass information about reversible and non-reversible aggregates in the sample solution. Some of the critical factors are measuring the lot-to-lot and aliquot-to-aliquot (reproducibility) variance.

3.2.1 Aggregation profile: SEC is the compendial method for the characterization of aggregated, soluble, non-reversible, HMWS in mAb's [45, 46]. SEC analysis was the reference method for measuring the degree of aggregation of rIgG from three different bulk lots (Lot1,2,4, n=26). The % of aggregates area calculated from the chromatograms showed a different aggregation profile for each lot (18.2±0.9, 12.90±0.6, and 11.2±0.3%) while the storage degraded samples (Soln.8) had the highest concentration of soluble dimers/trimers and higher order aggregates (22.5%). For the mechanically stressed samples (Soln.7) there was an increase in SEC detectable aggregates for the longer (>2 minutes) stress times. However, this was also accompanied by very significant protein loss (by A_{280}) of 8, 10, and 14% for 2, 4, and 8 minutes of mechanical stress respectively, presumably due to either insoluble aggregate formation or adherence to the container walls. Smaller changes were noticed when Lot2 samples (soln.6) were stressed using smaller vial headspace, which is known to impact on the degree of structure disruption [26].

What was very interesting is that data from the simple spectroscopic methods did not correlate with the SEC results (**Table 1** and Figure S11, SI). The UV-AI data was ambiguous showing a value of 0.47 ± 0.22 % (unstressed Lots 1, 2, and 4, n=26), which suggested a very low aggregate content percentage of

aggregates in each sample (<1%). The RSD across these 26 samples was higher for UV-AI[1] (> 45%) than for the SEC[2] measurements (22%). FL-AI values suggested a lower aggregation content; however, these did not correlate with the SEC measurements despite having a similar RSD (21%) across the sample set. One problem with UV-AI was that it is based on an absorbance measurement at 350 nm which is typically a very small value for non-, or mildly stressed samples and this may bias the results. For instance, the absorbance values obtained for the fifteen Lot1 IgG samples at 350 nm was 0.014 ± 0.030 a.u., which has a high standard deviation and mean value lower than the calculated limit of quantification (LOQ) of 0.020 a.u. (LOQ=10 × stdev calculated using 10 blank solutions). In conclusion, this method is not suitable for QA/QC screening of protein lots because it does not correlate with SEC measurements and it is only sensitive to large changes (i.e. detecting the presence of large particles with hydrodynamic radii > 200 nm) in the protein solutions and cannot identify changes in the soluble aggregates [47]. These simple spectroscopic methods are obviously measuring a different population of aggregates compared to the reference SEC method.

The next step was to try and extract more information from the absorption spectra by using ROBPCA to better examine the spectral variance and see if a correlation could be established with SEC measurements. However, since there was no significant change in spectral profiles, apart from the rising background due to scatter, ROBPCA (Figure 3A) could not discriminate the samples according to their SEC determined aggregate content. The situation with $EEM_{\parallel}$ data was very different and the PC1-PC2 scores plot from ROBPCA modelling of the fluorescence signal showed clear sample clustering according to SEC determined aggregation content. The loadings plots (Figure 3C and D) show the most influential spectral features which contributed to discrimination. For PC1, the loadings indicate that the positive contribution is related to phenylalanine absorption (λ_{max} ~260 nm) and tyrosine emission (λ_{max} ~320 nm) while the negative part of the loadings plots corresponds most strongly to Trp and the longer wavelength emission. This indicates that the spectral change which correlates with aggregation profile is FRET related. ROBPC2, on the other hand, is more likely to be related to sample preparation issues and small changes in concentration and this will be discussed in the next section.

[1] All measured samples from Lot1, 2, and 4 were used for AI RSD calculation.
[2] All measured samples from Lot1, 2, and 4 (including measurement replicates) were used for SEC RSD calculation.

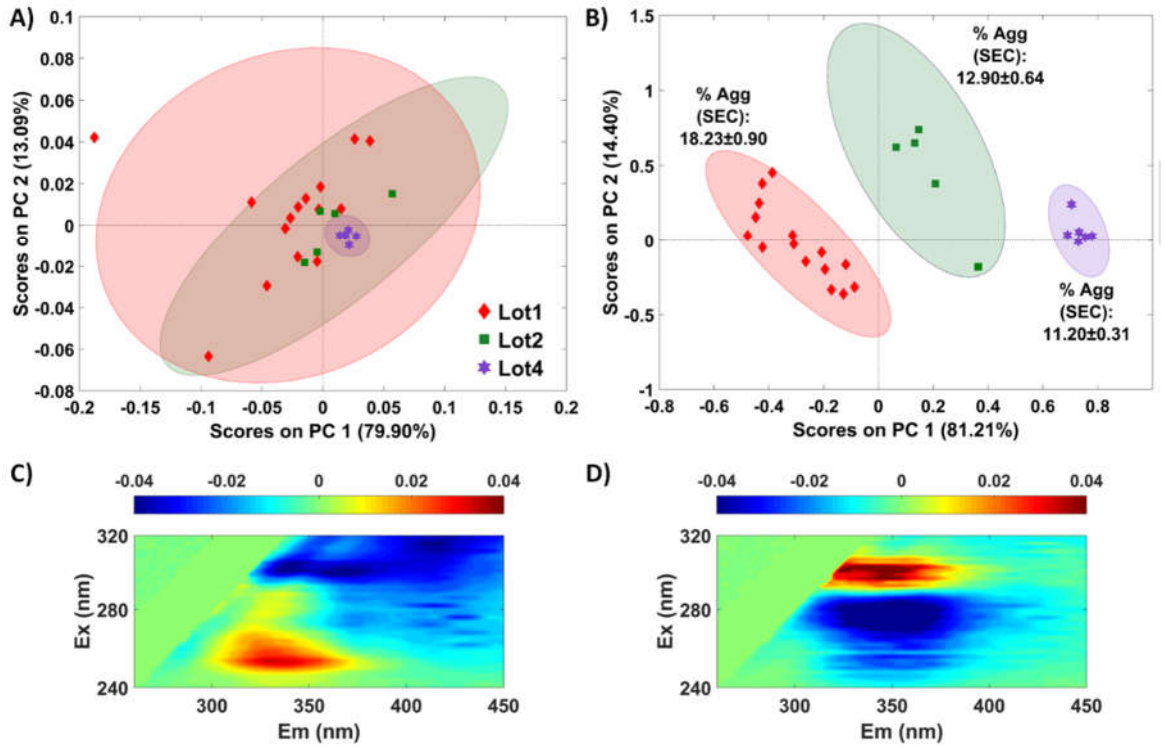

Figure 3: *Scores plots for lot-to-lot rIgG discrimination analysis by: (A) UV-visible spectroscopy, and (B) by EEM‖ measurements. Ellipses are the 95% confidence interval for each cluster. ROBPC1 (C) and ROBPC2 (D) loadings plots showing regions of most significant spectral variance.*

Based on the ROBPCA findings we built quantitative U-PLS models for predicting SEC aggregation content from EEM‖ measurements (Figure S7/S8). These models suggest that a quantitative correlation can be obtained. There is a caveat in that the spread of SEC aggregate data is small (10-24%) with only four clusters of values. It is useful to note that the model cross validation improves significantly when the 2-8 minute mechanically stressed samples were omitted which may be evidence that a different type of aggregate being formed. Mechanically stressed IgG is thought to form aggregates by displacement from the air-water interface [48] which is a different mechanism from that due to improper storage or potentially the manufacturing process. Further investigations are required to validate and ideally this should be done using a more representative industrial protein sample set where production and orthogonal analytical data is available (unfortunately we do not have access to such a sample set). However, these proof of concept models do demonstrate that it could be feasible to predict the SEC derived aggregation profile from EEM‖ measurements.

3.2.2 Reproducibility (Intra lot variability): Sample preparation is critical in protein analysis, and the complex structure and behaviour of these macromolecules does cause some issues. Here the changes

are expected to be more subtle than for aggregation, and mostly related to concentration and conformational changes caused by uncontrolled variations in sample preparation. To assess this source of variability we examined the Lot 1 (stock solutions 1-5) samples in detail as their aggregation profile was very consistent with a monomer to aggregate ratio of 4.52 ± 0.28 and low variability, RSD < 2%, (Table 1). The UV-AI and Fl-AI$_{\|}$ measurements on the other hand showed significant variation (RSD= 53% and 24% respectively) for the Lot 1 samples which indicated that the solution tertiary composition was different to that of the SEC analysis (we are confident that the measurement variability was low for these methods). ROBPCA of EEM$_{\|}$ required two components to adequately describe sample variation and the scores plot (Figure 4A) shows two significant issues.

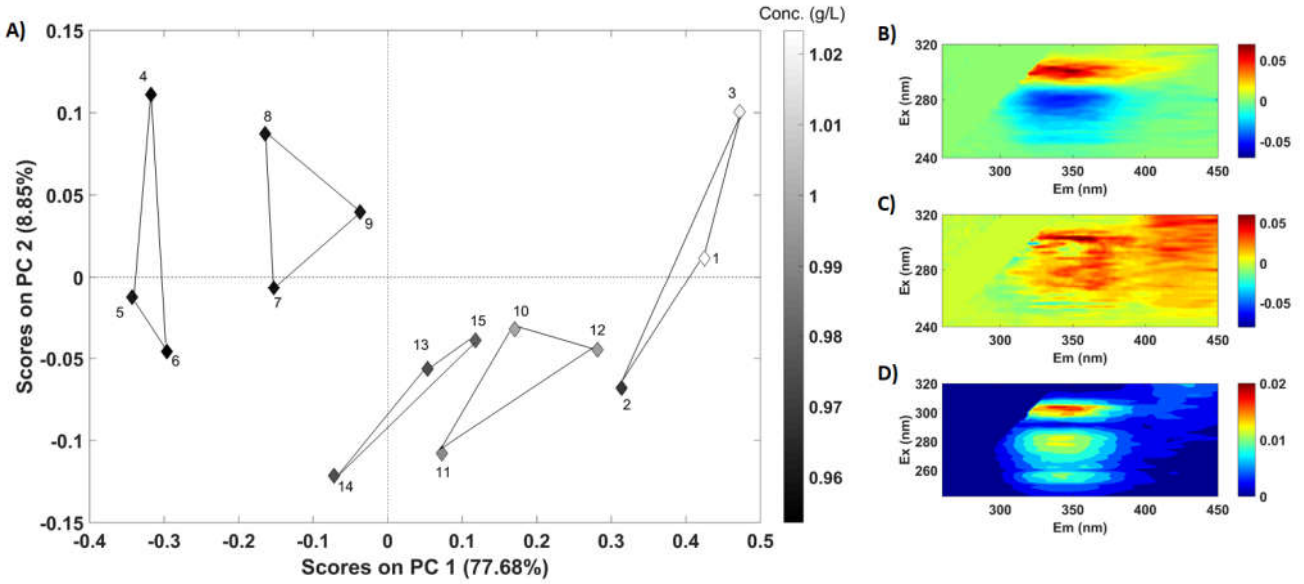

Figure 4: ***(A)*** *ROBPC1 versus ROBPC2 scores plot with samples numbers as per Table 1, and the points connected by lines are aliquots from the same stock solution. B and C are the refolded loadings plot (ROBPC1 and 2 respectively) showing the major areas of spectral variability in the rIgG samples from the same bulk lot. Plot (D) is the standard deviation plot calculated for all Lot1 samples. Analysis based on EEM$_{\|}$ data.*

The RobustPC1 and 2 loadings (Figure 4B/C), show the spectral areas of most significance. ROBPC1 shows changes in Tyr and Trp excitation and emission whereas ROBPC2 covers a much wider spectral range. The separation along PC1 correlates with the stock solution concentrations (**Table 1**) whereas the second PC does not correlate with any other measurements and thus represents some form of sample handling error that we have not yet been able to identify. This variation while significant in the scores plot does only represent a relatively small fraction (<9%) of the spectral variance because the spectral standard deviation (**Figure 4**D) calculated for these samples was only twice that of the experimental variability (Figure S5 and S6). We cannot yet fully exclude the possibility that the PC2 variance might be attributable to reversible oligomer formation in solution due to different buffer incubation times,

however, we do not have measurement methods of sufficient sensitivity to accurately measure these species in-situ. Thus, this plot graphically represents the minimum measurement error achievable with the pEEM measurement method for this sample type.

3.3 Protein Concentration:

The determination of concentration is another IgG critical quality attribute and the use of absorbance at 280 nm (A_{280}) is the simplest and commonest, non-destructive method [49]. Colorimetric methods such as Bradford assay are alternatives but do require the addition of reagents to the sample and thus are destructive to a certain degree. The use of IPF for *in-situ* protein quantification can be a challenge because of the potentially high working concentrations (>1 g/L) and thus significant IFE effects which might limit a simple intensity based quantification assay [50]. This is because there will be a non-linear dependence of fluorescence intensity to protein concentration as the chemical composition of the reaction solution changes [37]. Here, however, one can use normalized pEEM spectra with U-PLS regression to estimate protein concentration only by considering the changes in EEM profile because of the impact of changing IFE as concentration varied. For this, samples were split into calibration and test sets using the Kennard-stone algorithm [51] and the concentrations calculated from UV-Vis measurements were used as the nominal values. The U-PLS model required four LVs and resulted in a small $RMSE_{CV}$ (0.02g/L) which was 2% of the average rIgG concentration and equivalent to the error in the nominal concentration values. The model produced good results for both calibration and prediction sets (R^2_{Cal} =0.94 and R^2_{Pred}=0.96 and small errors (RMSE of calibration and prediction were 0.01g/L). The Elliptical Joint Confidence Region (EJCR) test was used to evaluate the accuracy of the predictions, this showed (Figure S13, SI) that the ideal point (1,0) for the slope and intercept was within the ellipse, indicating the absence of bias within the 95 % confidence level [52, 53]. The ellipse was narrow indicating good precision, but it was not centred about the ideal point because of the limited concentration range used. The loading plots (**Figure 5**B/C) show the regions of most variance, with LV1 representing Tyr excitation but Trp emission (i.e. FRET) and LV2 direct excitation of Trp.

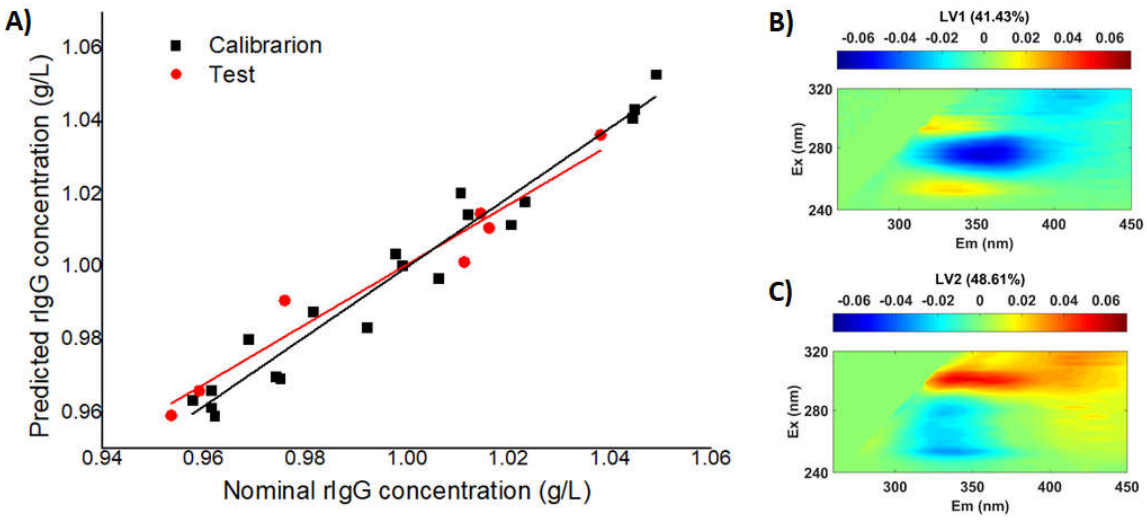

Figure 5- *Results of the U-PLS regression of EEM used for quantitative analysis: (A) Predicted versus nominal concentration for calibration and test set. Contour plots of latent variable loadings: (B) LV1 and(C) LV2. The nominal rIgG concentration was calculated from absorbance spectra.*

3.4 Sample screening:

The goal here was to determine which of the spectroscopic measurement methods provided the quickest and most effective screening method, based on ROBPCA analysis, for assessing protein variability which could be used for low-cost, routine QC applications, instead of SEC. To compare the various methods, we used the complete sample set which included 26 unstressed, two storage stressed, and six mechanically stressed samples (**Table 1**). An effective method for visualising the screening tool performance for the different spectroscopic methods is using the Hotelling T^2 vs. Q residuals plots (***Figure 6***). In these plots the upper right quadrant, should contain the outliers of most significance in terms of compositional differences [14].

ROBPCA of the absorbance spectra, (Figure S9A, SI), required four principal components to explain 93.06% of the total variance and easily differentiated some mechanically stressed samples as outliers in Figure 6A. This was mostly due to increased light scatter at longer wavelengths caused by the progressive increases in particle size /insoluble aggregates and, to a loss of light intensity in the 250-270 nm region due to the increased light scatter (Figure S9B-C, SI). The absorbance spectra did not clearly identify the Lot3 storage degraded samples as being outliers because they appear on the confidence interval ellipse boundary in the scores plot (Figure S9a, SI). This was possibly because these samples did not have many insoluble particles (and higher light scattering), but only a higher percentage of soluble oligomers compared to the other lots. ROBPCA of the normalized pEEM (4 PCs explaining

97% of the total variance) showed that the samples identified as outliers (Figure 6B and Figure S9D, SI) were the storage degraded and mechanically (4-8 minute) stressed samples.

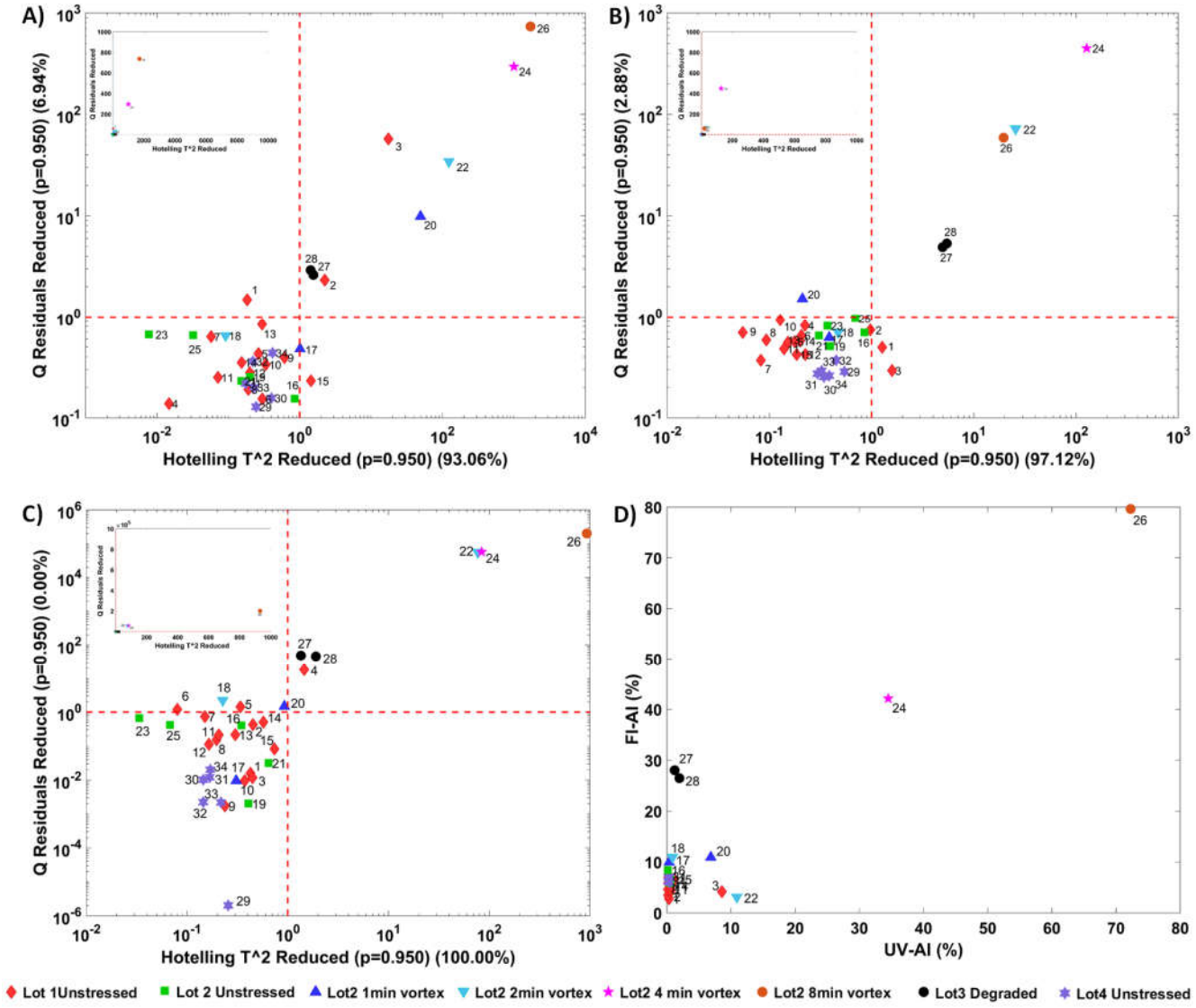

Figure 6: *ROBPCA outlier diagnostic plot for: (A) UV-Vis data, (B) EEM$_{\parallel}$, and (C) RS$_{\parallel}$ and sample distribution based on Fl-AI and UV-AI values (D). Plots A-C were plotted using log scale to facilitate outlier visualisation. The sample numbering is provided in Table 1.*

From the stresses applied one would expect between five (#28, 27, 26, 24, 22) and six (#20) significant outliers. The ROBPCA models built using UV-vis spectra identified 8 major outliers whereas the EEM and RS$_{\parallel}$ derived models yielded 5 and 6 respectively. The UV-AI and Flu-AI scatter plot (***Figure 6***D) clearly identified the most stressed samples (#24 and 26), and one can visually separate another five. It is clear that relying on a single measurement to determine outlier protein samples is not perfectly reliable. In particular it seems that measurements involving scatter are the most sensitive, generating a high rate

of false positives, for example, the good Lot 1 samples #2, 3, and 4 (Figure 6A/C). The $RS_{\|}$ spectra used here (Figure S12, SI), only covered a relatively small spectral range and thus probably does not accurately capture an accurate picture of the physical changes. Furthermore, the data collection methodology used in the fluorescence measurements (single scan per sample) was not ideal and there was significant noise artefacts. The $RS_{\|}$ spectra from the Lot1 samples had an RSD for the area of 6.3 % and one might expect that this would be considerably reduced by averaging multiple scans, thus leading to more accurate outlier identification.

Published version at: **DOI:** 10.1016/j.aca.2019.12.020

Table 1: SEC and simple fluorescence analysis of the five stock and one stressed rIgG solutions. The values in the table represent the mean ± std of three SEC runs of each stock solution (except for solution 4.).

Lot	rIgG Stock Solution	Sample no.	Cert. Analysis Sigma	Conc. (mg/mL)	% Loss[1]	% Mon. area	% Aggreg. Area[2]	M/A Ratio	UV-AI (%)	Fl-AI‖ (%)	RS_{296}/Max Intensity	RS_{316}/Max Intensity
	Trp	--	--	0.01	--	--	--	--	--	0.00±0.00	0.02±0.00	0.02±0.00
		--	--	0.02	--	--	--	--	--	0.00±0.00	0.02±0.00	0.04±0.01
Lot1 (SLBP7449V)	soln. 1 Non Stressed (n=3)	1/2/3	Purity: 96% H_2O:4%	1.00±0.03	--	81.93± 0.96	18.07± 0.96	4.55± 0.29	0.26±0.10	3.50 ± 0.70	1.49±0.02	4.49±0.08
	soln. 2 Non Stressed (n=3)	4/5/6		0.96±0.00	--	82.77± 0.56	**17.93± 1.09**	**4.81± 0.19**	0.54±0.04	5.25±1.01	1.50±0.09	4.33±0.27
	soln. 3 Non Stressed (n=3)	7/8/9		**0.96**±0.00	--	**81.48± 0.63**	**18.52± 0.63**	**4.41± 0.18**	0.20±0.03	5.82±1.22	1.60±0.12	4.66±0.40
	soln. 4 Non Stressed (n=3)	10/11/12		1.00±0.00	--	81.78± 0.94	18.22± 0.94	4.50± 0.28	0.62±0.04	5.75±1.00	1.62±0.12	4.84±0.37
	soln. 5 Non Stressed (n=3)	13/14/15		**0.98**±0.00	--	**81.54± 0.7**	18.46± 0.73	4.43± 0.22	0.85±0.27	6.19±0.71	1.61±0.09	4.71±0.32
	Overall (n=15)	--		0.98±0.02	--	81.86 ± 0.93 (RSD 1.14%)	18.16 ± 0.93 (RSD 5.11%)	4.52 ±0.28 (RSD 6.25%)	0.51±0.27 (RSD 53.19%)	0.51±0.27 (RSD 53.19%)	5.30±1.27 (RSD 23.94%)	1.56±0.10 (RSD 6.38%)
Lot 2 (SLBW8687)	soln. 6 Non-stressed (n=1)	16	Purity: 99% H_2O:2%	0.97	--	86.82±0.37	13.18±0.37	6.59±0.21	0.14	8.43	1.47	3.80
	soln. 6	17		0.99	--	87.03±0.64	12.97±0.64	6.73±0.37	0.37	9.95	1.65	4.45

	1 min vortex (n=1)											
	soln. 6 2 min vortex(n=1)	18		0.98	--	86.58±0.73	13.42±0.73	6.47±0.40	0.86	10.93	1.65	4.30
	soln. 7 Non-stressed (n=4)	19/21/ 23/25		1.05±0.00	--	87.21± 0.96	12.79 ± 0.96	6.84± 0.68	0.50±0.09	6.12±0.35	1.59±0.06	4.46±0.23
	soln. 7 1min vortex (n=1)	20		1.04	1%	85.66± 0.75	14.34 ± 0.75	5.99 ±0.38	6.86	10.93	2.06	5.87
	soln. 7 2min vortex (n=1)	22		1.03	8%	86.27± 0.61	13.73 ± 0.61	6.30± 0.32	10.91	3.11	1.66	7.22
	soln. 7 4min vortex (n=1)	24		1.00	10%	87.81± 0.48	12.19 ± 0.48	7.22± 0.32	34.50	42.11	4.46	14.25
	soln. 7 8min vortex (n=1)	26		0.94	14%	90.91± 0.01	9.99 ± 1.14	9.14±1.09	72.27	79.63	9.07	16.95
Lot 3 (SLBM261 7V)	soln.8 (Poor storage) (n=2)	27/28	Purity: 96% H_2O:4%	0.83±0.00	--	77.47±0.35	22.53±0.35	3.44±0.07	1.58±0.53	27.25±1.11	2.47±0.06	6.19±0.17
Lot4 (SLBZ5214)	soln.9 (n=3)	29/30/31	Purity: 97% H_2O:3%	1.01±0.00	--	88.86±0.28	11.14±0.28	7.99±0.23	0.32±0.07	6.75±0.48	1.44±0.05	3.99±0.15
	soln.10 (n=3)	32/33/34		1.01±0.00	--	88.75±0.17	11.25±0.21	7.89±0.17	0.36±0.07	6.11±0.13	1.42±0.01	3.85±0.05

[1] % of loss was calculated comparing the total area under the curve (AUC) of main SEC peak before and after mechanical stress.
[2] rIgG R03 was an outlier probably due to the UV-Visible measurement and was taken out from the data for the calculation of mean and std of IgG data.

Published version at: **DOI:** 10.1016/j.aca.2019.12.020

4. Conclusions

We have shown here that the use of polarized EEM spectra for the characterisation of variance in immunoglobulins has significant benefits when compared to conventional methods. Despite being the compendial method for protein aggregation measurements, SEC analysis has limitations that affect protein analysis, and in particular that of reversible oligomers/aggregates in solution [54]. The SEC sample preparation (e.g. dilution and mobile phase composition) and chromatographic separation process often lead to the disassociation of these reversible, non-covalently bound, aggregates. All of those factors may lead to an aggregation profile which may not be representative of the situation in the protein the stock solution. This is a critical advantage for spectroscopic based methods which can probe, non-destructively the protein solution without perturbing the balance between reversible and non-reversible aggregates. This was clearly the case here where there was no correlation between the UV-AI and FL-AI measurements and the SEC results, but these simple spectroscopic measurements did clearly indicate the presence of varying amounts of aggregates in solution. UV-visible spectroscopy, although fast, easy, and non-destructive, has poor detection limits and low selectivity. Fl-AI has better sensitivity and reproducibility making it a useful qualitative measure of aggregation; however, this measurement does not provide much conclusive information about protein structure changes, particularly when these were relatively small.

The pEEM measurement in combination with conventional chemometric data analysis can be considered a more comprehensive source of information about protein quality in solution. Here we have shown that it can provide information about concentration, gross structure variation, more subtle structure changes, and aggregate/particle formation (from the Rayleigh scatter) in a single measurement. Using normalized pEEM measurement data makes the method a robust screening method which minimises the effects of lamp intensity variation. In this current iteration, it could be used as a rapid screening method to select samples for more detailed characterisation by the slower and more time-consuming SEC reference method. This proof-of concept study also shows that it could be possible to produce quantitative predictive models for the non-reversible aggregate content. In particular, the sensitivity of the pEEM method to protein structural changes makes it very suitable for measuring differences compared to a reference batch. It should also be noted that the polyclonal antibody case is considerably more complex than that for therapeutic mAbs. For mAbs sourced from biopharmaceutical manufacturing processes, we would expect that the reduction in protein species diversity should lead to

better correlations with SEC data, potentially leading to the development of accurate quantitative models for both the total and non-reversible (i.e. that determined by SEC) aggregate content.

pEEM because of its non-destructive nature coupled with the potential to make rapid measurements (<1-2 minutes) could be easily adapted for online or at-line measurements in protein purification and polishing processes. Further improvement of the method, with additional data sourced from more relevant protein sample types, could lead to the development of a robust Process Analytical Technology (PAT) method for downstream operations.

Supplemental information available

Supporting information is available which provides additional spectral data further details on the chemometric analysis.

Acknowledgments

This publication has emanated from research supported in part by a research grant from Science Foundation Ireland (SFI) and is co-funded under the European Regional Development Fund under Grand number (14/IA/2282, *Advanced Analytics for Biological Therapeutic Manufacture*, to AGR). We also thank Agilent Technologies (Mulgrave Victoria, Australia) for the loan of a fluorescence spectrometer.

References

[1] P.L. Turecek, M.J. Bossard, F. Schoetens, I.A. Ivens, PEGylation of Biopharmaceuticals: A Review of Chemistry and Nonclinical Safety Information of Approved Drugs, J. Pharm. Sci., 105 (2016) 460-475.

[2] A. Beck, L. Goetsch, C. Dumontet, N. Corvaia, Strategies and challenges for the next generation of antibody drug conjugates, Nature Reviews Drug Discovery, 16 (2017) 315-337.

[3] H.C. Liu, K. May, Disulfide bond structures of IgG molecules Structural variations, chemical modifications and possible impacts to stability and biological function, mAbs, 4 (2012) 17-23.

[4] A.P. Chapman, PEGylated antibodies and antibody fragments for improved therapy: a review, Adv Drug Deliver Rev, 54 (2002) 531-545.

[5] S.C. Alley, N.M. Okeley, P.D. Senter, Antibody-drug conjugates: targeted drug delivery for cancer, Current Opinion in Chemical Biology, 14 (2010) 529-537.

[6] F. Kraeber-Bodere, C. Bodet-Milin, C. Rousseau, T. Eugene, A. Pallardy, E. Frampas, T. Carlier, L. Ferrer, J. Gaschet, F. Davodeau, J.F. Gestin, A. Faivre-Chauvet, J. Barbet, M. Cherel, Radioimmunoconjugates for the Treatment of Cancer, Seminars in Oncology, 41 (2014) 613-622.

[7] S. Kaur, G. Venktaraman, M. Jain, S. Senapati, P.K. Garg, S.K. Batra, Recent trends in antibody-based oncologic imaging, Cancer Letters, 315 (2012) 97-111.

[8] S. Wanning, R. Suverkrup, A. Lamprecht, Pharmaceutical spray freeze drying, Int. J. Pharm., 488 (2015) 136-153.
[9] A.S. Rathore, H. Winkle, Quality by design for biopharmaceuticals, Nat Biotechnol, 27 (2009) 26-34.
[10] V. Dotz, R. Haselberg, A. Shubhakar, R.P. Kozak, D. Falck, Y. Rombouts, D. Reusch, G.W. Somsen, D.L. Fernandes, M. Wuhrer, Mass spectrometry for glycosylation analysis of biopharmaceuticals, Trac-Trends in Analytical Chemistry, 73 (2015) 1-9.
[11] M. Sklepari, A. Rodger, A. Reason, S. Jamshidi, I. Prokes, C.A. Blindauer, Biophysical characterization of a protein for structure comparison: methods for identifying insulin structural changes, Analytical Methods, 8 (2016) 7460-7471.
[12] A.G. Ryder, C.A. Stedmon, N. Harrit, R. Bro, Calibration, standardization, and quantitative analysis of multidimensional fluorescence (MDF) measurements on complex mixtures (IUPAC Technical Report), Pure Appl. Chem., 89 (2017) 1849-1870.
[13] J. Bridgeman, M. Bieroza, A. Baker, The application of fluorescence spectroscopy to organic matter characterisation in drinking water treatment, Reviews in Environmental Science and Bio-Technology, 10 (2011) 277-290.
[14] B. Li, P.W. Ryan, M. Shanahan, K.J. Leister, A.G. Ryder, Fluorescence excitation-emission matrix (EEM) spectroscopy for rapid identification and quality evaluation of cell culture media components, Appl. Spectrosc., 65 (2011) 1240-1249.
[15] C. Calvet, B. Li, A.G. Ryder, Rapid quantification of tryptophan and tyrosine in chemically defined cell culture media using fluorescence spectroscopy., J. Pharm. Biomed. Anal., 71 (2012) 89-98.
[16] A. Calvet, B. Li, A.G. Ryder, A rapid fluorescence based method for the quantitative analysis of cell culture media photo-degradation, Anal. Chim. Acta, 807 (2014) 111-119.
[17] S. Heidari, B. Hemmateenejad, S. Yousefinejad, A.A. Moosavi-Movahedi, Excitation-emission matrix fluorescence spectroscopy combined with three-way chemometrics analysis to follow denatured states of secondary structure of bovine serum albumin, J Lumin, 203 (2018) 90-99.
[18] R.C. Groza, B. Li, A.G. Ryder, Anisotropy resolved multidimensional emission spectroscopy (ARMES): A new tool for protein analysis, Anal. Chim. Acta, 886 (2015) 133-142.
[19] Y. Casamayou-Boucau, A.G. Ryder, Accurate anisotropy recovery from fluorophore mixtures using Multivariate Curve Resolution (MCR), Anal. Chim. Acta, 1000 (2018) 132-143.
[20] M. Steiner-Browne, S. Elcoroaristizabal, Y. Casamayou-Boucau, A.G. Ryder, Investigating native state fluorescence emission of Immunoglobulin G using polarized Excitation Emission Matrix (pEEM) spectroscopy and PARAFAC, Chemometrics Intellig. Lab. Syst., 185 (2019) 1-11.
[21] M. Steiner-Browne, S. Elcoroaristizabal, A.G. Ryder, Using polarized Total Synchronous Fluorescence Spectroscopy (pTSFS) with PARAFAC analysis for characterizing intrinsic protein emission, Chemometrics Intellig. Lab. Syst., (2019).
[22] N.S. Lipman, L.R. Jackson, L.J. Trudel, F. Weis-Garcia, Monoclonal versus polyclonal antibodies: Distinguishing characteristics, applications, and information resources, Ilar Journal, 46 (2005) 258-268.

[23] F.C. Wiberg, S.K. Rasmussen, T.P. Frandsen, L.K. Rasmussen, K. Tengbjerg, V.W. Coljee, J. Sharon, C.Y. Yang, S. Bregenholt, L.S. Nielsen, J.S. Haurum, A.B. Tolstrup, Production of target-specific recombinant human polyclonal antibodies in mammalian cells, Biotechnol. Bioeng., 94 (2006) 396-405.
[24] T.P. Frandsen, H. Naested, S.K. Rasmussen, P. Hauptig, F.C. Wiberg, L.K. Rasmussen, A.M.V. Jensen, P. Persson, M. Wiken, A. Engstrom, Y. Jiang, S.J. Thorpe, C. Forberg, A.B. Tolstrup, Consistent Manufacturing and Quality Control of a Highly Complex Recombinant Polyclonal Antibody Product for Human Therapeutic Use, Biotechnol. Bioeng., 108 (2011) 2171-2181.
[25] E. Diamant, A. Torgeman, E. Ozeri, R. Zichel, Monoclonal Antibody Combinations that Present Synergistic Neutralizing Activity: A Platform for Next-Generation Anti-Toxin Drugs, Toxins, 7 (2015) 1854-1881.
[26] A. Eppler, M. Weigandt, A. Hanefeld, H. Bunjes, Relevant shaking stress conditions for antibody preformulation development, European Journal of Pharmaceutics and Biopharmaceutics, 74 (2010) 139-147.
[27] Y. Casamayou-Boucau, A.G. Ryder, Extended wavelength anisotropy resolved multidimensional emission spectroscopy (ARMES) measurements: better filters, validation standards, and Rayleigh scatter removal methods, Methods Appl Fluoresc, 5 (2017) 037001.
[28] D.S. Katayama, R. Nayar, D.K. Chou, J. Campos, J. Cooper, D.G. Vander Velde, L. Villarete, C.P. Liu, M.C. Manning, Solution behavior of a novel type 1 interferon, interferon-tau, J Pharm Sci-Us, 94 (2005) 2703-2715.
[29] A. Hawe, J.C. Kasper, W. Friess, W. Jiskoot, Structural properties of monoclonal antibody aggregates induced by freeze-thawing and thermal stress, European Journal of Pharmaceutical Sciences, 38 (2009) 79-87.
[30] Y. Nomine, T. Ristriani, C. Laurent, J.F. Lefevre, E. Weiss, G. Trave, A strategy for optimizing the monodispersity of fusion proteins: application to purification of recombinant HPV E6 oncoprotein, Protein Eng, 14 (2001) 297-305.
[31] M. Ameloot, M. vandeVen, A.U. Acuna, B. Valeur, Fluorescence anisotropy measurements in solution: Methods and reference materials (IUPAC Technical Report), Pure Appl. Chem., 85 (2013) 589-608.
[32] A. Rinnan, K.S. Booksh, R. Bro, First order Rayleigh scatter as a separate component in the decomposition of fluorescence landscapes, Anal. Chim. Acta, 537 (2005) 349-358.
[33] Y. Casamayou-Boucau, A.G. Ryder, Extended wavelength anisotropy resolved multidimensional emission spectroscopy (ARMES) measurements: better filters, validation standards, and Rayleigh scatter removal methods, Methods Appl Fluores, 5 (2017).
[34] S.Z. Zhang, F.L. Zhao, K.A. Li, S.Y. Tong, Determination of glycogen by Rayleigh light scattering, Anal. Chim. Acta, 431 (2001) 133-139.
[35] M. Hubert, P.J. Rousseeuw, K. Vanden Branden, ROBPCA: A new approach to robust principal component analysis, Technometrics, 47 (2005) 64-79.
[36] D.M. Haaland, E.V. Thomas, Partial Least-Squares Methods for Spectral Analyses .1. Relation to Other Quantitative Calibration Methods and the Extraction of Qualitative Information, Analytical Chemistry, 60 (1988) 1193-1202.
[37] J.R. Lakowicz, Principles of fluorescence spectroscopy, 3rd ed., Springer, New York, 2006.

[38] I.M. Warner, G.D. Christian, E.R. Davidson, J.B. Callis, Analysis of Multicomponent Fluorescence Data, Anal. Chem., 49 (1977) 564-573.
[39] F.T.S. Chan, G.S.K. Schierle, J.R. Kumita, C.W. Bertoncini, C.M. Dobson, C.F. Kaminski, Protein amyloids develop an intrinsic fluorescence signature during aggregation, Analyst, 138 (2013) 2156-2162.
[40] T.N. Tikhonova, N.R. Rovnyagina, A.Y. Zherebker, N.N. Sluchanko, A.A. Rubekina, A.S. Orekhov, E.N. Nikolaev, V.V. Fadeev, V.N. Uversky, E.A. Shirshin, Dissection of the deep-blue autofluorescence changes accompanying amyloid fibrillation, Arch. Biochem. Biophys., 651 (2018) 13-20.
[41] S.K. Panigrahi, A.K. Mishra, Inner filter effect in fluorescence spectroscopy: As a problem and as a solution, Journal of Photochemistry and Photobiology C: Photochemistry Reviews, 41 (2019).
[42] P.H.C. Eilers, P.M. Kroonenberg, Modeling and correction of Raman and Rayleigh scatter in fluorescence landscapes, Chemometr Intell Lab, 130 (2014) 1-5.
[43] S. Bhattacharjee, DLS and zeta potential - What they are and what they are not?, J. Controlled Release, 235 (2016) 337-351.
[44] C.M. Maguire, M. Rosslein, P. Wick, A. Prina-Mello, Characterisation of particles in solution - a perspective on light scattering and comparative technologies, Science and Technology of Advanced Materials, 19 (2018) 732-745.
[45] P. Hong, S. Koza, E.S.P. Bouvier, A Review Size-Exclusion Chromatography for the Analysis of Protein Biotherapeutics and Their Aggregates, J Liq Chromatogr R T, 35 (2012) 2923-2950.
[46] V. Joshi, T. Shivach, N. Yadav, A.S. Rathore, Circular Dichroism Spectroscopy as a Tool for Monitoring Aggregation in Monoclonal Antibody Therapeutics, Anal. Chem., 86 (2014) 11606-11613.
[47] B. Raynal, P. Lenormand, B. Baron, S. Hoos, P. England, Quality assessment and optimization of purified protein samples: why and how?, Microbial Cell Factories, 13 (2014) 10.
[48] S. Rudiuk, L. Cohen-Tannoudji, S. Huille, C. Tribet, Importance of the dynamics of adsorption and of a transient interfacial stress on the formation of aggregates of IgG antibodies, Soft Matter, 8 (2012) 2651-2661.
[49] A. Aitken, M.P. Learmonth, Protein Determination by UV Absorption, in: J.M. Walker (Ed.) Protein Protocols Handbook, Third Edition, Humana Press Inc, Totowa, 2009, pp. 3-6.
[50] R. van der Kant, A.R. Karow-Zwick, J. Van Durme, M. Blech, R. Gallardo, D. Seeliger, K. Assfalg, P. Baatsen, G. Compernolle, A. Gils, J.M. Studts, P. Schulz, P. Garidel, J. Schymkowitz, F. Rousseau, Prediction and Reduction of the Aggregation of Monoclonal Antibodies, J Mol Biol, 429 (2017) 1244-1261.
[51] R.W. Kennard, L.A. Stone, Computer Aided Design of Experiments, Technometrics, 11 (1969) 137-&.
[52] A.G. Gonzalez, M.A. Herrador, A.G. Asuero, Intra-laboratory testing of method accuracy from recovery assays, Talanta, 48 (1999) 729-736.
[53] J. Mandel, F.J. Linnig, Study of Accuracy in Chemical Analysis using Linear Calibration Curves, Anal. Chem., 29 (1957) 743-749.
[54] J.S. Philo, Is any measurement method optimal for all aggregate sizes and types?, Aaps Journal, 8 (2006) E564-E571.