

# Monitoring Antibody Aggregation in Early Drug Development Using Raman Spectroscopy and Perturbation-Correlation Moving Windows

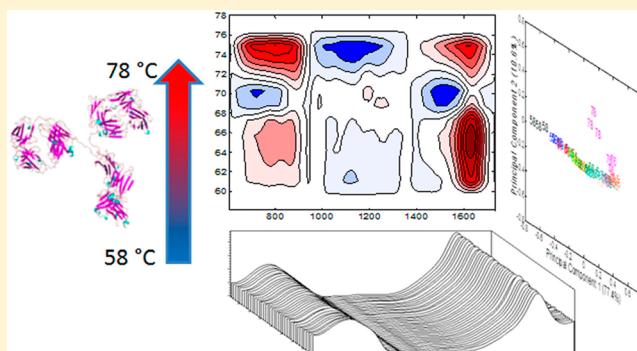
Ramón Gómez de la Cuesta,<sup>†</sup> Royston Goodacre,<sup>‡</sup> and Lorna Ashton<sup>\*,‡,§</sup>

<sup>†</sup>Lonza Biologics plc, Granta Park, Great Abington, Cambridge, CB21 6GS, U.K.

<sup>‡</sup>School of Chemistry, Manchester Institute of Biotechnology, University of Manchester, Manchester, M1 7DN, U.K.

## Supporting Information

**ABSTRACT:** In this study, we demonstrate the sensitivity of two-dimensional perturbation-correlation moving windows (PCMW) to characterize conformational transitions in antibodies. An understanding of how physicochemical properties affect protein stability and instigate aggregation is essential for the engineering of antibodies. In order to establish the potential of PCMW as a technique for early identification of aggregation mechanisms during antibody development, five antibodies with varying propensity to aggregate were compared. Raman spectra were acquired, using a 532 nm excitation wavelength as the protein samples were heated from 56 to 78 °C and analyzed with PCMW. Initial principal component analysis confirmed a trend between the observed spectral variations and increasing temperature for all five samples. Analysis using PCMW revealed that when spectral variations were directly related to temperature, distinct differences in conformational changes could be determined between samples related to protein stability, providing a greater understanding of the aggregation mechanisms of problematic antibody variants.



Protein-based biopharmaceuticals are becoming increasingly popular therapeutic agents, yet the production and characterization of such therapeutics are extremely expensive and continue to pose numerous analytical challenges. Over 90% of drug candidates will fail during clinical development, and this figure increases if preclinical stages of development are included. More and more pharmaceutical and biotechnology companies are adopting the concept of “developability”, that is, the application of design criteria and problem solving at the very early stage of drug development, thereby avoiding future complications after large investments of time and money.<sup>1,2</sup> Raman spectroscopy combined with two-dimensional perturbation-correlation moving windows (PCMW) has been shown to be an effective tool in the investigation of perturbation-induced structural changes in proteins.<sup>3–6</sup> However, these studies have tended to focus on single variants of well-characterized proteins, and the technique is yet to be applied in the development of protein-based biopharmaceuticals. In this study, we demonstrate the vast potential of Raman spectroscopy and PCMW as analytical tools for protein stability in the early stages of protein engineering and drug design.

An important focus for developability is the understanding of protein structure and how this affects protein aggregation, as poor protein stability not only influences production but also activation, longevity, and shelf life.<sup>7–9</sup> Aggregation occurs when patches of hydrophobic areas of protein molecules are attracted to each other, involving noncovalent interactions, often forming large insoluble particles. In some proteins an associated state is

the actual native form and aggregation is necessary for protein activity. However, in the majority of therapeutic proteins, aggregation is highly undesirable, as it may affect the immunogenicity of the drug and in larger aggregate particles can lead to an adverse effect upon administration.<sup>7,10</sup> Several techniques are well-established and frequently used for the quantification of protein aggregation, including size exclusion chromatography<sup>10,11</sup> and analytical ultracentrifugation.<sup>12,13</sup> While these traditional techniques are useful in assessing the amount of aggregation induced by various factors, such as high temperature, acidic pH, and agitation, they do not provide any further insight into the mechanisms of aggregation. Consequently, in order to improve engineering of new protein therapeutics including antibodies, there is a need for alternative techniques that can provide a greater understanding of the effects of physicochemical properties on protein stability and aggregation.

Raman spectroscopy, which has the advantage of being nondestructive, label-free, and insensitive to water, has long been used for therapeutic protein analysis and is ideal for online monitoring of protein production.<sup>14–17</sup> When combined with 2D correlation analysis (2DCOS), Raman spectroscopy has the distinctive ability to follow dynamic fluctuations in spectra induced by an external perturbation.

**Received:** June 6, 2014

**Accepted:** October 20, 2014

**Published:** October 20, 2014

In this study we have applied 2DCOS PCMW<sup>18</sup> to Raman spectra acquired during heating of five protein variants important in the engineering of new antibodies, one control antibody and four mutated antibodies, each with a different propensity to aggregate. We show that by using Raman spectroscopy and PCMW to compare conformational transitions that occur during aggregation, new insights into the aggregation mechanisms of specially engineered mutations can be gained, which will aid in the future development of therapeutic antibodies.

## ■ EXPERIMENTAL SECTION

**Samples.** Five antibody samples, each with a different propensity to aggregate (Table 1) were kindly provided by

**Table 1. Percentage of Soluble Aggregates As Measured by HPLC for the Five Antibody Samples before and after Heating to 60 °C for 2 h (Accuracy ±0.3%)**

sample	% soluble aggregates (HPLC)	
	native	heated 60 °C (2 h)
1	0.78	0
2	62.2	64
3	5	63.8
4	14.7	12.9
5	0	0

Lonza Biologics (Cambridge, U.K.) and labeled 1–5. Sample 1 is a control antibody with negligible levels of aggregation, sample 2 is an antibody used as a model of high aggregation for internal development work in Lonza, and samples 3–5 are variants of sample 2 in which point mutations have been introduced to reduce the level of aggregation. All five antibodies were diluted at 1 µg/µL in 500 µL of sodium phosphate buffer at pH 7.3 and stored at –20 °C. Size exclusion chromatography was used to determine relative amounts of monomer and aggregates present in the purified samples. Chromatography was carried out using an Agilent 1200 binary pump HPLC system equipped with a Zorbax GF-250 HPLC column (Agilent). Antibodies were detected at 280 nm and the peak chromatograms were analyzed using Agilent Chemstation software. The proportion of monomer and aggregates in samples was determined by calculation of the integrated peak areas of each component relative to the total integrated peak areas.

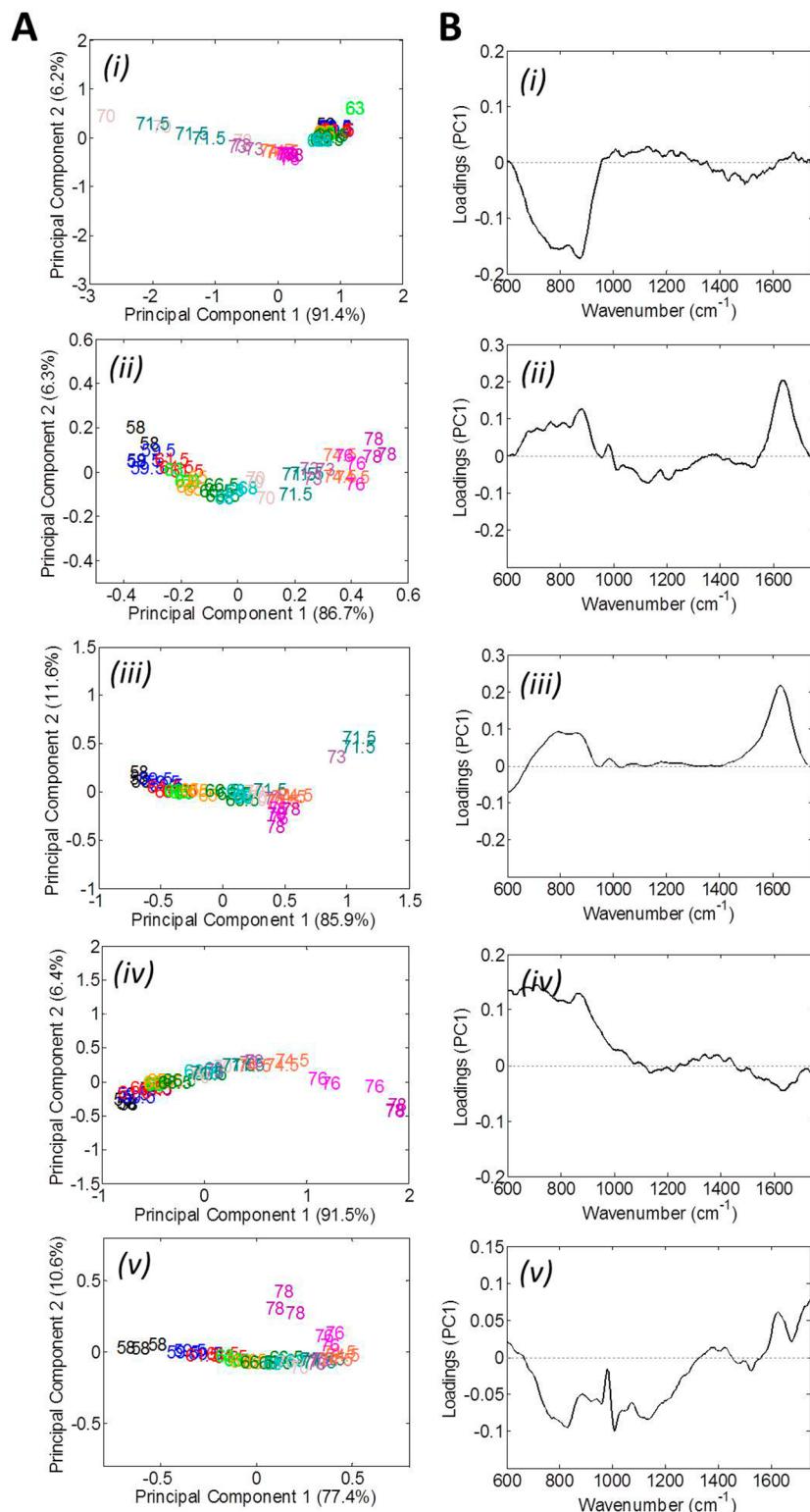
**Raman Spectra.** Raman spectra were measured using a benchtop Ocean Optics fiber optic probe with an excitation wavelength of 532 nm and laser power at the sample of ~8 mW. A 500 µL portion of each sample was placed in a 1 mL shell vial and tightly sealed with a plastic stopper to prevent evaporation. Each sample vial was placed into a water-heated cell holder, and the probe was aligned to focus directly onto the sample before covering over the complete Raman setup. The temperature of the cell holder was controlled using a heating water circulator (Julabo model F12-ED) and sample temperature was calibrated using buffer with an error of ±1 °C. For each sample, the sample temperature was initially set at 55 °C and left for 30 min, and three repeat spectra were then consecutively acquired for 60 s each. After the third spectrum was recorded, the temperature setting of the water heater was increased by 2 °C and left for a further 4 min before the next three spectra were measured. This was repeated until a maximum heater temperature of 86 °C was reached. From

the buffer calibration, the sample temperature for each set of three repeat spectra was determined as 58, 59.5, 61.5, 63, 65, 66.5, 68, 70, 71.5, 73, 74.5, 76, and 78 °C. Despite acquiring spectra for 60 s, each visual inspection of the raw data revealed poor signal-to-noise and some interference from fluorescent background as a result of the experimental setup (Figure S-1, Supporting Information). However, through the careful application of well-established preprocessing methods (as detailed below), which were robustly checked not to induce spectral distortions and applied consistently to all data sets, it was possible to compare temperature-induced spectral variations across the five samples.

**Data Preprocessing.** With the exception of the perturbation-moving windows calculation, where 2Dshige freeware (<https://sites.google.com/site/shigemorita/home/2dshige>) was used, all other data processing was carried out using MATLAB software version 2011a (The Math Works). In order to compare the data sets for each of the five samples directly, spectra were normalized (standard normal variate), smoothed, and then baseline-corrected (Figure S-2, Supporting Information). Principal component analysis (PCA) was carried out on the full data set for each sample for the wavenumber range 600–1800 cm<sup>-1</sup> in order to investigate protein conformational changes. For two-dimensional correlation analysis (2DCOS), the three repeat spectra were averaged. It has been previously reported that in order to apply 2D correlation analysis calculations, the perturbation steps need to be evenly spaced, that is, measured using a fixed increment.<sup>19</sup> As the final calibration temperature steps were not precisely evenly spaced, the data sets were converted into evenly spaced data using an interpolation procedure. This procedure produces results that are based upon the underlying assumption that the dynamic system does not change significantly throughout the missing sections of data. As the data collected here varied in a consistent nature between samples, an interpolation procedure was performed in MATLAB using a Savitzky–Golay spline estimation to generate interpolated data with a temperature range from ~58 to 78 °C in ~1.53 °C steps.

## ■ RESULTS AND DISCUSSION

**Principal Component Analysis.** Before applying 2D correlation spectroscopy (2DCOS) techniques, principal component analysis (PCA) was used to confirm the presence of temperature-induced variability in the Raman spectral data sets for all five samples. PCA is a well-established and extensively used technique that reduces the dimension of multivariate data to a small number of uncorrelated variables referred to as principal components (PCs), which represent the natural variance in the data. It is frequently used with Raman spectral sets to demonstrate reproducibility of spectra and to determine trends or patterns within the data. Figure 1A displays the PCA scores plots of PC1 versus PC2 for each of the five samples. A trend/trajecory with increasing temperature can be observed predominantly along PC1 for all five samples with percentage variance greater than 77%. As PCA is an unsupervised method, the fact that the data points are arranged in increasing temperature order confirms that the spectra do indeed contain enough structural information to effect temperature-induced discrimination. Figure 1A(i,iii) does reveal some potential outliers (data points positioned outside of the general trend, one for sample 1 and three for sample 3) possibly due to experimental variation not removed by data processing. Consequently, these outlying spectra were not included in the

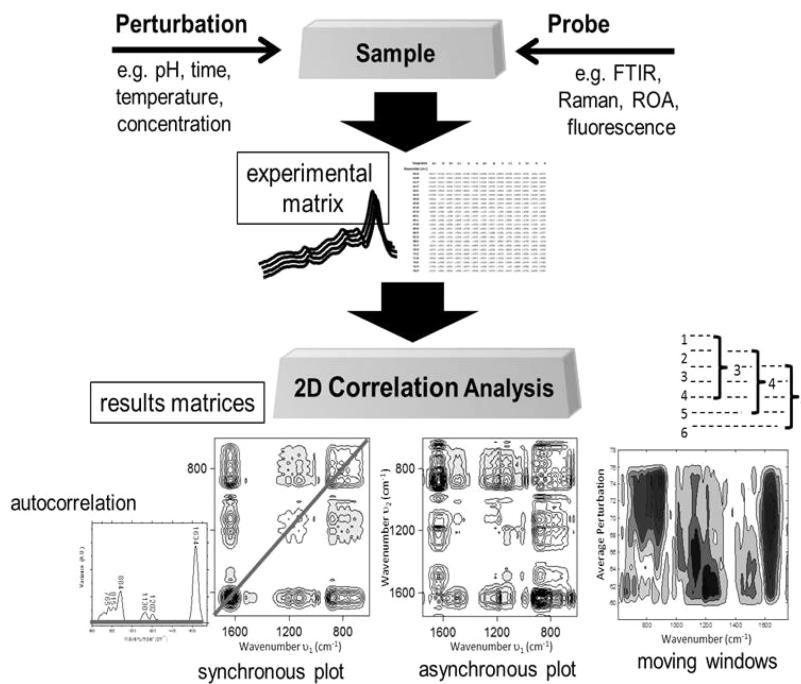


**Figure 1.** PCA of Raman spectral sets monitoring temperature-induced changes: (A) scores plots of PC1 versus PC2 and (B) loadings plots of PC1 for (i) sample 1, (ii) sample 2, (iii) sample 3, (iv) sample 4, and (v) sample 5.

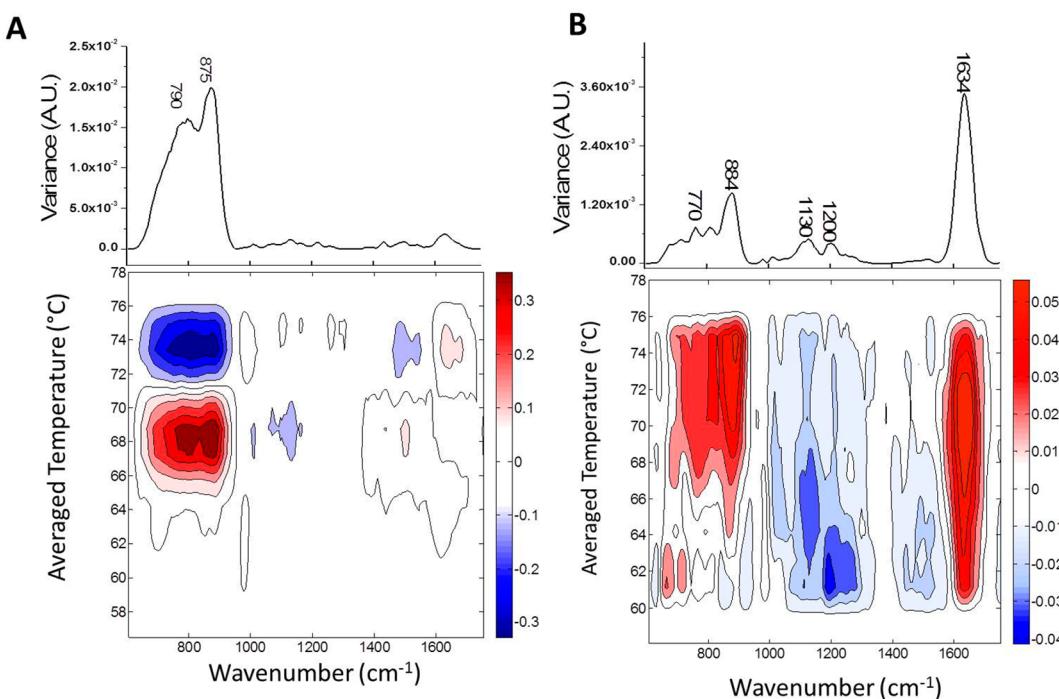
2DCOS analysis. What is also noticeable in Figure 1A is that although all three repeat spectra for each sample and time point were acquired within 3 min of each other, a trend within each group of three can also be observed, suggesting that spectral variation occurred during this short time period. The trend with temperature combined with the large percentage variance

strongly indicates that the spectral changes occurring in the Raman data sets are temperature-induced.

PCA loadings plots can be used to determine the specific wavenumber regions that account for the variance in each PC. Figure 1B compares the loadings plots of PC1 for all five samples. Although the relative intensities of the peaks shown in loadings plots vary, the specific wavenumber regions of variance



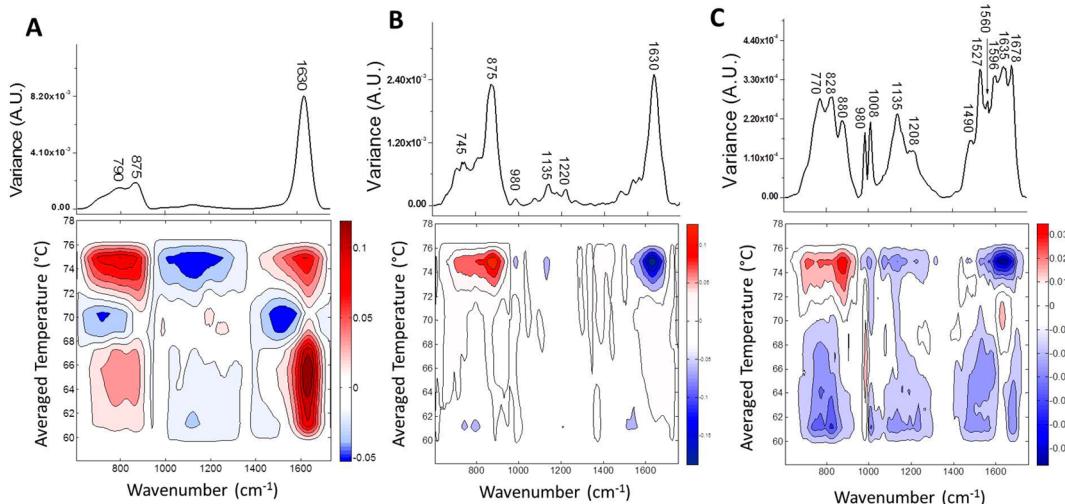
**Figure 2.** Schematic of 2D correlation analysis. A data set is ordered in the direction of perturbation-induced spectral variations to form an experimental matrix. A cross-correlation calculation is then applied to the experimental matrix, resulting in a new set of matrices, including synchronous, asynchronous, autocorrelation, and moving window plots, with the aim to improve visualization and interpretation of the perturbation-induced spectral variations.



**Figure 3.** PCMW plots with autocorrelations calculated from (A) sample 1 and (B) sample 2 spectral data sets. PCMW plotted as a function of spectral wavenumber and average translating window temperature. Contours shaded red indicate peaks that are increasing in intensity with increasing temperature, while blue shading indicates decreasing peak intensity with increasing temperature. The darker the shade of blue or red and the closer together the contours, the greater the change in intensity, as indicated by the color shading bar. The scale on this bar is of arbitrary units.

are somewhat consistent, particularly for samples 1–4. The key regions accounting for the largest percentage of spectral variance can be observed in the region  $\sim 680$ – $900\text{ cm}^{-1}$  associated with H-bonding of side chain residues, the phenylalanine assigned to the peak at  $\sim 1008\text{ cm}^{-1}$  and the

amide I region ( $1600$ – $1700\text{ cm}^{-1}$ ) associated with protein secondary structure. What cannot be determined from PCA is the exact time course of the observed variations which is needed to elucidate the evolution of these temperature-dependent spectral changes. Only the spectral regions that



**Figure 4.** PCMW plots with autocorrelations calculated from (A) sample 3, (B) sample 4, and (C) sample 5 spectral data sets. PCMW plotted as a function of spectral wavenumber and average translating window temperature. Contours shaded red indicate peaks that are increasing in intensity with increasing temperature, while blue shading indicates decreasing peak intensity with increasing temperature. The darker the shade of blue or red and the closer together the contours, the greater the change in intensity, as indicated by the color shading bar. The scale on this bar is of arbitrary units.

are changing to the largest extent across the complete time course can be identified and not how specific temperatures affect the spectral features. In order to determine more precisely how the spectral variations relate to specific, smaller temperature ranges, we have applied two-dimensional PCMW.<sup>18</sup>

**Two-Dimensional Perturbation-Correlation Moving Windows.** Since its development during the 1990s, 2DCOS has been increasingly applied to the investigation of biomolecules due to its ability to follow dynamic fluctuations in spectra induced by an external perturbation.<sup>3,19</sup> Figure 2 illustrates the general schematic for 2DCOS, where it is clear that the technique is not limited to any specific type of spectroscopic probe or perturbation and only requires the collection of a data set ordered in the direction of relatively consistent perturbation-induced changes to form an experimental matrix. 2DCOS techniques are based on cross-correlation calculations that are applied to the experimental matrix, resulting in a new set of matrices that aim to improve visualization and therefore interpretation of the spectral variations. The original 2DCOS technique, referred to as generalized two-dimensional spectroscopy, focused on the calculation and interpretation of the synchronous and asynchronous matrices, from which similarities and differences in behaviors between peaks could be more readily identified.<sup>20</sup> The advantage of PCMW is that it includes both the synchronous and asynchronous data on a plane between a spectral variable axis and a perturbation axis, thereby directly relating the perturbation to the direction of intensity changes.<sup>21</sup> In a PCMW calculation, the complete spectral set is subdivided into smaller “windows” and labeled as the perturbation average. For the antibody samples data sets used in this study, the spectra were divided into subsets of five spectra each, with the first subset containing the first to fifth spectra, the second subset the second to sixth spectra, and so on, until all spectra were included in a subset covering a 6 °C temperature increase. The results are plotted as a function of spectral wavenumber and the average translating perturbation (temperature), and for the purposes of this paper, red and blue shading has been used

to aid interpretation: contours shaded red indicate peaks that are *increasing* in intensity with increasing temperature, while blue shading indicates peaks *decreasing* in intensity with increasing temperature. Furthermore, the relative extent of intensity changes can be directly compared by the color shading and the number of contours; the darker the shade of blue or red and the closer together the contours, the greater the change in intensity, as identified by the color bars shown in Figures 3 and 4.

**Samples 1 (Control) and 2 (Model of High Aggregation).** Figure 3A,B displays the PCMW (lower panels) along with the autocorrelation (upper panels) for the sample 1 data set (control sample), alongside the sample 2 data set. As also observed in the PCA loadings of PC1 for sample 1 [Figure 1B(i)], features in the region 650–900 cm<sup>-1</sup> dominate and therefore changed to the greatest extent as temperature was increased. Some features can be observed in sample 1 PCMW outside of this region, but as determined by the autocorrelation, they are extremely weak compared to the peak intensities at ~770 and 875 cm<sup>-1</sup>. Raman bands at ~760–770 and 875–880 cm<sup>-1</sup> have been extensively assigned to tryptophan residuals (Table 2) the exposure of which to the solvent environment has previously been reported to change following heat-induced aggregation of antibodies.<sup>22</sup> It can also be determined from sample 1 PCMW that spectral variations transpired in two distinct temperature phases from 66 to 70 and 72 to 76 °C, again supported by the PCA scores [Figure 1A(i)], where the data points can be divided into two separate groups from ~58 to 68 and 70 to 78 °C. As sample 1 is the control antibody, which does not readily aggregate upon heating (Table 1), two phases of change are to be expected with temperature-induced denaturation.<sup>23–25</sup> Studies of thermal denaturation of antibodies using differential scanning calorimetry reported two transitions, one with a denaturation temperature ( $T_m$ ) of 61 °C and a second at 71 °C, attributed to the denaturation of the F<sub>ab</sub> domain and the F<sub>c</sub> domain, respectively.<sup>23,24</sup> The temperature values observed in sample 1 PCMW are slightly higher; however Vermeer and Norde<sup>23</sup> reported a heating rate dependency with  $T_m$  increasing with increased heating rates.

**Table 2. Proposed Raman Assignments<sup>a</sup> Based on the PCMW Results**

wavenumber (cm <sup>-1</sup> )	assignment
755	tryptophan <sup>26,35</sup>
770	tryptophan <sup>26,27</sup>
828	tyrosine <sup>31,35</sup>
875–885	tryptophan <sup>26,27</sup>
980	glycine/β-sheet <sup>33,36</sup>
1008	phenylalanine <sup>36</sup>
1117–1135	glycine <sup>33</sup>
1200	turns/loops of β-structure <sup>28</sup>
1208	tyrosine <sup>37</sup>
1490	histidine <sup>31</sup>
1550–1565	tryptophan <sup>27,36</sup>
1596	glycine <sup>33</sup>
1620–1635	H-bonding of β-structure <sup>7,34</sup>
1678	β-structure <sup>26,31</sup>

<sup>a</sup>Bands identified as being important from PCMW analyses.

The samples in this study were heated at a rate of ~3.5 °C/min, and our results are consistent with the temperatures reported by these authors for higher heating rates.<sup>23</sup> Furthermore, it can be determined from sample 1 PCMW (Figure 3A) that from 66 to 70 °C spectral changes are due to an increase in intensity (red shading), while spectral variations from 72 to 76 °C are due to a decrease in intensity (blue shading). The decrease in intensity of the Raman bands at 700 and 875 cm<sup>-1</sup> indicates that the tryptophan residues in sample 1 were increasingly exposed to the solvent at the highest temperatures, 72 to 76 °C,<sup>26,27</sup> most likely as a result of the unfolding of the F<sub>c</sub> region. From 66 to 70 °C, the bands at 700 and 875 cm<sup>-1</sup> increased in intensity, indicating that tryptophan residues became less exposed to the solvent,<sup>26,27</sup> which may be due to a change in the orientation of tryptophan residues in the F<sub>c</sub> domain as the F<sub>ab</sub> region unfolded.

Whereas sample 1 is extremely stable, showing no signs of aggregation even after heating at 60 °C for 2 h, sample 2 is extremely unstable, with ~62% of the sample already aggregated before any heat treatment was applied (Table 1). The autocorrelation and PCMW calculated from the sample 2 spectral set (Figure 3B) indicate that spectral variation occurred across the full temperature range with no distinct phases. As sample 2 had already aggregated to a considerable extent at the start of the experiment, spectral variation must be monitoring further aggregation, rather than any structural change due to temperature-induced unfolding. As with sample 1, Raman bands assigned to tryptophan residues at 770 and 884 cm<sup>-1</sup> can be observed in sample 2 PCMW to increase in intensity with increasing temperature, indicating decreased exposure as further aggregation occurred. From the autocorrelations for sample 2, the largest increase in intensity can be observed to have occurred in the peak at 1634 cm<sup>-1</sup>, as also indicated by the PC1 loadings [Figure 1B(ii)] and can be determined from the PCMW to have increased steadily throughout the increasing temperature range. Raman bands in the region of ~1620–1635 cm<sup>-1</sup> are frequently assigned to α-helical structure in protein;<sup>28–30</sup> however, there is only a very small amount of α-helical structure in this specific antibody, and we would not expect to see such dominant changes as indicated by the PCMW. Further possible Raman assignments for the peak at ~1634 cm<sup>-1</sup> include typtophan,<sup>30,31</sup> glycine residues,<sup>32,33</sup> or increases in intermolecular H-bonding of β-sheet structure that

occur during protein aggregation, with possible associations with fibril formation.<sup>7,34</sup> As this peak is not observed to vary significantly in the sample 1 PCMW (Figure 3A), the increase in intensity observed in sample 2 (Figure 3B) is most likely due to aggregation and therefore can be assigned to increasing intermolecular H-bonding of β-structure.

Features decreasing in intensity can also be observed in Figure 3B; however, the autocorrelation can again be used to establish that the majority of these spectral variations are weaker than those observed at 770, 884, and 1634 cm<sup>-1</sup>. The peaks seen at ~1200 and 1130 cm<sup>-1</sup> can be observed in the PCMW to change to their greatest extent at ~61 and 66 °C, respectively. The Raman band at 1200 cm<sup>-1</sup> has been previously associated with β-structure turns and loops, while the Raman band at 1117–1130 cm<sup>-1</sup> has been assigned to glycine (Table 2). Glycine residues are present in the loops that connect β-strands in antibodies. The spectral variations observed here may be due to conformation changes of the aggregates during initial heating before further aggregation and, more specifically, reorientation of the loops connecting the β-strands.

**Antibody Variant Samples 3–5.** When the autocorrelation for sample 2 (Figure 3B) is compared with those of samples 3 and 4 (Figure 4A,B), they look similar, with the most significant changes occurring in the regions of 750–900 and 1575–1700 cm<sup>-1</sup>, despite the fact that the three samples have very different propensities to aggregate (Table 1). However, when the PCMW plots are directly compared, very different features can be observed, demonstrating the potential of PCMW for determination of specific aggregation mechanisms in various antibody variants.

The PCMW for sample 3 (Figure 4A) indicates three potential transitions phases occurring at ~62–68, 68–71, and 73–76 °C, although as identified by the reduced number of contours and the color shading, changes between 68 and 71 °C are far weaker than those observed in the other two temperature phases. As well as identifying the two separate unfolding steps for the two domains (F<sub>ab</sub> at 61 °C and F<sub>c</sub> at 71 °C), Vermeer and Norde<sup>23,24</sup> also suggested that unfolding is immediately followed by a third irreversible aggregation step, and hence, the denaturation of the antibody domains should be described as a three-state process. In the region of 770–900 cm<sup>-1</sup> (Figure 4A), two initial phases can be determined, with increasing intensity (shaded red) from 62 to 68 °C followed by decreasing intensity (shaded blue) from 68 to 71 °C. Although occurring at different temperature ranges, these changes are the same as those observed in sample 1 PCMW (Figure 3A). It is therefore feasible that similar unfolding of the two domains is occurring for sample 3 with the F<sub>ab</sub> domain unfolding resulting in less exposed tryptophan residues within the F<sub>c</sub> domain before the F<sub>c</sub> domain then unfolds. A similar loss of intensity can also be determined by the contour pattern and shading in the region of 1470–1575 cm<sup>-1</sup> between 62 and 68 °C in Figure 4A. Raman bands at ~1490 cm<sup>-1</sup> have been assigned to histidine, while bands at 1550–1565 cm<sup>-1</sup> have been assigned to tryptophan residues (Table 2), supporting the suggestion that amino acid residues are becoming more exposed to the solvent with the unfolding of the F<sub>c</sub> domain. The lower temperature ranges at which these two unfolding transitions occur indicate the reduced stability of sample 3 compared to sample 1. The final transition phase observed for sample 3 for the region of 770–900 cm<sup>-1</sup> from 73 to 76 °C is similar to the spectral variations observed for sample 2, suggesting that

tryptophan residues become less exposed to the solvent as the sample aggregates.

In Figure 4A, a number of dark contours can also be observed at  $\sim 1630\text{ cm}^{-1}$  from 73 to  $76\text{ }^{\circ}\text{C}$ , which are also consistent with sample 2 observations and, as previously discussed, assigned here to the increasing intermolecular H-bonding of the  $\beta$ -structure. However, a similar feature is also observed from 62 to  $68\text{ }^{\circ}\text{C}$  in Figure 4A, which may indicate aggregation is occurring before the two domains have unfolded. Sample 3 has been shown to be relatively stable at native temperatures but upon heating to aggregate extensively; consequently, it may be that with rapid heating, as applied here, there is a mixture of partially unfolded molecules with aggregating populations. Spectral variations in the region of  $1100\text{--}1250\text{ cm}^{-1}$  are also similar to those observed in sample 2 and may be monitoring conformational changes in  $\beta$ -strand turns and loops. In the case of sample 3, these variations are divided into two phases from 60 to 65 and 73 to  $76\text{ }^{\circ}\text{C}$  and may additionally indicate some aggregation before complete unfolding of all antibody molecules followed by further aggregation.

From the HPLC results comparing the percentage of soluble aggregates in the native sample with the sample after heating at  $60\text{ }^{\circ}\text{C}$  for 2 h (Table 1), we would expect sample 5, and possibly sample 4, to behave in a similar way to sample 1, as none of these antibodies show increased aggregation with heating. When the autocorrelations are initially compared (Figures 3A and 4B,C), they look very different, but this can be accounted for by the large variation in actual intensity counts of the variance (*y*-axis). While sample 5 autocorrelation may show the most detailed variations, these are actually very weak in intensity (maximum count of  $\sim 4 \times 10^{-4}$ ) compared to those of sample 1 (maximum count of  $2 \times 10^{-2}$ ). These results clearly highlight the need for caution when comparing contour plots as to the choice of maximum contour levels and color shading as well as a need to compare the absolute intensities of the autocorrelation plots. Although the spectral variations identified for sample 5 are weaker than those observed for sample 1, the PCA scores plot [Figure 1A(v)] does confirm that the changes being observed are related to temperature. In the PC1 loadings plot for sample 5 [Figure 1B(v)], as well as in the autocorrelation (Figure 4C), a large number of sharp peaks can be observed that correlate with Raman bands assigned to amino acid residues and  $\beta$ -structure (Table 2). As previously discussed, the large temperature-induced spectral variations observed for the control sample 1 indicate that the  $F_{ab}$  domain unfolds before the  $F_c$  domain, denaturing the protein. The weaker temperature-induced spectral changes observed for sample 5 may indicate that neither domain completely unfolds, so the antibody is not denatured, but more subtle structural changes occur, and any further unfolding is restricted due to the specific amino acid sequence. The point mutations introduced in sample 5 lead to an increase in the number of internal salt bridges, which could account for the increase in stability of the antibody and explain these results. The increase in stability may also account for the slightly lower percentage variation in PC1 (77%) compared to the other samples (>86%).

What is interesting to note is that very similar spectral variations can be observed from 73 to  $76\text{ }^{\circ}\text{C}$  in the regions of  $650\text{--}900$  and  $1600\text{--}1670\text{ cm}^{-1}$  for both samples 4 and 5 (Figure 4B,C). Although native sample 4 had some aggregation ( $\sim 14\%$  soluble aggregates), this did not increase when heated at  $60\text{ }^{\circ}\text{C}$  for 2 h and, in fact, is shown to have reduced by  $\sim 2\%$

of soluble aggregates. The PCMW therefore suggests that the antibody is relatively stable up to temperatures of  $72\text{ }^{\circ}\text{C}$ . However, while the peaks at  $770$  and  $880\text{ cm}^{-1}$  increase in intensity after  $72\text{ }^{\circ}\text{C}$ , as observed for samples 2 during aggregation, the peak at  $1630\text{ cm}^{-1}$  is observed to decrease in intensity. The decrease at  $1630\text{ cm}^{-1}$  may indicate a loss of H-bonding of  $\beta$ -structure or, alternatively, a general loss of overall secondary structure. Consequently, the spectral sets for samples 2 and 3 may be monitoring a different temperature-induced process than those for samples 1 and 4, where aggregation occurs.

## CONCLUSIONS

We have demonstrated the sensitivity of Raman spectroscopy to compare and characterize conformational transitions as well as aggregation mechanisms in five antibodies, each with a different propensity to aggregate. Although standard techniques used in early antibody development, such as HPLC, provide information on the level of aggregation, they reveal little information on specific conformational transitions in protein domains, secondary structures, or residues. From the PCA results in this study it was possible to identify differences in temperature-induced trends affecting secondary structure conformations in each antibody variant, but only the region of greatest spectral variance across the complete temperature range could be determined. It was only by the application of PCMW that observed spectral variations could be related to potential conformational transitions identified at specific temperature ranges. Conformational changes identified from the PCMW for sample 1 were consistent with previously reported unfolding mechanisms.<sup>23,24</sup> By comparing the PCMW of the other four samples, distinct differences in aggregation mechanisms as a result of the various mutations could be determined. This type of information provided by PCMW can greatly aid in the initial screening for promising antibody variants. Although the concentrations used in early screening (as used in this study) are far below the final product concentrations (>50 mg/mL), aggregation at this early stage of development is a strong indicator of future problems. By understanding the precise mechanism of aggregation for each variant, only the most promising candidates can be selected for further development, thus saving time and money. Consequently, Raman spectroscopy combined with 2DCOS techniques has the potential to be an effective tool in the investigation of how physiochemical properties affect protein stability and instigate aggregation, essential for the early engineering of new protein therapeutics.

## ASSOCIATED CONTENT

### S Supporting Information

Figure S-1, the raw Raman data sets, and Figure S-2, the preprocessed Raman data discussed in the text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: l.ashton@lancaster.ac.uk.

### Present Address

<sup>§</sup>Department of Chemistry, Faraday Building, Lancaster University, Lancaster, LA14YB, U.K.

### Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We would like to thank the U.K. EPSRC and BBSRC and the industrial members of the Bioprocessing Research Industry Club (BRIC) for funding.

## REFERENCES

- (1) Lauer, T. M.; Agrawal, N. J.; Chennamsetty, N.; Egodage, K.; Helk, B.; Trout, B. L. *J. Pharm. Sci.* **2011**, *101*, 102–115.
- (2) Zurdo, J. *Pharm. Bioprocess.* **2013**, *1*, 29–50.
- (3) Noda, I. *Vib. Spectrosc.* **2012**, *60*, 146–153.
- (4) Shashilov, V. A.; Lednev, I. K. *J. Raman Spectrosc.* **2009**, *1749*–1758.
- (5) Brewster, V. L.; Ashton, L.; Goodacre, R. *Anal. Chem.* **2011**, *83*, 6074–6081.
- (6) Ashton, L.; Pudney, P. D. A.; Blanch, E. W.; Yakubov, G. E. *Adv. Colloid Interface Sci.* **2013**, *199*, 66–77.
- (7) Nielsen, L.; Frokjaer, S.; Carpenter, J. F.; Brange, J. *J. Pharm. Sci.* **2001**, *90*, 29–37.
- (8) Goddard, P. *Adv. Drug Delivery Rev.* **1991**, *6*, 103–131.
- (9) Beck, A.; Wagner-Rousset, E.; Ayoub, D.; Van Dorsselaer, A.; Sanglier-Cianferani, S. *Anal. Chem.* **2013**, *85*, 715–736.
- (10) Cromwell, M. E. M.; Hilario, E.; Jacobson, F. *AAPS J.* **2006**, *8*, E572–E579.
- (11) Garcia-Fruitos, E.; Vazquez, E.; Gonzalez-Montalban, N.; Ferrer-Miralles, N.; Villaverde, A. *Curr. Pharm. Biotechnol.* **2011**, *12*, 1530–1536.
- (12) Berkowitz, S. A. *AAPS J.* **2006**, *8*, E590–E605.
- (13) Gabrielson, J. P.; Arthur, K. K.; Stoner, M. R.; Winn, B. C.; Kendrick, B. S.; Razinkov, V.; Svitel, J.; Jiang, Y. J.; Voelker, P. J.; Fernandes, C. A.; Ridgeway, R. *Anal. Biochem.* **2010**, *396*, 231–241.
- (14) Wen, Z.-Q. *J. Pharm. Sci.* **2007**, *96*, 2861–2878.
- (15) Abu-Absi, R.; Kenty, B. M.; Cuellar, M. E.; Borys, M. C.; Sakhamuri, S.; Strachan, D. J.; Hausladen, M. C.; Li, Z. *J. Biotechnol. Bioeng.* **2010**, *108*, 1215–1221.
- (16) Li, B.; Ryan, P. W.; Ray, B. H.; Leister, K. J.; Sirimuthu, N. M. S.; Ryder, A. G. *Biotechnol. Bioeng.* **2010**, *107*, 290–301.
- (17) Ashton, L.; Xu, Y.; Brewster, V. L.; Cowcher, D. P.; Sellick, C. A.; Dickson, A. J.; Stephens, G. M.; Goodacre, R. *Analyst* **2013**, *138*, 6977–6985.
- (18) Morita, S.; Shinzawa, H.; Noda, I.; Ozaki, Y. *Appl. Spectrosc.* **2006**, *60*, 398–406.
- (19) Noda, I.; Ozaki, Y. *Two-Dimensional Correlation Spectroscopy: Applications in Vibrational and Optical Spectroscopy*; John Wiley and Sons Ltd: Chichester, UK, 2004.
- (20) Noda, I. *Appl. Spectrosc.* **1993**, *47*, 1329–1336.
- (21) Morita, S.; Shinzawa, H.; Tsenkova, R.; Noda, I.; Ozaki, Y. *J. Mol. Struct.* **2006**, *799*, 111–120.
- (22) McCarthy, D. A.; Drake, A. F. *Mol. Immunol.* **1989**, *26*, 875–881.
- (23) Vermeer, A. W. P.; Norde, W. *Biophys. J.* **2000**, *78*, 394–404.
- (24) Vermeer, A. W. P.; Norde, W.; van Amerongen, A. *Biophys. J.* **2000**, *79*, 2150–2154.
- (25) Goto, Y.; Ichimura, N.; Hamaguchi, K. *Biochemistry* **1988**, *27*, 1670–1677.
- (26) Liang, M.; Chen, V. Y. T.; Chen, H.-L.; Chen, W. *Talanta* **2006**, *69*, 1269–1277.
- (27) Miura, T.; Takeuchi, H.; Harada, I. *J. Raman Spectrosc.* **1989**, *20*, 667–671.
- (28) Mikhonin, A. V.; Asher, S. A. *J. Am. Chem. Soc.* **2006**, *128*, 13789–13795.
- (29) Tsuboi, M.; Suzuki, M.; Overman, S. A.; Thomas, G. J. *Biochemistry* **2000**, *39*, 2677–2684.
- (30) Ashton, L.; Blanch, E. W. *J. Mol. Struct.* **2010**, *974*, 132–138.
- (31) Lord, R. C.; Yu, N.-T. *J. Mol. Biol.* **1970**, *51*, 203–213.
- (32) Kumar, S.; Rai, A. K.; Singh, V. B.; Rai, S. B. *Spectrochim. Acta Part A* **2005**, *61*, 2742–2746.
- (33) Derbel, N.; Hernandez, B.; Pfluger, F.; Liquier, J.; Geinguenaud, F.; Jaidane, N.; Lakhdar, Z. B.; Ghomi, M. *J. Phys. Chem. B* **2007**, *111*, 1470–1477.
- (34) Webster, G. T.; Dusting, J.; Balabani, S.; Blanch, E. W. *J. Phys. Chem. B* **2011**, *115*, 2617–2626.
- (35) Vohnik, S.; Hanson, C.; Tuma, R.; Fuchs, J. A.; Woodward, C.; Thomas, J. G. *J. Protein Sci.* **1998**, *7*, 193–200.
- (36) Howell, N.; Li-Chan, E. *Int. J. Food Sci. Technol.* **1996**, *31*, 439–451.
- (37) Prevelige, J. P. E.; Thomas, D.; Aubrey, K. L.; Towse, S. A.; Thomas, J. G. *J. Biochemistry* **1993**, *32*, 537–543.