The challenge of applying Raman spectroscopy to monitor recombinant antibody production

Lorna Ashton, Yun Xu, Victoria L. Brewster, David P. Cowcher, Christopher A. Sellick, Alan J. Dickson, Gill M. Stephens and Royston Goodacre

UV resonance Raman (UVRR) spectroscopy combined with chemometric techniques was investigated as a physiochemical tool for monitoring secreted recombinant antibody production in cultures of Chinese hamster ovary (CHO) cells. Due to the enhanced selectivity of the UVRR, spectral variations arising from protein, small molecule substrates, and nucleic acid medium components could be measured simultaneously and we have successfully determined antibody titre. Medium samples were taken during culture of three CHO cell lines: two antibody-producing cell lines and a non-producing cell line, and analysed by UVRR spectroscopy using an excitation laser of 244 nm. Principal component analysis (PCA) was applied to the spectral sets and showed a linear trend over time for the antibody-producing cell lines that was not observed in the non-producing cell line. Partial least squares regression (PLSR) was used to predict antibody titres, glucose utilization and lactate accumulation, and compared very favourably with gold standard data acquired with the much slower techniques of ELISA and liquid chromatography. Further analysis of the UVRR spectral sets using two-dimensional correlation moving windows also revealed that spectral variations due to protein and nucleic acid concentrations in the medium during cell culture varied between each of the three cell lines investigated.

Introduction

With increasing demand for new therapeutic treatments, including products generated from mammalian cell lines, the need for robust, reliable and effective process analytical technology (PAT) has never been more paramount. One rapidly advancing area that is being increasingly applied to PAT is spectroscopy including fluorescence, near-infrared, mid-infrared spectroscopy, and more recently Raman spectroscopy. Raman spectroscopy offers significant potential for the biopharmaceutical industry as it utilises the interaction of light with molecules to measure functional group vibrations, and provides both quantitative and qualitative information. It is non-invasive, non-destructive, rapid and requires little or no sample preparation. The chemical and structural sensitivity of Raman spectra enables multiple analytes to be measured simultaneously and, therefore, a wide number of key components present in cell culture can be determined.

Although Raman spectroscopy has successfully been used to identify and quality-assess the complex constituents of mammalian cell culture media, to date, no studies have effectively monitored antibody titre. The process of recombinant antibody production in mammalian cell cultures is extremely complex. Not only do supernatant media samples contain a large variety of components resulting in multifaceted spectral features but also each component concentration will uniquely vary throughout the antibody production process. For example, glucose levels decrease during antibody production while lactate and protein levels increase, which rules out any type of spectral background subtraction such as subtraction of the starting medium spectrum or individual component spectra (as there are too many and not all components are known). The large variety of components and their time-related changes in concentration make the identification of spectral variations purely from antibody levels challenging and it requires careful application of appropriate chemometric methods. A further complication when monitoring on-line protein production is the requirement for large numbers of frequently acquired Raman spectra if reliable and detailed data analysis is to be carried out.

Here, we demonstrate how, with the use of robust cellular (biological) controls and appropriate chemometric methods, changes in mammalian cell line supernatant composition and antibody concentration can be determined from UV resonance Raman (UVRR) spectra. Measuring Raman spectra with UV lasers (excitation wavelength 180–260 nm) has the advantage of no fluorescent background and creating a resonance Raman effect, enhancing the signal by a factor of $10^{3}–10^{5}$ over...
conventional Raman in the visible range. In resonance Raman (RR) spectroscopy, the signal is enhanced if the energy of the incident laser is within the molecular absorption bands of chromophores, or at 244 nm excitation, aromatics in the molecules. This effect gives UV resonance Raman spectroscopy the advantage of selectivity; by changing UV wavelength individual species in complex samples can be selectively enhanced. Using an excitation laser of 244 nm, which is employed in this present study, features from both proteins (aromatic amino acids) and nucleic acids can be enhanced.

We have compared UVRR spectra of medium samples collected daily from three different mammalian cell cultures; two producing an IgG4 antibody (LB01 and 60B1) and one non-producer (CL8). By applying a combination of three chemometric techniques, principal component analysis (PCA), partial least squares regression (PLSR) modelling and two-dimensional correlation analysis we have overcome some of the key challenges facing the application of Raman spectroscopy to monitor recombinant antibody production.

Methods and material

**Media and reagents**

CD-CHO was obtained from Invitrogen (Carlsbad, USA); the chemical composition is defined but confidential. All other reagents were obtained from Sigma unless otherwise stated.

**Cell lines**

Three GS-CHO cell lines, LB01, 60B1 and CL8, were kindly obtained from Lonza Biologics (Slough, UK). The GS-CHO derived suspension cell lines (LB01 and 60B1) were generated by stable transfection of CHOK1SV cells with a construct containing glutamine synthetase (GS) and a chimeric IgG4 (CB72.3). A non-producing GS-CHO cell line (CL8) was generated by transfection with the same construct lacking the antibody sequences.

**Cell culture**

Stocks of GS-CHO cell lines were revived in 20 mL CD-CHO medium (Invitrogen), supplemented with 23 μM methionine sulfoximine (MSX). All cell lines were sub-cultured every 3–4 days with a seeding density of 0.2 × 10^6 cells mL⁻¹. The cells were grown in 250 mL Erlenmeyer flasks in a volume of 50 mL medium. All cultures were grown at 37 °C with 100 rpm shaking. Growth was assessed by light microscopy using an improved Neubauer haemocytometer at 24 h intervals. Samples were appropriately diluted and mixed 1 : 1 with 0.5% Trypan Blue in PBS. Samples were collected for UVRR, ELISA and HPLC analysis on days 3, 4, 5, 6, 7 and 10 and clarified by centrifugation at 20 000×g for 1 min.

**Sandwich ELISA assay**

Secretion of IgG was monitored using a sandwich ELISA assay. Microtitre plates were coated with a Fab2 goat anti-human IgG Fc (Jackson Immunoresearch Labs). Samples and IgG standards were incubated overnight at 37 °C, and bound antibody was detected using Anti-human IgG kappa HRP conjugate (Invitrogen). The complex was visualised with TMD chromogenic substrate (Sigma), and colour development was measured at 450 nm in a plate reader. Standards of known antibody titre were included on each plate and the standard curve was used to calculate IgG concentrations in medium samples. The typical error from this standard curve was typically 11%. Values are given in Table 1.

**HPLC analysis**

Clarified medium samples were analysed by HPLC using a Shimadzu instrument (Shimadzu Corp., Nakagyo-ku, Japan), equipped with an autosampler and variable wavelength UV detector. Separation was achieved using a 300 × 7.8 mm Rezex ROA column with 4 × 3 mm Carbo-H guard column (Phenomenex Inc., Macclesfield, UK) at 60 °C and 5 mM sulfuric acid mobile phase at 0.5 mL min⁻¹ isocratic flow. Detection was by UV absorbance at 195 nm (glucose) and 215 nm (lactate). Samples were filtered using 0.2 μm Phenex RC membrane syringe filters (Phenomenex Inc., Macclesfield, UK) before injection. The error from these calibration curves was typically 8%. Values are given in Table 1.

**UVRR spectroscopy**

UVRR spectra were acquired using a Renishaw Raman 1000 system (Renishaw, Wotton-under-edge, Gloucestershire, UK). 40 μL of each clarified medium sample were pipetted into a small well on the lid of a polystyrene microplate, as shown in

| Table 1 Concentration (g L⁻¹) of antibody as determined by ELISA assay, glucose and lactate as determined by HPLC analysis for the three cell lines, LB01, 60B1 and CL8. No antibody was detected in the non-producing cell line, CL8. |
|---|---|---|---|---|---|
| Sample | LB01 | | | 60B1 | | | CL8* | |
| | Antibody | Glucose | Lactate | Antibody | Glucose | Lactate | Antibody | Glucose | Lactate |
| Day 3 | 0.04 | 3.95 | 1.24 | 0.03 | 3.89 | 1.39 | 3.45 | 1.93 |
| Day 4 | 0.06 | 3.53 | 1.66 | 0.04 | 3.47 | 1.55 | 3.19 | 1.85 |
| Day 5 | Not available | 3.05 | 1.64 | Not available | 2.87 | 1.61 | 2.86 | 1.82 |
| Day 6 | 0.27 | 2.3 | 1.75 | 0.12 | 2.03 | 1.95 | 2.38 | 1.86 |
| Day 7 | 0.49 | 0.81 | 2.53 | 0.16 | 1.39 | 2.59 | 1.7 | 1.98 |
| Day 10 | 0.58 | 0.48 | 2.97 | 0.37 | 0.61 | 2.89 | 0.77 | 2.52 |

* No antibody was detected for CL8.
Fig. 1. Approximately 0.2 mW of power were delivered to the sampling point using a Lexel Model 95 Ion Laser emitting at 244 nm. The UV laser was focused into the solution with care being taken not to focus on the lid itself. The well was continuously rotated under the laser to avoid photodegradation; and if on line analysis was used in the future the movement of the liquid within the vessel would also negate photo-damage. Spectra were collected with acquisition time of 60 s and recorded from several wells carefully monitored for signs of evaporation or degradation. Only spectra with no demonstrable photodegradation and no signal from the plate were used for analysis.

Data preprocessing and chemometric analysis

Before chemometric processing cosmic spikes were electronically removed from spectra using Grams, all further data processing and analysis was applied using MATLAB software version 2011a (The Math Works, MA, USA).

In order to compare spectra directly the data were normalised by standard normal variate (SNV) so that each spectrum has a 0 mean and standard deviation of 1. Principal components analysis (PCA) was applied to the normalised data after smoothing using a triangular sliding average.

Partial least squares regression (PLSR) models were built using the normalised spectra only to find correlation between the UVRR spectra and the concentration of the antibody, glucose and lactate. The models were validated using a double k-fold cross-validation. For each fold of the cross-validation, the samples of one concentration of the predictor were removed and used as a blind test set. The remaining samples were used as the training set, the number of PLS components were chosen by using another cross-validation on the training set only and the model built on the training set, the concentration levels of the test set were then predicted by using the model. This process is repeated until every concentration level has been used as a blind test set and predicted once.

Two-dimensional correlation moving window calculations were performed on preprocessed UVRR spectral sets using 2Dshige freeware available from http://sci-tech.ksc.kwansei.ac.jp/~ozaki/e_2D.htm and moving windows plots were generated with MATLAB software.

Results

Complexity of spectral data

Raman spectroscopy has the advantage of being sensitive to a large range of biomolecules, as demonstrated in Fig. 2a–c, where distinctive spectral features can be observed in the UVRR spectra of glucose, lactate and antibody. However, one of the key challenges of applying Raman spectroscopy to monitor recombinant protein production is that the acquired spectra contain information from these and many more chemical species (metabolites, lipids, proteins, nucleotides etc.) within the supernatant sample. Numerous features observed in Fig. 2a–c can be observed to have merged together and therefore are difficult to determine in the UVRR spectrum of supernatant sample collected on day 5 from the LB01, 60B1 and CL8 cell lines (Fig. 2d). Identifying individual features is further complicated by the presence of additional components such as host protein, DNA and RNA, all of which contribute to the final spectrum. Many chemical species are already present in the cell culture medium and therefore complicate the spectrum even before the production of antibody. Despite this complexity,

![UVRR experimental set up. A polystyrene microplate lid was placed on a turntable and 40 μL of each clarified medium sample was pipetted into the small wells onto which the UVRR objective was carefully focused. The plate was continually rotated during spectral acquisition.](image)
when IgG4 is spiked into pure CD-CHO medium, increases in peak intensity from the presence of IgG4 can be observed in the UVRR spectra to occur at $\nu/C_24$ 1240, 1350 and 1620 cm$^{-1}$ (Fig. 3) suggesting that it should be possible to monitor increasing antibody concentration alongside other changing chemical species. Note that these spikes that result in a visible change in the baseline spectra contain considerably more antibody than produced by these cell lines.

A further complication for the purposes of monitoring online protein production is the requirement of large numbers of spectra for efficient and reliable data analysis. Fig. 4 displays the three complete spectral sets used in this study with 6 spectra for each of the 6 days (36 per set). Even with this relatively small number of spectra identifying spectral variations and assigning peaks to specific supernatant components, specifically antibody concentration, is a challenge. Nevertheless, this sensitivity of Raman spectroscopy does have enormous potential benefits for PAT since the spectra do contain a vast amount of information not only about antibody production but also about processes linked to cell growth, including substrate use and waste product formation. It has been extensively reported that good analysis of complex Raman data, including biopharmaceutical related samples, requires not only visual examination of the spectra but also the application of chemometrics including PCA and PLSR.$^6,7,14$

**Principal component analysis (PCA)**

PCA is a well-established technique that reduces the dimension of multivariate data to a small number of uncorrelated variables referred to as principal components (PCs) that represent the natural variance in the data. As PCA is an unsupervised method it has the advantage of not requiring a priori information (e.g. antibody titre or sampling time) to aid in analysis. To establish if any trends within the UVRR spectral sets could be related to antibody production PCA was applied to the full Raman spectra.

Fig. 5 displays the PCA scores plots of PC1 versus PC2 for LB01, 60B1 and CL8 UVRR spectra. Although no clear separation between days or antibody levels is observed, a trend with respect to time can be readily observed for both LB01 and 60B1 spectra. A trend with increasing days and antibody concentration can be observed from left to right, along PC1 (accounting for $\sim$59.1% of total explained variance (TEV)) in the LB01 spectral set (Fig. 5a) and diagonally, bottom right to top left therefore combining PC1 and PC2 together, accounting for $\sim$72.3% TEV in 60B1 spectral set (Fig. 5b). What is not possible from these two PCA plots alone is to establish whether the observed trends, and therefore the major cause of variance in the spectral sets, is actually due to antibody concentration or
only due to a variety of time-related changes also occurring in the medium as a result of more general cell growth processes. Comparisons with the non-antibody producing cell line are therefore essential to ensure that any build up of natural host protein in the medium (rather than the recombinant antibody), or glucose consumption and metabolite waste accumulation (e.g. lactate) during growth are not mistaken for antibody. When the PCA scores plot of the non-producing cell line (Fig. 5c) is compared with the two antibody producing cell lines (Fig. 5a and b) a very different distribution can be observed with the spectral data divided into three groups, day 7, day 10 and a mix of days 3 to 6 rather than a more direct trend with days. The lack of trend with time from days 3 to 6 in the scores plot for the non-antibody producing cell line compared to the producing cell line does clearly indicate that the production of antibody is influencing trends in spectral variation, either through the increase in antibody concentration or differences the cell growth process.

The separation of days 7 and 10 from the previous days in Fig. 5c may be due to an increase in host proteins and nucleic acid content as a result of the a loss of cell viability that occurs when the IgG4 cell cultures reach day 7. Specific investigations into host cell protein dynamics in CHO cell lines using 2D-page and LC-MS determined that the majority of host protein in the supernatant is due to cell lysis or breakage associated with loss of viability. Visual examination of the spectral sets (Fig. 4) does reveal an increase in intensity from day 6 to day 10 in the bands observed at ~1478 cm\(^{-1}\). In the UVRR spectra this band is assigned to adenosine and guanine moieties and therefore the increase observed from day 6 to day 10 for all three spectral sets may be due to the release of DNA and RNA into the supernatant as a consequence of cell lysis.

The large contribution of the band at ~1478 cm\(^{-1}\) to spectral variance in the PCA score plots can also be determined from the corresponding loadings plots (Fig. 6) where it is the most dominant peak. As well as the adenosine and guanine assigned band, the wavenumber region ~1595–1645 cm\(^{-1}\) is also shown in the loadings plots to contribute strongly to spectral variance. UVRR bands in this region are extensively assigned to aromatic amino acid residues, tryptophan, tyrosine and phenylalanine.

Further peaks can be observed in the PC1 loadings plots at ~1324 cm\(^{-1}\) and in the PC2 loadings plot at ~1000 and 1188 cm\(^{-1}\) also assigned to tyrosine and phenylalanine. The UVRR peak at ~1324 cm\(^{-1}\) has also been previously assigned to adenosine and guanine moieties (Table 2). What is important to note is that the loadings plots are similar for all three cell lines, and consequently, the specific spectral regions which vary over time are regardless of antibody production. However, the different trends over time observed in the PCA scores plots (Fig. 5) does suggest that although the same peaks are changing in intensity the exact rate and extent of change over time is specific to wavenumber as well as to cell line.

**Partial least squares regression (PLSR)**

PLSR was also applied to the UVRR spectral sets to try to establish if spectral differences could be determined due to antibody production alone and not the large variety of time-related changes also occurring in the medium. Unlike PCA, PLSR is a supervised regression method where the a priori information is used to train the model to extract relevant spectral regions which best correlate with given target variables. In this study we have used target information of (i) glucose and
(ii) lactate levels quantified by HPLC, as well as (iii) IgG4 levels measured using ELISA (Table 1). These primary reference data (or gold standard data) may also contain some errors (ELISA typically 11% and HPLC 8%, vide supra) and it is important to note that as UVRR is calibrated with these the PLSR models will contain this unavoidable variability. For the two producing cell lines (LB01 and 60B1), PLS-2 models (where all three determinants were simultaneously predicted) were built first to find correlations between the UVRR spectra and the concentration of the antibody, glucose and lactate. The models were validated using a double k-fold cross validation (see Materials and methods for further information). Good correlations were observed between predicted and actual levels for antibody titre, glucose consumption and lactate accumulation, in terms of the cross-validation coefficient ($Q^2$) (Table 3).

A further PLS-2 model was built using the non-producing cell line, CL8, in an attempt to correlate the Raman spectra and the concentration of glucose and lactate using the cross-validation method. The correlations for glucose (Table 3) were lower than the antibody producing cell lines. This suggests that the PLS models of the antibody producing cell lines was not modelling general growth factors, such as changes in glucose and lactate concentrations but had extracted pertinent changes as a result of antibody production. The predicted versus actual levels (Fig. 7) are well correlated for the lower concentrations.

Table 2 Proposed assignments for the components of UVRR spectra of supernatant samples as shown in Fig. 4

<table>
<thead>
<tr>
<th>Wavenumber (cm$^{-1}$)</th>
<th>Proposed assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1615</td>
<td>Tryptophan &amp; tyrosine$^{19,20}$</td>
</tr>
<tr>
<td>1595–1645</td>
<td>Phenylalanine, tryptophan &amp; tyrosine$^{8,18,19}$</td>
</tr>
<tr>
<td>1478</td>
<td>Guanosine &amp; adenosine$^{8,18,20}$</td>
</tr>
<tr>
<td>1324</td>
<td>Tyrosine, adenosine &amp; thymidine$^{8,18}$</td>
</tr>
<tr>
<td>1188</td>
<td>Phenylalanine &amp; tyrosine$^{8,18}$</td>
</tr>
<tr>
<td>1000</td>
<td>Phenylalanine &amp; tyrosine$^{8,18}$</td>
</tr>
</tbody>
</table>

Table 3 PLS-R $Q^2$ from the leave-one day-out cross-validation. The number of the optimal latent factors is given in parentheses.

<table>
<thead>
<tr>
<th>Cell line</th>
<th>Antibody</th>
<th>Glucose</th>
<th>Lactate</th>
</tr>
</thead>
<tbody>
<tr>
<td>CL8</td>
<td>N/A</td>
<td>0.1911 (3)</td>
<td>0.4452 (2)</td>
</tr>
<tr>
<td>60B1</td>
<td>0.8650 (4)</td>
<td>0.6414 (2)</td>
<td>0.5201 (3)</td>
</tr>
<tr>
<td>LB01</td>
<td>0.8259 (2)</td>
<td>0.4315 (3)</td>
<td>0.3267 (3)</td>
</tr>
</tbody>
</table>

Fig. 6 PCA loading plots of UVRR spectral sets monitoring IgG4 fermentation. PC1 loadings are shown in black while PC2 loadings are in red for (a) LB01, (b) 60B1 and (c) CL8 UVRR spectral sets.

Fig. 7 Partial least squares regression (PLSR) prediction models. Actual concentration versus predicted concentration of antibody titre from (a) LB01 and (b) 60B1 UVRR spectral sets.
but are not as successful for higher concentrations. These higher antibody levels are measured from day 7 onwards which, as previously discussed, is also the stage at which loss in cell viability occurs and it may be that the increase in host protein is influencing the PLSR results after day 7. Variable Importance in Projection (VIP) plots of the PLS models are displayed in Fig. 8 identifying which regions of the Raman spectra have made important contribution to the correlation found by the PLS-2 models. From the plots two peaks can be observed suggesting that in the spectra of LB01 and 60B1 samples the most significant regions of change occur at ~1607 and 1478 cm\(^{-1}\). For the CL8 VIP only one large peak occurs at ~1615 cm\(^{-1}\) suggesting spectral variations around ~1615 cm\(^{-1}\) are largely as a result in growth characteristics whilst spectral variations at ~1478 cm\(^{-1}\), although visible in all spectral sets (Fig. 2), are significantly affected by antibody production.

**Two-dimensional correlation analysis**

The PCA and PLSR results indicate that different UVRR spectral variations can be determined between the antibody producing and non-producing spectral sets but what cannot be determined from either of these techniques is the relative time course of these spectral variations. A further chemometric technique that has been successfully applied to monitor perturbation-induced transitions in biomolecules is two-dimensional (2D) correlation moving windows. This approach identifies dynamic fluctuations in spectra induced by an external perturbation on the system under analysis, in this case, the effect over time of cell growth and antibody production on the protein, small molecule substrate as well as nucleic acid composition of the supernatant samples.

In a moving window calculation the spectral set is subdivided into smaller ‘windows’. For the data presented here, because of the small number of perturbation steps, a subset of 3 was applied, therefore the first window comprised of the averaged spectra from days 3, 4 and 5, the second window comprised of averaged spectra from days 4, 5 and 6 and so forth to include all spectra. For each subset a covariance or cross-product matrix was calculated (by multiplying the 2D data matrix of the subset spectra by its transpose) and the variance or diagonal values of the covariance matrix were plotted as a function of spectral wavenumber and the average translating perturbation (days). When these variance values are displayed as contour plots (Fig. 9) the dynamic changes occurring at specific wavenumbers for each cell line can be directly related to the time it occurs. For all three cell lines the groups of contours occur at very similar wavenumbers to those already identified in Fig. 8.
the PCA loadings plots and VIP plots (Fig. 4 and 6, respectively). However, what can now be determined, along the y-axis, is how these spectral variations relate to the day of production and different arrangements of contours with time can be observed for each cell line.

As with the PCA and PLSR techniques, the largest spectral changes (indicated by the largest number of clustered contours in Fig. 9) can be observed at \( \sim 1478 \text{ cm}^{-1} \), which for all three cell lines occur from day 6 to day 10. As previously discussed, this UVRR band is assigned to adenosine and guanine\(^{17,18} \) and the increase in intensity is most likely due to an increase in host DNA and RNA in the supernatant occurring with the breakdown of cell viability that occurs after day 6 in these cell cultures.\(^{15} \) Significant spectral changes are also indicated by large numbers of contours in the wavenumber region \( \sim 1595–1640 \text{ cm}^{-1} \), however in this region the distribution of contours varies between cell lines. In the CL8 moving window plot changes in this region occur at a constant rate from day 4 to day 10, whereas for the LB01 and 60B1 cell lines the division of contours into separate groups indicates different rates of change dependent on day. In the LB01 contour plot early spectral changes occur at \( \sim 1640 \text{ cm}^{-1} \) with contours centred at day 4 followed by changes in the peaks at \( \sim 1595 \) and \( 1615 \text{ cm}^{-1} \) centred at day 5. A further set of contours at \( \sim 1615 \text{ cm}^{-1} \) can be observed from day 6 to day 10. In comparison in the 60B1 moving window plot peaks at \( \sim 1595 \) and \( 1615 \text{ cm}^{-1} \) are observed to change at day 4 while the peak at \( \sim 1600 \text{ cm}^{-1} \) changes from day 6 to day 10. These results are very similar to those observed in the VIP plots (Fig. 8) suggesting the most significant spectral variations in the none-producing cell line occur in the tryptophan and tyrosine bands.\(^{19,20} \)\( \) UVRR assigned peak at \( 1615 \text{ cm}^{-1} \) while in the antibody producing cell lines further important changes are observed at \( 1595, 1600 \) and \( 1640 \text{ cm}^{-1} \) assigned to phenylalanine, tryptophan and tyrosine as well as protein secondary structure.\(^{17–19} \) Ideally for detailed interpretation of the moving window plots much smaller time steps should be measured requiring a much larger data set and care needs to be taken to not over interpret these results but even with these limitations the moving window plots do indicate different time courses for each cell line.

**Conclusions**

We have shown that despite the many challenges UVRR spectra can be used to monitor secreted recombinant antibody production in mammalian cell lines, providing information on changes in both protein and nucleic acid composition in medium samples; i.e., the growth supernatant or what is called ‘footprint’ following cell growth.\(^{24} \) Comparisons between antibody-producing cell lines and a non-producing cell lines were essential to avoid inaccurate modelling of spectral variations so that antibody specific predictions are generated. When initially examined the variations across UVRR spectral sets appeared similar for all three cell lines, however with the careful application of chemometric techniques spectral changes due to either antibody production or more general growth characteristics could be distinguished. PCA plots revealed trends occurring with antibody production that could not be observed in the spectra from the non-producing cell line where variance appeared to be dominated by the effects of cell lysis. While PLSR modelling predicted antibody production, glucose utilisation and lactate accumulation in medium samples comparisons of non-producing and antibody-producing models further indicated that the model predictions for the antibody-producing cell lines were a result of factors relating to antibody production rather than general growth factors.

One specific advantage of Raman spectroscopy is that changes in supernatant composition can be monitored with time, however the vast amount of detail in large spectral sets means that these time-related changes are difficult to monitor using standard chemometric techniques. Using 2D correlation moving windows we identified dynamic fluctuations in spectra dependent on day of cell growth that varied not only between the antibody producing and non-producing cell lines but also between the two antibody producing cell lines. Although analysis of the 2D correlation moving windows plots was limited by the small spectral sets the plots did clearly show differences in the protein small molecule substrate and nucleic acid composition of the supernatant samples. The plots also revealed how the dominating spectral variations identified in the PCA and PLSR plots at \( 1478 \text{ cm}^{-1} \) actually only occur after day 6 but the relatively large extent of change compared to other peaks (due to the sensitivity of UVRR spectroscopy to nucleic acids) overshadows spectral variations arising from protein bands. These results further highlight the complexity, and therefore care required when monitoring antibody production with spectroscopic techniques.

In conclusion, Raman spectroscopy offers huge potential for the monitoring of bioprocesses and with the development of compact, user-friendly fibre optic probes real-time, on-line applications are feasible. The acquired spectra offer a vast range of information not only of changes in media components relating growth characteristics but also for monitoring product quantity and quality; here we demonstrate this for the simultaneous measurements of antibody production, glucose consumption and lactic acid production without recourse to lengthy chromatography prior to UVRR spectroscopy. The challenge of Raman spectroscopy for PAT is to be able to extract the required information with a consistent and reliable approach, and to deploy the technique in an on-line or at-line setting. As demonstrated here, by careful application of chemometric techniques and the use of appropriate benchmarks Raman spectra can provide valuable information for bioprocess development.

**Acknowledgements**

We would like to thank the U.K. EPSRC and BBSRC and the industrial members of the Bioprocessing Research Industry Club (BRIC) for funding. We are grateful to Lonza Biologics (Slough, UK) for providing the cell lines utilized in this study.

**References**

17 Z. Q. Wen and G. J. Thomas, Biopolymers, 1997, 45, 247–256.