

# The Action of all-*trans*-Retinoic Acid (ATRA) and Synthetic Retinoid Analogues (EC19 and EC23) on Human Pluripotent Stem Cells Differentiation Investigated using Single Cell Infrared Microspectroscopy

Graeme Clemens,<sup>\*a</sup> Kevin R. Flower,<sup>b</sup> Andrew P. Henderson,<sup>c,d,e</sup> Andrew Whiting,<sup>d</sup> Stefan A. Przyborski,<sup>c,e</sup> Melody Jimenez-Hernandez,<sup>a</sup> Francis Ball,<sup>a</sup> Paul Bassan,<sup>a</sup> Gianfelice Cinque<sup>f</sup> and Peter Gardner.<sup>\*a,b</sup>

Received (in XXX, XXX) Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXXX 20XX

DOI: 10.1039/b000000x

All *trans*-retinoic acid (ATRA) is widely used to direct the differentiation of cultured stem cells. When exposed to the pluripotent human embryonal carcinoma (EC) stem cell line, TERA2.cl.SP12, ATRA induces ectoderm differentiation and the formation of neuronal cell types. We have previously generated synthetic analogues of retinoic acid (EC23 and EC19) which also induce the differentiation of EC cells. Even though EC23 and EC19 have similar chemical structures, they have differing biochemical effects in terms of EC cell differentiation. EC23 induces neuronal differentiation in a manner similar to ATRA, whereas EC19 directs the cells to form epithelial-like derivatives.

Previous MALDI-TOFMS analysis examined the response of TERA2.cl.SP12 cells after exposure to ATRA, EC23 and EC19 and further demonstrated the similarity in the effect of ATRA and EC23 activity whilst responses to EC19 were very different. In this study, we show that Fourier Transform Infrared Micro-Spectroscopy FT-IRMS coupled with appropriate scatter correction and multivariate analysis can be used as an effective tool to further investigate the differentiation of human pluripotent stem cells and monitor the alternative effects different retinoid compounds have on the induction of differentiation. FT-IRMS detected differences between cell populations as early as 3 days of compound treatment. Populations of cells treated with different retinoid compounds could easily be distinguished from one another during the early stages of cell differentiation. These data demonstrate that FT-IRMS technology can be used as a sensitive screening technique to monitor the status of the stem cell phenotype and progression of differentiation along alternative pathways in response to different compounds.

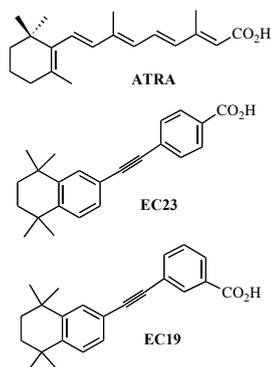
[a] Manchester Institute of Biotechnology, Manchester University, Manchester, 131 Princess Street, M1 7DN, UK; [b] Department of Chemistry, Manchester University, Manchester, UK; [c] Department of Biological Sciences, Durham University, Science Laboratories, South Road, Durham, DH1 3LE, UK; [d] Department of Chemistry, Durham University, Science Laboratories, South Road, Durham, DH1 3LE, UK; [e] Reinnervate Limited, NETPark Incubator, Thomas Wright Way, Sedgefield, TS21 3FD, UK; [f] Diamond Light Source, Diamond House, Harwell Science and Innovation Campus, Didcot, Oxfordshire, OX11 0DE UK.

## Introduction

The use of stem cells in biomedical applications is currently an area of intense interest given the potential that stem cells offer in terms of a renewable source of material for the production of differentiated human tissues. There are however a number of challenges in stem cell biology and regenerative medicine which include: (i) maintenance of the stem cell phenotype; (ii) control of cell differentiation in a robust and reproducible manner; and (iii) ensuring the purity of cell populations.<sup>1,2</sup> Monitoring the phenotype of stem cells and their differentiated progeny is therefore important. There are a number of approaches that utilise biological assays making use of biomarkers or labels to follow these processes.<sup>3,4</sup> There are issues, however, with these approaches due to a limited number of known biomarkers, or the lack of biomarkers specificity.<sup>5-7</sup> Approaches based on such biomarkers that use fluorescent or magnetic labels are therefore also shown to have drawbacks.<sup>8-10</sup>

Embryonal carcinoma (EC) cells are pluripotent stem cells derived from teracarcinomas and are considered to be the malignant counterparts of human embryonic stem (ES) cells. The EC stem cell line TERA2.cl.SP12 is one such line that has been derived,<sup>11</sup> and used as a model for cellular development in humans.<sup>12</sup> Cultures of TERA2.cl.SP12 have proven to be a robust model responsive to growth factors and compounds that interact with known biological pathways involved in the control of cell differentiation.

Naturally occurring retinoids play essential roles in a range of biological processes including embryonic development and are particularly important in the development of the central nervous system (CNS).<sup>13</sup> For example, all *trans*-retinoic acid (ATRA) is the principal active form during early embryonic CNS development. Recently, we have designed and produced some synthetic analogues of ATRA, notably 4-(5,5,8,8-tetramethyl-5,6,7,8-tetrahydronaphthalen-2-ylethynyl) benzoic acid (EC23) and 3-(5,5,8,8-tetramethyl-5,6,7,8-tetrahydronaphthalen-2-ylethynyl) benzoic acid (EC19) (Fig. 1).<sup>14</sup>



**Fig. 1** Molecular structures of the retinoid analogues ATRA, EC23 and EC19.

We have previously reported that EC23 induces neural differentiation in a manner similar to ATRA, whereas EC19 produces very few neurons and a large number of epithelial-like cells.<sup>14</sup> At present, it is not entirely clear how relatively small structural changes in these types of small molecules effect the different development processes, and this is therefore an important and developing area of study. Recent proteomic profiling (MALDI-TOF MS matrix assisted laser desorption ionisation time of flight mass spectrometry) and identification of biomarkers of the stem cell response to retinoic acid and the synthetic retinoid analogues has been described and supports the previous findings into functional relationships.<sup>15</sup> Of particular note is the fact that the results indicate that there is a subtle but significant difference in the response of TERA2.cl.SP12 cells to ATRA and synthetic compounds such as EC23, despite both molecules inducing the formation of neural cells.<sup>14</sup> There is a need to be able to study more directly the structural and chemical effects that these types of compounds can elicit and to be able to monitor these effects over time to give greater differentiation between these developmental effects as a function of their structure. Furthermore, careful characterisation of pluripotent stem cells is required given their inherent ability for differentiate into multiple tissue types and form complex teratoma tumours.<sup>16,17</sup>

This has led to the suggestion that a non-invasive method that has no or minimal impact on the cell biology is required for monitoring the status of the stem cell phenotype or progression of their differentiation. A brief review of this area and its potential has recently been published,<sup>18</sup> and a number of other studies have tackled the approach of studying stem cell differentiation in a label-free manner.<sup>19-22</sup> One such method that is currently being developed is infrared micro-spectroscopy. Infrared spectroscopy of single cells was first demonstrated by Jamin *et al.*<sup>23</sup> in the late 1990s using synchrotron radiation generated infrared as the light source. This generated considerable interest in the field but real progress was hampered by severe scattering effects that distorted the spectra of the intact cells. In a series of papers, this distortion was studied and identified as Resonant Mie scattering.<sup>24-26</sup> Importantly, it was shown that failure to remove the scattering

distortion from the spectra could lead to apparent shifts in positions of absorption bands resulting in wrongly interpreting the differences as changes in cellular biochemistry. More recently, we developed a correction algorithm to tackle this problem that enables the scattering component of the spectrum to be calculated and removed from the measured spectrum, leaving the pure absorption spectrum of the intact cell.<sup>27-28</sup> This correction algorithm has been used successfully on single cell spectra enabling different cell types, including stem-like cells, to be separated based on their infrared spectral features.<sup>29-30</sup> A review of the literature, however, shows that other infrared studies of stem cell differentiation have not been subjected to any form of scatter correction. Krafft *et al*<sup>31</sup> and Bently *et al*,<sup>32</sup> used average spectra rather than single cell spectra,<sup>33</sup> or used EMSC without taking into account resonance effects. It is possible, therefore, that spectral separation between cell phenotypes may be based in-part on changing cell morphology rather than exclusive changes in cell chemistry. In addition, recent studies have shown that single cell spectra recorded in transfection mode using low-e or metallic substrates can also be subject to distortion, resulting from the so called electric-field standing wave effect.<sup>34,35</sup> Although the use of the second derivative<sup>33,36</sup> can mitigate to some extent against this distortion, it is not fool-proof and errors in assignment/classification can still occur.<sup>35</sup> In this study, we show that FT-IRMS coupled with the RMieS-EMSC correction algorithm and multivariate analysis can be used as an effective tool to investigate the differentiation of human pluripotent stem cells, based on the spectral changes associated with changing biochemistry alone. Using this methodology, we show that it is then possible to observe subtle spectral differences induced by various different retinoid compounds that reflect real biochemical changes within the cells.

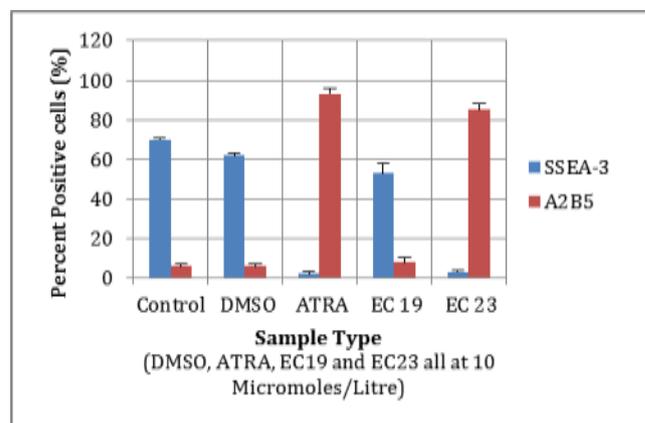
## Experimental

### Preparation and analysis of cell cultures

**Retinoid Solutions:** In preparation for use in cell culture experiments, stock solutions of synthetic retinoids (EC23, EC19) and all-*trans*-retinoic acid (ATRA) (Sigma) were prepared in DMSO (Sigma) to concentrations of 10 mM. Aliquot stock solutions were stored at  $-80^{\circ}\text{C}$  in the dark.

**Tissue Culture:** Human pluripotent TERA2.cl.SP12 embryonal carcinoma stem cells were maintained under standard laboratory conditions. In brief, cells were cultured in a humidified atmosphere of 5 %  $\text{CO}_2$  in air at  $37^{\circ}\text{C}$ , in a Sanyo  $\text{CO}_2$  incubator, cultured in DMEM (Sigma) supplemented with 10 % FCS (Gibco), 2mM L-glutamine and 100 active units each of penicillin and streptomycin (Gibco). Cultures were passaged using acid-washed glass beads (VWR) unless a single-cell suspension was required for counting, in which case a 0.25 % trypsin EDTA (Cambrex) solution was used. All cultures were handled in reduced light conditions, to account for the instability (photosensitivity) of ATRA. Cultures intended for flow cytometry and IR analysis were set up in T25 flasks (BD Falcon). After 7 days the cultures were split in a ratio of 1:9, with the smaller fraction used for analysis by flow cytometry and the larger fixed in 4% formalin solution.

**Flow Cytometry:** Flow cytometry analysis was carried out on live cells using antibodies recognising cell surface markers. The expression of markers indicative of the stem cell (SSEA-3 (University of Iowa Hybridoma Bank) or neural cell (A2B5, R&D Systems) phenotype was determined to indicate the status of cellular differentiation by TERA2.cl.SP12 cells. Suspensions of single EC cells of their differentiated derivatives were formed by the addition of 1 mL 0.25% trypsin/EDTA solution. The cell suspension was divided accordingly for flow cytometry analysis and fixing accordingly (see above). Cells were added to a 96-well plate as a suspension in wash buffer (0.1 % BSA in PBS) for incubation with primary (1:10 SSEA-3 and 1:20 A2B5) and FITC-conjugated secondary antibody IgM (Sigma, 1:100). Labelled cells were analysed in a Guave EasyCyte Plus System (Millipore) flow cytometer. Thresholds determining the numbers of positively expressing cells were set against the negative control antibody P3X.



**Fig. 2** Expression of SSEA-3 (Stem cell) and A2B5 (Neural) markers for the DMSO, ATRA, EC19 and EC23 treated TERA2.cl.SP12 cells and the untreated control cells after 7 days.

5 Analysis of the cell surface markers indicative of the status of stem cell phenotype or commitment toward differentiation were assessed by flow cytometry (Fig. 2). The data show that ATRA and EC23 induce the differentiation of TERA2.cl.SP12 EC cells by increasing the expression of the neural marker A2B5 and  
10 reducing the expression of the stem cell marker SSEA3. Cultures treated with EC19 retain a high level of SSEA3 and A2B5 expression is low indicating that these cultures remain largely undifferentiated.

## 15 FTIR sample preparation

**Cytospinning onto 1 mm CaF<sub>2</sub> plates:** All sample and control culture solutions were agitated to provide a homogeneous solution. 100 microliters of all culture solutions were then  
20 dispensed into individual labelled cytospinning cuvettes. The loaded cuvettes were spun at 950 rpm for five minutes firing a number of TERA2.cl.SP12 embryonal carcinoma stem cells onto the CaF<sub>2</sub> substrate window. The micro-slides were left to dry overnight, dipped carefully into double distilled water to remove  
25 any residual salt and left to air dry.

## Experimental set-up for FTIR

**FT-IR imaging:** The infrared data were recorded on a Varian 670 FTIR spectrometer interfaced with a Varian-620 imaging  
30 infrared microscope. The microscope was equipped with a 128 × 128 liquid nitrogen cooled MCT focal plane array detector (FPA) with a pixel effective size on the sample of 5.5 μm. The infrared spectral images were collected in transmission mode (500 scans at 4 cm<sup>-1</sup> resolution) and the background image was recorded  
35 from a clean CaF<sub>2</sub> slide.

**Synchrotron FT-IR microspectroscopy:** The synchrotron infrared microscopy was carried out using the infrared beamline (MIRIAM) at the Diamond Light source, Didcot Oxford UK.<sup>37</sup>  
40 The data were recorded on a Bruker Vertex 80V vacuum spectrometer coupled to a Hyperion 3000 infrared microscope. The infrared beam from the synchrotron was optimised by adjusting the 36x magnification optical condenser on the microscope enhancing the lower wavenumber range (1200 - 1000

45 cm<sup>-1</sup>). The aperture was set to 14 × 14 μm<sup>2</sup> such that the beam just covered a single cell on the CaF<sub>2</sub> slide. The microscope is equipped with a mid-band high sensitivity MCT detector which, via the 36x magnification objective used, has an effective area of 14×14 μm<sup>2</sup> at the sample. Both the background and sample  
50 spectra were recorded at 4 cm<sup>-1</sup> spectral resolution and 256 scans. For each biological replicate approximately 40 cell spectra were recorded.

## Data pre-processing and chemometrics

55 **Pre-processing for FT-IR imaging:** An in house function was written for and used in MATLAB that identifies individual cells within the spectral images collected from sample slides. The function also removes spectra from non-mono-layer distributed  
60 cells and single cell spectra with poor signal to noise. The spectra obtained from each pixel in a given cell area are averaged to give a single unique spectrum for that single cell. All of the mean cell spectra were subjected to the RMieS – EMSC correction algorithm (20 iterations).<sup>25-28</sup> The corrected spectra were then run  
65 through a noise reduction algorithm, vector normalised and transformed to the second derivative (11 point Savitsky-Golay smoothing, polynomial order 3). The fingerprint region 1000-1475 cm<sup>-1</sup> was analysed after mean centring using principal component analysis (PCA) and principal component linear  
70 discriminant analysis (PC-LDA).

**Pre-processing for synchrotron single point data:** All synchrotron collected data with poor signal to noise were removed and the remaining spectra corrected using the RMieS –  
75 EMSC correction algorithm (20 iterations).<sup>25-28,38</sup> The corrected spectra were then run through a noise reduction algorithm, vector normalised and transformed to the second derivative (9 point Savitsky-Golay smoothing, polynomial order 3). The fingerprint region 1000-1475 cm<sup>-1</sup> was analysed after mean centring using  
80 principal component analysis (PCA) and principal component linear discriminant analysis (PC-LDA).

**Statistical Analysis:** PCA analysis was performed on all pre-processed data to generate a new data set consisting of principle  
85 components (PCs) that contains the spectral variance present in the original data set.<sup>39</sup> These PCs representing the spectral variance, were extracted and inputted into the Linear Discriminant Analysis (LDA) supervised multivariate algorithm.

It is important to ensure that no noise is being extracted into the LDA model as over fitting can occur. To ensure noise was not a significant component in the LDA analysis, predicted residual sum of squares (PRESS) was used to estimate the predictive ability of the PC-LDA model using leave one out (LOO) cross validation.<sup>40</sup> PRESS results reveal the PCs from the PCA analysis which have started to contain noise. Also, the number of PCs selected contained less than 95 % of the overall spectral variance explained. Example scree plots of the percent variance explained per PC and PRESS results are provided in the supplementary notes. PC-LDA models have also been subjected to 10 fold cross validation to test the performance of the PC-LDA model<sup>41</sup>; see supplementary material.

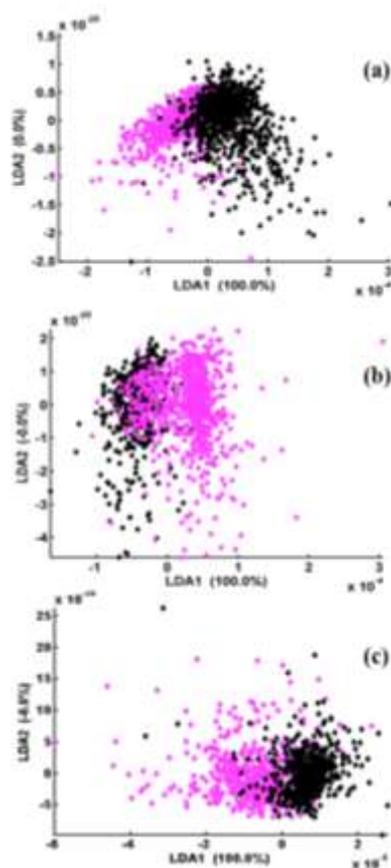
**Sample and Control reproducibility:** All Samples and controls have been studied in multiple replicates (generally triplicate) and verification for internal consistency of the control replicates was undertaken and can be found in the supplementary data.

## Results and Discussion

### Identification and justification of control

It was necessary to initially dissolve the compounds ATRA, EC23 and EC19 in the organic solvent DMSO prior to adding them to the aqueous based media solution at the desired concentration. DMSO is therefore a vehicle and it was therefore necessary to identify any spectral changes effected by DMSO itself. Several studies have shown that DMSO can induce differentiation of both ES and EC stem cells with key stem cell markers, such as Oct-4 being down regulated in gene expression tests at certain concentrations.<sup>42-44</sup> The flow cytometry results in this study show that the TERA2.cl.SP12 cells treated with DMSO at 10 mM concentration show no significant differences from the untreated TERA2.cl.SP12 cells. There was, however, a small drop in the SSEA-3 stem cell marker in the DMSO treated cells, Fig. 2, but not large enough to indicate that differentiation had begun to any significant level. This was consistent with previous work reported by Christie *et al.*<sup>14</sup> However, using PC-LDA to compare the infrared spectral fingerprints of the untreated and DMSO treated TERA2.cl.SP12 cells shows clear spectral differences are visible at 3, 5 and 7 days (Fig. 3). Accordingly, even though no apparent down regulation of the stem cell marker (SSEA-3) was observed by flow cytometry for the DMSO treated samples, there is clearly a visible cell response to the DMSO in

the infrared spectra. This difference must be considered for accurate interpretation of the spectra generated for ATRA and synthetic retinoid treated cells to dissect out the affect of the DMSO vehicle. DMSO treated samples were therefore selected as the control. For the purposes of being concise, only spectral data concerning the effects of EC23 treatment of TERA2.cl.SP12 cells over the experimental time periods (3, 5 and 7 days) with direct comparisons to the other systems, where appropriate, are shown. Full equivalent data relating to ATRA and EC19 can be found in the supplementary material.



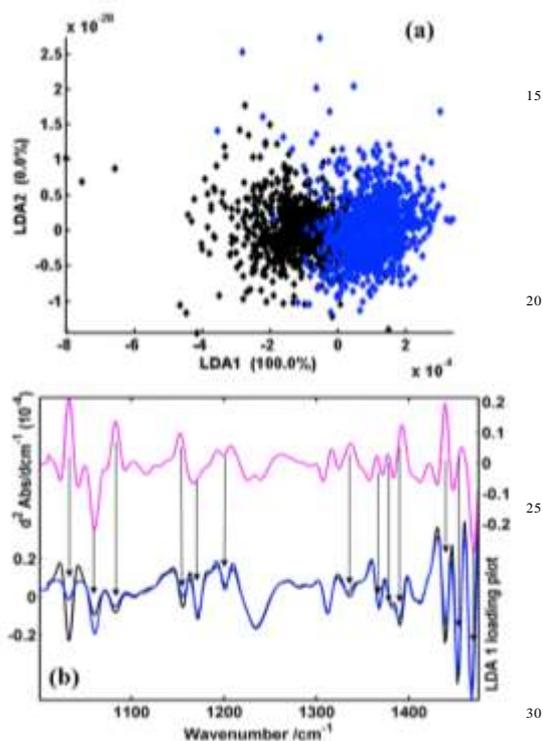
**Fig. 3** PC-LDA of control (pink) vs. DMSO (black) treated TERA2.cl.SP12 cells at days 3 (a), 5 (b) and 7 (c) respectively. (7 days 6 PC's used, 5 days 10 PC's used, 3 days 6 PC's used)

### Comparison of Control and EC23 treated TERA2.cl.SP12 cells after 7 days

Flow cytometric analysis of TERA2.cl.SP12 cells treated with EC23 for 7 days indicate induction of differentiation toward the neural lineage (Fig. 2). This is consistent with previous work of Maltman *et al.*<sup>15</sup> The PCA-LDA plot, Fig. 4a, shows the separation between the DMSO and EC23 treated TERA2.cl.SP12 cells along LDA 1 and is related to the LDA loadings plot, Fig.

4b, which identifies the wavenumbers of absorption bands responsible for the separation.

Comparison of either the mean spectra or the loadings plot for the EC23 treated TERA2.cl.SP12 cells in Fig. 4b show no obvious new spectral bands, rather subtle spectral differences that relate to changes in relative absorption. The increase in spectral absorption at  $\sim 1060\text{ cm}^{-1}$  and  $\sim 1383\text{ cm}^{-1}$  assignable to polysaccharides and glycoproteins respectively,<sup>45</sup> could reasonably represent an increase in glycoproteins<sup>46-47</sup> and also fits with the increase in absorption at  $\sim 1468\text{ cm}^{-1}$  attributed to spectral absorptions of lipids and proteins<sup>45</sup> first described by Abaskharoun *et al.*<sup>48</sup>



**Fig. 4** (a) PC-LDA of EC23 (blue) and DMSO (black) treated TERA2.cl.SP12 cells at day 7 (8 PCs used) (b) (Lower curves) mean spectra of the control and EC23 treated TERA2.cl.SP12 cells, (upper curve) LDA 1 loading plot from the EC23 vs. DMSO PC-LDA analysis at 7 days.

This increase in polysaccharide could also indicate increased proteoglycan synthesis a known cell protection process invoked during differentiation. In addition, extracellular proteoglycan ligands play a major role in extracellular signalling for differentiation.<sup>46</sup> The increase in absorption at  $\sim 1468\text{ cm}^{-1}$  and decreases at  $\sim 1392\text{ cm}^{-1}$ ,<sup>33</sup>  $1439\text{ cm}^{-1}$  and  $1454\text{ cm}^{-1}$  may indicate down-regulation of glycolipid synthesis and up-regulation of new sphingolipids and gangliosides.<sup>49</sup> This is

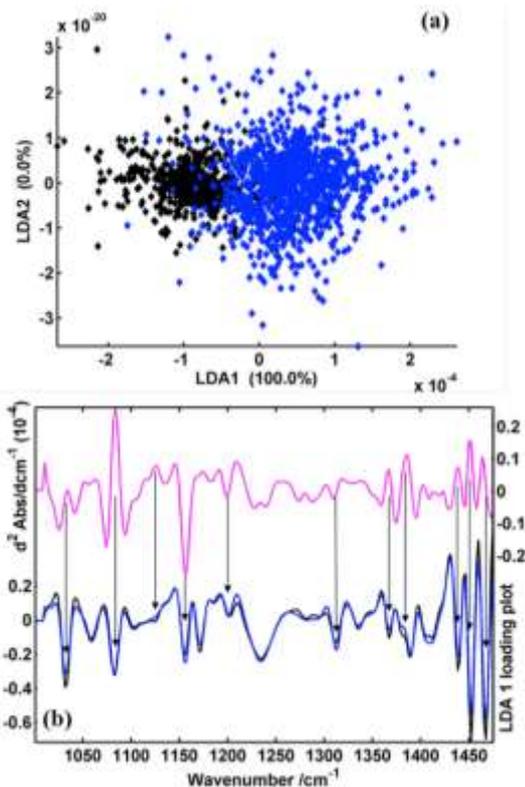
consistent with major roles identified for sphingolipids such as, membrane stabilisers, regulation of apoptosis, cell polarity, differentiation of pluripotent embryonic stem cells and part of the genome signalling pathways.<sup>49</sup> Ganglioside volume is also known to increase during neuronal differentiation of embryonal carcinoma stem cells induced by retinoid acid.<sup>50</sup> This would also agree with the up-regulation of the A2B5 (neural) marker in Fig. 2, which is a cell surface ganglioside epitope. Another significant difference is the reduced absorption at  $1032\text{ cm}^{-1}$ . This band is ascribed to glycogen<sup>45</sup> and a reduction in its intensity is indicative of a reduction in glycogen brought about by the need fuel the synthesis of new proteins, lipids and nucleic acids during the differentiation process. Reduced absorption can also be seen from the EC23 treated cells at bands  $\sim 1083\text{ cm}^{-1}$ ,  $1155\text{ cm}^{-1}$ ,  $1338\text{ cm}^{-1}$  attributed to RNA<sup>6</sup>, carbohydrate and collagen.<sup>45</sup> Changes in spectral absorption seen for the ATRA and EC19 treated TERA2.cl.SP12 cells can be seen from in the supplementary data and Table 1.

#### 65 Comparison of Control and EC23 treated TERA2.cl.SP12 cells after 5 days

Since PC-LDA of the fingerprint region allowed separation of the EC23 treated and untreated TERA2.cl.SP12 cells at day 7 the investigation was repeated at day 5 to see if: (i) spectral differences due to cell differentiation could be identified at this time point: and (ii) if cell differentiation could be seen at day 5, whether the same or different spectral features are responsible for the separation (*vide infra*); see Fig. 5. It was found, as at day 7, that separation of the EC23 treated from untreated TERA2.cl.SP12 cells could be achieved using PC-LDA, Fig. 5(a). Also see supplementary data for analogous plots for EC19 and ATRA treated TERA2.cl.SP12 cells. The main absorption bands responsible for the separation seen in Fig. 5(a) are identified from the loadings plot, Fig. 5(b).

In previous work by Heraud *et al.*<sup>33</sup> using FTIR spectroscopy to distinguish between undifferentiated hESCs and differentiating cells at 4 days, the principal spectral differences observed were reduced absorption from the fingerprint ( $1000 - 1450\text{ cm}^{-1}$ ) and lipid spectral regions ( $2830 - 2975\text{ cm}^{-1}$ ) and increased absorption from the amide spectral region ( $1500 - 1700\text{ cm}^{-1}$ ). Reduced absorption at  $\sim 1080\text{ cm}^{-1}$  was assigned to the phosphate stretching vibration of RNA<sup>33</sup> and had previously been described as important in differentiating cells by Notingher and Walsh.<sup>51-54</sup> Notingher *et al.* suggested that the decrease in RNA observed in

differentiating cells is due to the cells utilising a pool of dormant mRNA to synthesise new specific proteins required by the new phenotype.<sup>52-54</sup> TERA2.cl.SP12 cells treated with EC23 in Fig. 5, show a similar drop in absorption at  $\sim 1080-84\text{ cm}^{-1}$  along LDA 1 with increases in absorption in bands at  $\sim 1155\text{ cm}^{-1}$ ,  $\sim 1204\text{ cm}^{-1}$



25

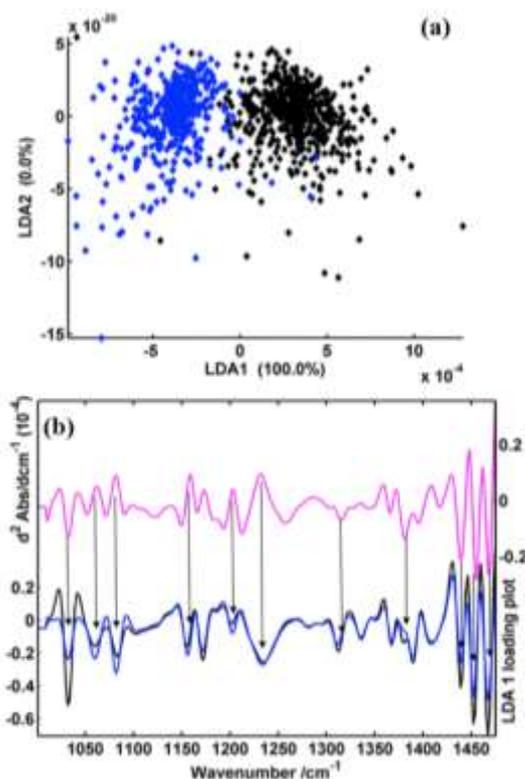
**Fig. 5** (a) PC-LDA of EC23 (blue) and DMSO (black) treated TERA2.cl.SP12 cells at day 5 (8 PCs used) (b) (Lower curves) mean spectra of the control and EC23 treated TERA2.cl.SP12 cells, (upper curve) LDA 1 loading plot from the EC23 vs. DMSO PC-LDA analysis at 5 days

assigned to carbohydrates, proteins-amide III and collagen respectively. Decreases in absorption of the EC23 treated TERA2.cl.SP12 cells assigned to glycogen, nucleic acids, lipids and proteins can also be seen,<sup>33,45</sup> see Table 1. An increase in mean protein absorption at  $\sim 1658\text{ cm}^{-1}$  is also observable for the differentiating EC23 TERA2.cl.SP12 cells when compared to the control cells, at day 5, see supplementary material. This is in agreement with the hypothesis that the differentiating EC23 treated cells at day 5 are using up their glycogen and mRNA stores in order to synthesise the new proteins needed for the new neuronal phenotype.<sup>52,54</sup> Overall spectral absorption decreases in RNA/DNA and increases in protein absorption during differentiation have also been witnessed by Ami *et al.*<sup>18</sup> in a study looking at murine embryonic stem cells. Similar spectral changes are seen from the EC19 and ATRA treated TERA2.cl.SP12 cells, see supplementary data and Table 1.

50

and  $\sim 1338\text{ cm}^{-1}$ ,<sup>45</sup> these absorption bands are

55



60

65

10

15

20

**Fig. 6 (a)** PC-LDA of EC23 (**blue**) and DMSO (**black**) treated TERA2.cl.SP12 cells at day 3 (6 PCs used) **(b)** (Lower curves) mean spectra of the control and EC 23 treated TERA2.cl.SP12 cells, (upper curve) LDA 1 loading plot from the EC23 vs. DMSO PC-LDA analysis at 3 days

### Comparison of Control and EC23 treated TERA2.cl.SP12 cells after 3 days

The PC-LDA fingerprint region allowed separation of the EC23 treated and untreated TERA2.cl.SP12 cells at day 5, therefore the investigation was repeated at day 3 to determine whether the technique is sensitive enough to identify any biochemical differences due to the introduction of the retinoids at this earlier time period (Fig. 6 (a)).

It was found, as at days 7 and 5, separation of the EC23 treated TERA2.cl.SP12 cells from the control could be achieved using PC-LDA, Fig. 6 (a). See supplementary data for analogous plots for EC19 and ATRA treated TERA2.cl.SP12 cells. The main absorption bands responsible for the separation seen in Fig. 6 (a) were identified in the loadings plot in Fig. 6 (b).

Figure 6(b) shows that the EC23 treated TERA2.cl.SP12 cells have decreased absorption bands attributable to glycogen, carbohydrates, collagen, glycoproteins, lipids and proteins and increases in band absorptions for nucleic acids (proteoglycan and RNA), carbohydrates, proteins amide and protein-amide III, see Table 1. The main spectral differences seen in the loading plot, Fig. 6 (b), tend to be above 1400 wavenumbers with the most significant spectral difference being the reduction in absorption at bands  $\sim 1439\text{ cm}^{-1}$ ,  $\sim 1452\text{ cm}^{-1}$  and  $1468\text{ cm}^{-1}$ , relating to proteins and lipids,<sup>45</sup> see Table 1. These changes may well signify the down-regulation of proteins and lipids needed for normal cell cycle proliferation. Increases in absorptions attributed to nucleic acids (proteoglycan and RNA) and proteins may also signify the synthesis of new proteins and nucleic acids needed for differentiation. Similar changes in absorption are seen for the ATRA and EC19 treated TERA2.cl.SP12 cells (See supplementary data and Table 1).

### Comparison of EC23 treated TERA2.cl.SP12 cells after 7, 5 and 3 days

All EC23 treated TERA2.cl.SP12 cells can be distinguished from control cells at 3, 5 and 7 days respectively. It is, therefore, important to determine whether the spectral changes at each time

point are the same or different. Identification of the spectral markers as the differentiation process proceeds should provide insight into the nature of the changing biochemical processes involved in differentiation and this is an area of active current investigation.<sup>31-33, 51-57</sup>

PC-LDA analysis of the EC23 treated TERA2.cl.SP12 cells at 3, 5 and 7 days shows clear separation of the three time periods (Fig. 7 (a)). Separation is seen for the ATRA TERA2.cl.SP12 cells between day 7 and days 3 and 5 thus, suggesting that between 3 and 5 days after ATRA was added to the TERA2.cl.SP12 cells, little or no biochemical change has occurred. However, PC-LDA results of the control and EC19 cells at days 3, 5 and 7 (supplementary material), show that the control and EC19 cells at days 3 and 7 are spectrally similar. This shows more of a cyclic biochemical change in the EC19 treated and control cells through time. This also coincides with the flow cytometry results in Fig. 2, which show that the DMSO and EC19 treated cells appear to have mainly remained stem cell like at 7 days. This may suggest that this cyclic biochemical change is a product of cells that are still largely in the proliferating cell cycle. Previous work has shown that the EC19 synthetic retinoid compound has a much slower rate to which cells are committed to differentiate than ATRA and EC23.<sup>14</sup> Results in Fig.2 and the PC-LDA results of the EC19 cells at days 3, 5 and 7, would be in agreement with this and would suggest that many of the cells have yet to fully commit to differentiation at 7 days. However, a cell reaction to the introduction of the EC19 compound at all time periods has been seen when compared to the control cells, see supplementary information, which suggests that the cells have started on their differentiation pathway to form epithelial cells but have yet to fully commit to differentiation.

Being able to distinguish between the treated TERA2.cl.SP12 cells and the different spectral absorption changes seen at each time point shows that the spectral changes of the treated TERA2.cl.SP12 cells are different as the differentiation process progresses. By monitoring these spectral changes it will be possible to identify the spectral markers as the differentiation process proceeds. However, later and early time periods may be needed to fully investigate the differentiation process of this cell line. The spectral absorption increases in Fig. 7(b) for the EC23 treated cells at day 7, are assignable to nucleic acids, carbohydrates, glycoproteins and proteins and lipids, whereas the decreases in absorbance's are assigned to glycogen, RNA, carbohydrates, protein-amide III and lipids and proteins, see

---

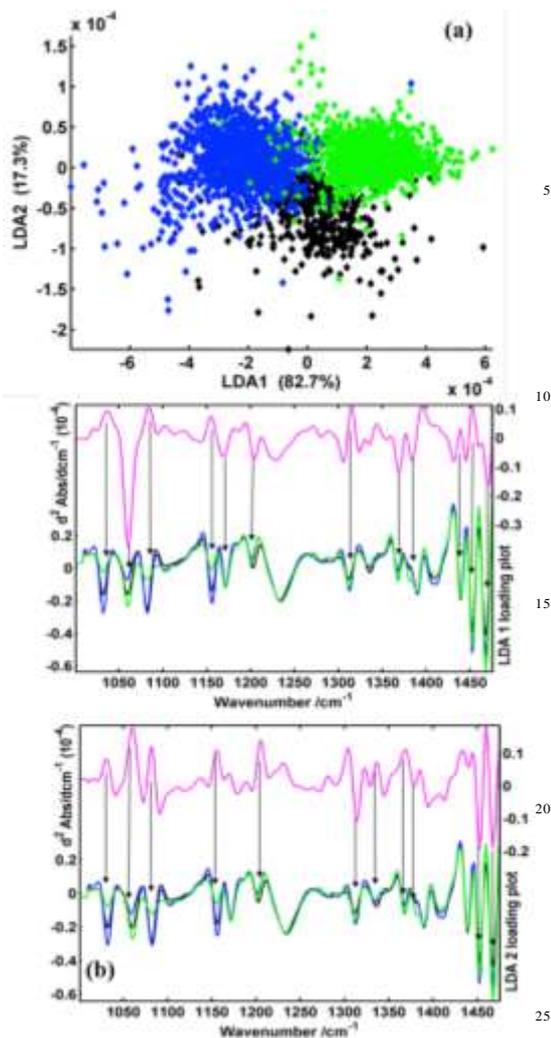
Table 1 for absorption band assignment. For LDA 2, increases in spectral absorbance can be seen for the EC23 treated cells at day 3 attributed to glycogen, nucleic acids, carbohydrates, collagen and amide-amide III, see Table 1.<sup>45</sup> Decreases in spectral absorption along LDA 2 can be assigned to protein-amide III at ~1311 cm<sup>-1</sup>, glycoprotein, lipids and proteins, see Table 1.<sup>45</sup>

From the mean spectral differences in Fig. 7 (a) and (b), it can be seen that at ~1338 cm<sup>-1</sup>, ~1383 cm<sup>-1</sup> and ~1468 cm<sup>-1</sup>, there are changes in absorbance with respect to time attributable to changes in collagen, glycoprotein, protein and lipids.<sup>45</sup> Changes in absorption with respect to time can also be seen at ~1468 cm<sup>-1</sup>, ~1454 cm<sup>-1</sup>, ~1439 cm<sup>-1</sup> and ~1383 cm<sup>-1</sup> from the ATRA treated TERA2.cl.SP12 cells. However, the control and EC19 treated

TERA2.cl.SP12 cells only show absorption changes with respect to time at ~1032 cm<sup>-1</sup> and 1236 cm<sup>-1</sup> respectively. The increase at ~1468 cm<sup>-1</sup> is attributed to absorption of proteins and lipids<sup>45</sup> and suggests that the ATRA and EC23 treated TERA2.cl.SP12 cells are up-regulating new proteins and lipids needed to form the new differentiated phenotype. This increase in absorption at ~1468 cm<sup>-1</sup> could therefore be a spectral biomarker of neuronal stem cell differentiation for this cell line. The increase at ~1236 cm<sup>-1</sup> for the EC19 treated TERA2.cl.SP12 cells could also indicate the first signs of new proteins being synthesised for epithelial differentiation and the spectral variability of the control cells may be a confirmation of the cells being at different stages of mitosis.

**Table 1. Assigning spectral bands to the loading plot weights**

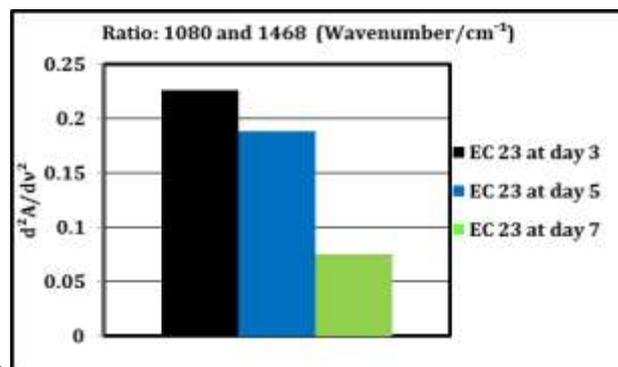
	EC 23		EC 19		ATRA		ATRA vs. EC 23		ATRA vs. EC 19		EC 23 vs. EC 19		(>) increase spectral absorption
	LDA 1		LDA 1		LDA 1		LDA 1		LDA 1		LDA 1		(<) decrease in spectral absorption
	> EC 23	< EC 23	> EC 19	< EC 19	> ATRA	< ATRA	> EC 23	< EC 23	> ATRA	< ATRA	> EC 23	< EC 23	Band Assignments (cm <sup>-1</sup> )
<b>Day 7</b>	~1061	~1032	~1061	~1034	~1084	~1032	~1061	~1082	~1084	~1059	~1061	~1084	<b>1025:</b> glycogen. <sup>45</sup> <b>1032:</b> glycofen and one of the triad peaks of nucleic acids. <sup>45, 66</sup> <b>1051:</b> C-O-C stretching of DNA and RNA. <sup>45</sup> <b>1062:</b> polysaccharide-cellulose and one of the triad peaks of nucleic acids. <sup>45</sup> Ribose RNA <sup>64</sup> and proteoglycans. <sup>50</sup> <b>1068:</b> Stretching C-O ribose. <sup>45</sup> <b>1084:</b> Stretching PO <sub>2</sub> <sup>-</sup> symmetric of DNA and RNA. <sup>6, 45, 66</sup> <b>1117:</b> C-O stretching vibration of C-OH group of ribose (RNA). <sup>45</sup> <b>1126:</b> DNA/RNA. <sup>68</sup> <b>1155:</b> C-O stretching vibration from carbohydrates. <sup>45</sup>
	~1117	~1084	~1082	~1155	~1155	~1063	~1117	~1153	~1180	~1117	~1313	~1155	<b>1173:</b> Non-hydrogen-bonded stretching mode of C-OH groups (carbohydrates). <sup>45</sup> <b>1204:</b> Vibrational modes of collagen proteins-amide III and polysaccharides. <sup>45</sup> <b>1236:</b> Amide III band PO <sub>2</sub> <sup>-</sup> of nucleic acids. <sup>45</sup> <b>1312:</b> Amide III band components of proteins. <sup>45</sup> <b>1338:</b> CH <sub>2</sub> wagging of collagen. <sup>45, 67</sup> <b>1367:</b> Stretching C-O, deformation C-H, deformation N-H. <sup>45</sup> <b>1380:</b> Stretching C-O, deformation C-H, deformation N-H. <sup>45</sup> <b>1385:</b> CH <sub>3</sub> stretching <sup>45</sup> and glycoproteins. <sup>53</sup> <b>1390:</b> Methyl, methylene and carboxylate groups of lipids. <sup>33</sup> <b>1396:</b> Symmetric CH <sub>3</sub> bending of the methyl groups of proteins. <sup>45</sup> <b>1404:</b> CH <sub>3</sub> asymmetric deformation <b>1412-1414:</b> Stretching C-N, deformation N-H, deformation C-H. <sup>45</sup> <b>1419:</b> Deformation C-H. <sup>45</sup> <b>1444:</b> Changes in lipids and fatty acids. <sup>45</sup> <b>1451:</b> Asymmetric CH <sub>3</sub> bending modes of the methyl groups of proteins. <sup>45</sup> <b>1454:</b> CH <sub>3</sub> scissoring of mainly lipids and asymmetric methyl deformation. <sup>45, 68</sup> <b>1458:</b> CH <sub>3</sub> of collagen. <sup>45</sup> <b>1465-1468:</b> CH <sub>2</sub> bending vibration (lipids and proteins). <sup>45</sup> <b>1470:</b> CH <sub>2</sub> bending of the methylene chains in lipids. <sup>45</sup>
	~1165	~1153	~1204	~1304	~1201	~1169	~1171	~1201	~1201	~1155	~1381	~1171	
	~1308	~1338	~1234	~1384	~1234	~1367	~1367	~1338	~1394	~1171	~1404	~1203	
	~1383	~1367		~1442	~1315	~1381	~1381	~1396	~1452	~1232	~1412	~1338	
	~1469	~1392		~1468		~1439	~1450		~1466	~1367	~1450	~1390	
		~1439				~1452	~1468			~1385	~1468	~1419	
		~1456				~1468				~1439		~1439	
<b>Day 5</b>	~1026	~1034	~1026	~1051	~1157	~1032	~1026	~1061	~1051	~1030	~1119	~1032	
	~1155	~1051	~1061	~1068	~1201	~1051	~1155	~1084	~1066	~1061	~1173	~1061	
	~1200	~1065	~1155	~1124	~1410	~1065	~1196	~1173	~1084	~1076	~1194	~1153	
	~1311	~1084	~1203	~1234		~1084	~1315	1203	~1126	~1155	~1313	~1203	
	~1472	~1126	~1335	~1369		~1234	~1448	~1336	~1173	~1201	~1404	~1336	
		~1367	~1392	~1383		~1365		~1367	~1236	~1331	~1410	~1367	
		~1387	~1439	~1412		~1381		~1387	~1387	~1365	~1450	~1379	
		~1439		~1450		~1437		~1456	~1452	~1381		~1392	
		~1452		~1469		~1452			~1469	~1394		~1441	
		~1466				~1466				~1441		~1469	
<b>Day 3</b>	~1063	~1032	~1061	~1032	~1059	~1034	~1159	~1032	~1032	~1061	~1084	~1061	
	~1082	~1315	~1201	~1051	~1082	~1051	~1173	~1194	~1051	~1200	~1155	~1198	
	~1159	~1365	~1076	~1084	~1157	~1367	~1201	~1381	~1082	~1232	~1173	~1315	
	~1173	~1381		~1151	~1201	~1385	~1232	~1394	~1153	~1313	~1205	~1379	
	~1203	~1439		~1336	~1232	~1468	~1410	~1439	~1336		~1232	~1394	
	~1232	~1456		~1367	~1375			~1456	~1381		~1336	~1458	
	~1410	~1468		~1381				~1468	~1389		~1369	~1473	
				~1390					~1410		~1410		
				~1439					~1439		~1450		
				~1452					1452		~1466		
				~1468					~1468				



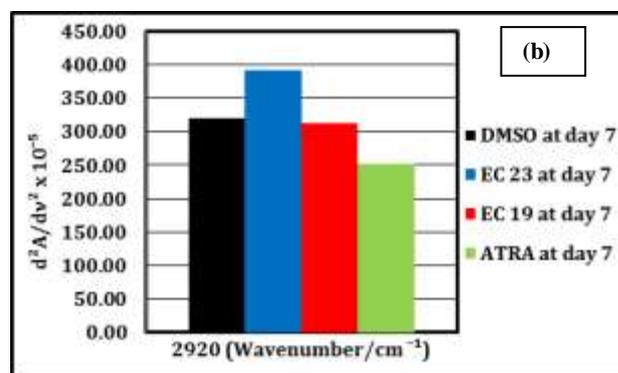
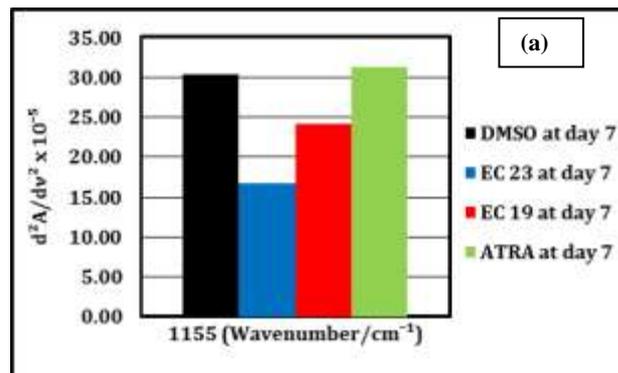
**Fig. 7** (a) PC-LDA of EC23 treated cells at 3 (black), 5 (blue) and 7 (green) days, 8 PCs used (b) Mean spectra of the EC23 treated TERA2.cl.SP12 cells at days 3, 5 and 7 (bottom) and the LDA 1 loading plot from the PC-LDA analysis of EC23 at days 3, 5 and 7 (top) (c) LDA 2 loading plot from the PC-LDA analysis of EC23 at days 3, 5 and 7.

Previously it has been suggested that embryonic stem cells have a higher nucleus to cytoplasm ratio than that of differentiating cells.<sup>18</sup> This is because it is believed that as the cells differentiate the nucleus shrinks causing a reduction in nucleic acid absorption and an increase in protein and lipid absorption from the larger cytoplasm.<sup>18</sup> To see if the TERA2.cl.SP12 cells follow a similar pattern, the nucleus to cytoplasm ratio was calculated using the ratio of the 1080 (nucleus) to 1468 (cytoplasm) bands Fig. 8. Results in Fig. 8, are consistent with previous work by Notingher *et al.*<sup>18, 52-54</sup> however, the nucleus to cytoplasm results for the EC19, ATRA treated and control TERA2.cl.SP12 cells are not, see supplementary data. Heraud *et al.* has also looked at specific spectral absorbance bands in order to try and explain

stem cell differentiation. Work by Heraud showed that hESC have increased carbohydrate and lipid absorptions at bands  $\sim 1155 \text{ cm}^{-1}$  and  $\sim 2920 \text{ cm}^{-1}$  respectively when compared to that of differentiating cell populations.<sup>33</sup> In Fig. 9, it can be seen that the EC23 treated cells are in agreement with the results seen by Heraud for the carbohydrate absorption band at  $\sim 1155 \text{ cm}^{-1}$ , however, there was an increase in absorption at  $\sim 2920 \text{ cm}^{-1}$  when compared to that of the control cells. In the case of the ATRA treated TERA2.cl.SP12 cells, Fig 9(b) there's a slight increase in absorption at  $\sim 1155 \text{ cm}^{-1}$  and a decrease at  $\sim 2920 \text{ cm}^{-1}$  when compared to the control cells in agreement with the results seen by Heraud thus, highlighting a possible difference in biochemical response of the TERA2.cl.SP12 cells when ATRA and EC23 are introduced for 7 days.



**Fig. 8** Bar graph representing the nucleus to cytoplasm ratio for the EC23 treated TERA2.cl.SP12 cells at days 3, 5 and 7



**Fig. 9 (a)** Bar graph representing the mean absorption at band  $1155\text{ cm}^{-1}$  for the control cells and all retinoid treated TERA2.cl.SP12 cells at day 7  
**(b)** Bar graph representing the mean absorption at band  $2920\text{ cm}^{-1}$  for the control cells and all retinoid treated TERA2.cl.SP12 cells at day 7.

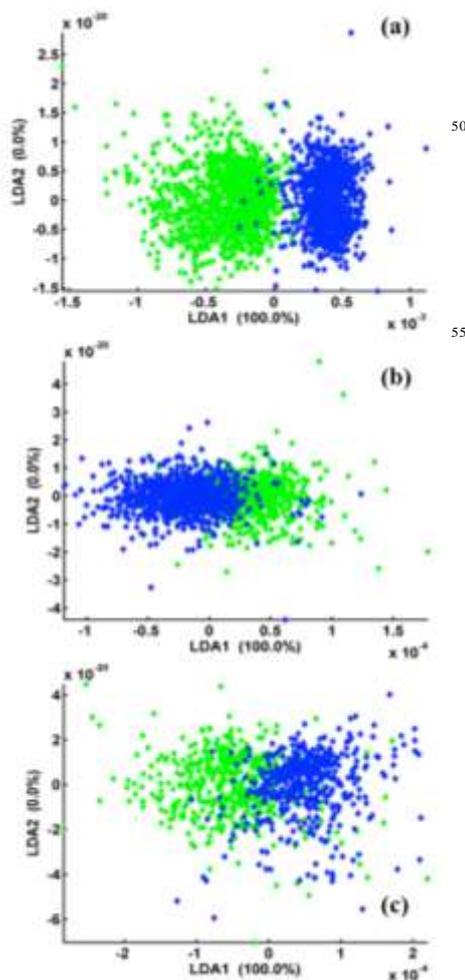
### Comparison of EC23 and ATRA treated TERA2.cl.SP12 cells

Previous work<sup>15</sup> has shown that the EC23 and ATRA treated TERA2.cl.SP12 cells produce very similar protein profiles during the cell differentiation process and therefore, a close functional relationship. However, subtle but significant difference in the response of the TERA2.cl.SP12 cells to both ATRA and EC23 was seen, primarily the cellular retinoic acid binding protein 1 (CRABP1).<sup>15</sup> Christie *et al.* also looked at the similar biological effects the two retinoids had on the TERA2.cl.SP12 cells and their study showed that the EC23 retinoid induced the production of neural tissue at a similar or possibly increased level to that seen when using ATRA.<sup>14</sup> The cells treated with EC23 and ATRA at 3, 5 and 7 days were compared to see if these subtle differences could be also identified in the infrared spectral fingerprints, Fig. 10.

Even though both ATRA and EC23 are known to induce cell differentiation to form the same cell type, the FT-IR results, Fig. 10, show that there are measurable spectral differences between the EC23 and ATRA treated TERA2.cl.SP12 cells at the same relative time points and this is consistent with previous observations seen using other techniques.<sup>14-15</sup> This could result from a number of possibilities: (i) the retinoid treated TERA2.cl.SP12 cells follow different differentiation pathways: or, (ii) the retinoid treated TERA2.cl.SP12 are following the same differentiation pathway but staggered relative to each other due to the differences in each initiators ability to start the differentiation process.

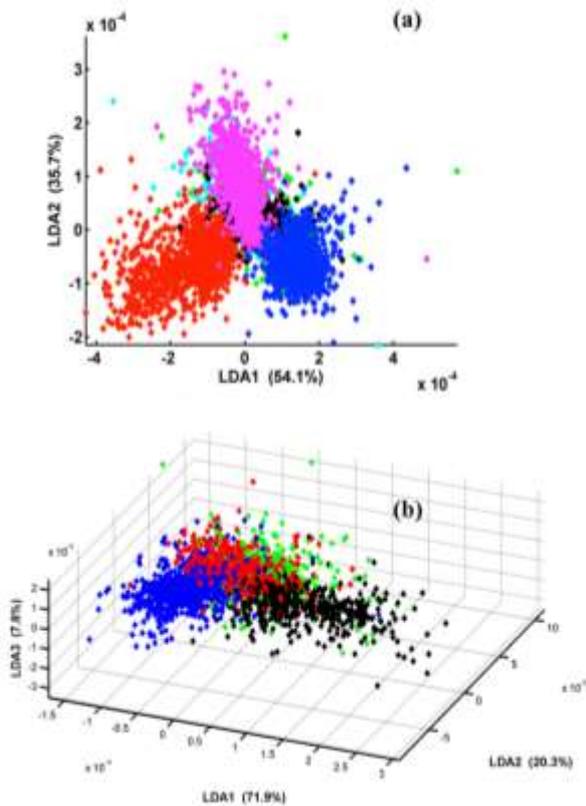
To test these possibilities, ATRA and EC23 treated TERA2.cl.SP12 cells were compared at all time periods, Fig. 11 (a). The score plot results in Fig. 11 (a) appear to show that the ATRA and the EC23 treated cells at days 3 and 5 are spectrally similar as they have both merged into a single cluster, whereas at day 7, there are significant spectral differences between the ATRA and EC23 treated cells. However, it has been shown in Fig. 10, that there are spectral differences between the ATRA and EC23 treated cells at both 3 and 5 days. Therefore, this cluster formation of the treated cells at days 3 and 5 may be a result of the spectral differences at day 7 being large and the spectral differences at days 5 and 3 being subtle, but still significant.

Spectral differences between the ATRA and EC23 treated TERA2.cl.SP12 cells can be seen in Fig.12 (a) and (b).



**Fig. 10** PC-LDA of EC23 (blue) and ATRA (green) treated TERA2.cl.SP12 cells at (a) 7, (b) 5 and (c) 3 days (7 days 6 PC's used, 5 days 10 PC's used, 3 days 6 PCs used).

The score plot results in Fig. 11 (b) show the ATRA treated cells at days 3 and 5 to be spectrally similar; this is in agreement with the PC-LDA results of the ATRA treated cells at days 3, 5 and 7 days found in the supplementary material. The EC23 treated cells at days 3 and 5 can be separated from the ATRA cell cluster. Coupling this with the fact that the ATRA treated cells at days 3 and 5 have shown reduced biochemical differences when compared to the control cells, see supplementary material, and the ATRA treated cells at days 3 and 5 are spectrally similar, this may suggest that EC23 induces differentiation of the TERA2.cl.SP12 cells to begin more rapidly than ATRA and is subsequently further on down the differentiation pathway at days 3 and 5 than the ATRA treated at the same time periods.

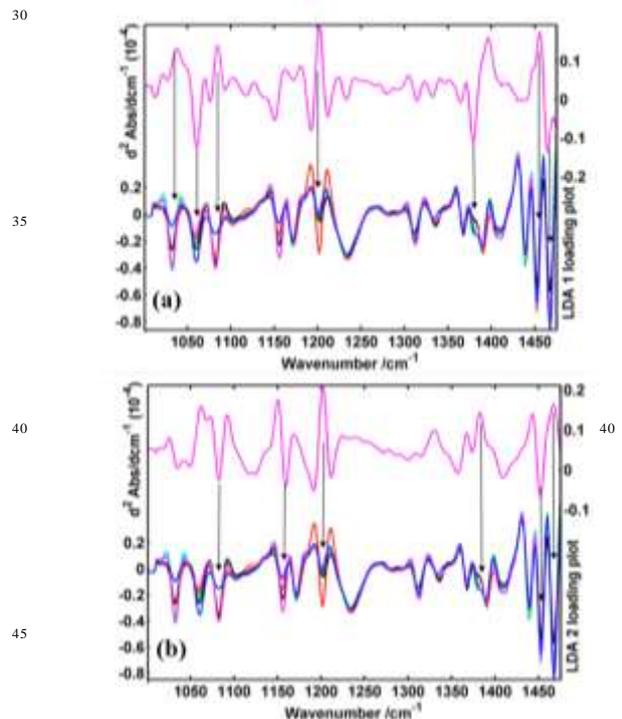


**Fig. 11** (a) PC-LDA of EC23 vs. ATRA at days 3, 5 and 7 (EC23 cells: day 3 (black), day 5 (pink), day 7 (blue), ATRA cells: day 3 (green), day 5 (cyan) and day 7 (red) (14 PCs used)) (b) PC-LDA of EC23 vs. ATRA at days 3 and 5 days (EC23 cells: day 3 (black), day 5 (blue), ATRA cells: day 3 (green), day 5 (red) (15 PCs used).

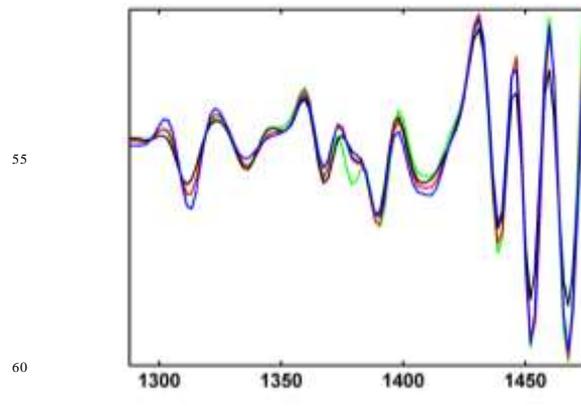
At day 7 though, there are large spectral differences seen between the differently treated TERA2.cl.SP12 cells in both Fig 10 and 11. Results in Fig. 10 and 11 could therefore be confirmation that at 7 days, EC23 induces a different differentiation pathway altogether to that of ATRA, this would be in agreement with results seen in Fig. 9. However, one can also not rule out that both ATRA and EC23 are both inducing the differentiation of the TERA2.cl.SP12 cells down the same biological pathway, but slightly staggered due to their ability in starting the differentiation process. In future studies the spectral changes of the ATRA and EC23 treated cells will be monitored over a longer time period to confirm whether EC23 induces the TERA2.cl.SP12 cells down a different biological pathway when compared to the natural retinoid ATRA.

LDA 1, Fig. 12 (a), separates the EC23 treated TERA2.cl.SP12 cells at day 7 from the ATRA treated cells at day 7. The primary spectral differences can be attributable to absorption increases for

the EC23 treated cells at day 7 in polysaccharide-cellulose, proteoglycan and proteins and lipids and decreases in glycogen, RNA, protein-amide III and lipid absorption. LDA 2, Fig. 12 (b), separates the ATRA and EC 23 treated cells at day 7 from the treated cells at days 5 and 3.



**Fig. 12** (a) Mean spectra and LDA 1 loading plot of the EC23 and ATRA treated cells at 3, 5 and 7 days (b) Mean spectra and LDA 2 loading plot of the EC23 and ATRA treated cells at 3, 5 and 7 days.



**Fig. 13** Mean spectra of EC23 and ATRA at 3 and 5 days (EC23 cells: day 3 (black), day 5 (blue), ATRA cells: day 3 (green), day 5 (red).

Here the main spectral differences can be attributable to a decrease in absorption for the treated cells at day 7 in RNA, carbohydrate and lipid absorption and increases in protein-amide III proteoglycan and protein lipid (Table 1). This may again suggest that as the ATRA and EC 23 treated TERA2.cl.SP12

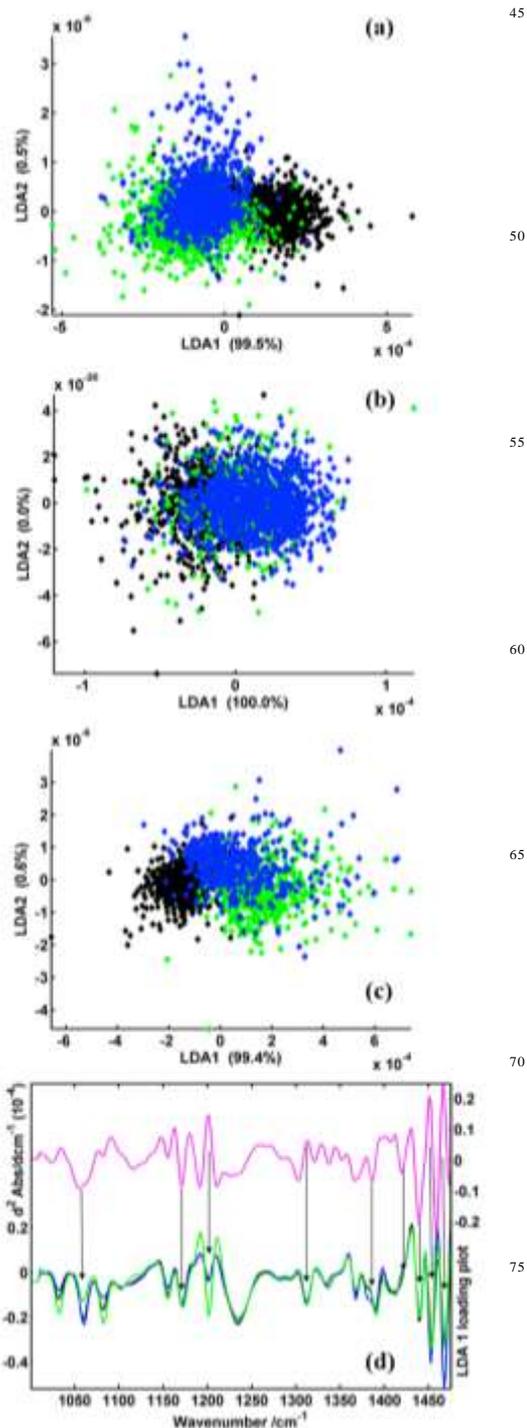
cells start to differentiate they use up their mRNA stores in order to synthesise new specific proteins required by the new neuronal phenotype being produced. The spectral absorption increase at  $\sim 1204\text{ cm}^{-1}$  attributed to proteins-amide III,<sup>45</sup> may also be a spectral biomarker of neural differentiation for this cell line, along with the increase in absorption at  $1468\text{ cm}^{-1}$ . If this were the case then it would suggest that the ATRA treated cells were a step further on in the differentiation process than the EC23 treated TERA2.cl.SP12 cells at the day 7 time period. By looking at the data from both the ATRA and EC23 treated TERA2.cl.SP12 cells at later time periods these spectral biomarkers could be confirmed.

Interestingly, when comparing the mean spectra of both ATRA and EC23 at days 3 and 5, Fig. 13, there was a mean spectral reduction in absorption at band  $\sim 1383\text{ cm}^{-1}$  attributed to glycoproteins<sup>32</sup> from all treated cells except from the ATRA cells at day 3. The score plot results in Fig. 11 (b) suggest that the ATRA treated cells at day 3 have the least spectral change therefore, this reduced absorption at  $\sim 1383\text{ cm}^{-1}$  could signify the start of neuronal differentiation. Also in Fig. 13, increases in spectral absorption with respect to time can be seen at  $\sim 1410\text{ cm}^{-1}$  attributed to stretching C-N, deformation N-H, deformation C-H.<sup>45</sup>

### 25 Comparison of the EC23 and ATRA treated TERA2.cl.SP12 cells with the EC19 treated TERA2.cl.SP12 cells

FT-IRMS coupled with PC-LDA has been able to discriminate between the ATRA and EC 23 treated cells showing that although they both differentiate EC cells to form neurons, there are significant differences in the two differentiating pathways. However, as both of these retinoids induce the differentiation of EC cells to produce neurons, it was necessary to test whether there is similarity between the ATRA and EC23 treated TERA.cl.SP12 cells. Further, can the neuronal pathway be distinguished from epithelial differentiation? To investigate further, the data from the ATRA and EC23 cells were classed as the same sample group and compared with the data from the EC19 treated cells using PC-LDA. PC-LDA analysis was performed for all time periods (Fig. 14 (a)).

Epithelial and neural differentiation can be classified after 7 days using spectral IR fingerprint (Fig. 14(a)). There was also reasonable classification at 5 and 3 days consequently suggesting that neural and epithelial differentiation may have started to begin as early as 3 days.



80 **Fig. 14** PC-LDA analysis of ATRA, EC23 and EC19 treated TERA2.cl.SP12 cells where two groups are stipulated at 7 (a), 5 (b) and 3 (c) days (10 PCs extracted from the PCA analysis for 7 days, 8 PCs used for 5 days and 10 PCs used for 3 days). EC19 (black) cells epithelial differentiation and (blue and green) EC23 treated cells and ATRA (d) Mean spectra of the ATRA, EC23 and EC19 treated TERA2.cl.SP12 cells and the LDA 1 loading plot from the neuronal vs. epithelial PC-LDA analysis at 7 days.

Therefore, even though there are spectral differences seen between the ATRA and EC23 treated cells at each time period, there are still significant similarities between the two differently treated TERA2.cl.SP12 cells. The absorption bands that contributed the most to the separation seen could potentially be spectral biomarkers of both neuronal and epithelial differentiation. The mean spectra and LDA 1 loading plot from the day 7 treated cells can be seen in Fig.14 (b). LDA 1, shows a reduction in absorbance for the neuronal differentiating cells in proteoglycan, carbohydrate, stretching C-O, deformation C-H, deformation N-H, methyl, methylene and carboxylate groups of lipids, deformation C-H and from changes in lipids and fatty acids at  $\sim 1439\text{ cm}^{-1}$ ,<sup>33,45,50</sup> (Table 1). Increases in absorption intensity of the neuronal differentiating cells can be seen from proteins-amide III and polysaccharides, proteins-amide III at  $\sim 1312\text{ cm}^{-1}$ ,  $\text{CH}_3$  scissoring of mainly lipids and asymmetric methyl deformation lipids and from proteins and lipids at  $\sim 1468\text{ cm}^{-1}$ .<sup>45</sup> The spectral differences shown in Fig. 14 (b) may be an indication that the neuronal differentiating cells are using up carbohydrate stores in order to synthesise the new proteins needed for the neuronal phenotype, also lipids and proteoglycan absorption is falling which again may indicate the down-regulation of glycolipid synthesis and up-regulation of new sphingolipids and gangliosides.<sup>49</sup> Interestingly, both the ATRA and EC23 treated TERA2.cl.SP12 cells at day 7 have increased absorptions at bands  $\sim 1204\text{ cm}^{-1}$  and  $\sim 1468\text{ cm}^{-1}$  when compared to the EC19 treated cells, this is in agreement with the results shown in Fig.12 thus, suggesting that these absorption bands may be potential biomarkers of neuronal differentiation.

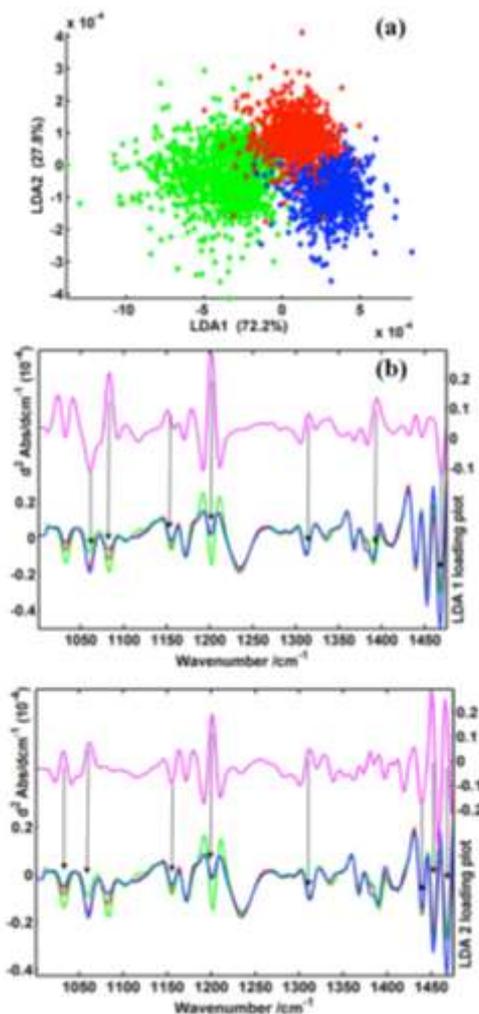
### Comparison of the EC23, EC19 and ATRA treated TERA2.cl.SP12 cells at 7 days

Previous research has demonstrated that the EC19 treated TERA2.cl.SP12 cells not only produce epithelial cells, but also a small number of neuronal cells when inducing differentiation.<sup>14</sup>

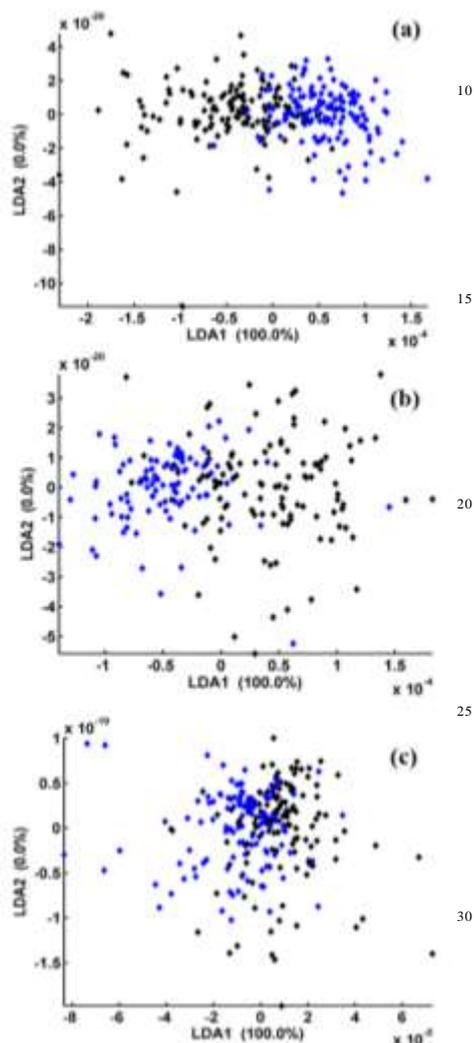
To see if the FT-IR data confirms this, PC-LDA analysis was performed of the ATRA, EC23 and EC19 treated cells at 7 days, where three individual group classes are stipulated (Fig. 15 (a)). The score plot results in Fig. 15 (a), tends to support this independently verified observation.

The data results in Fig. 15 (a), may also suggest that the EC23 synthetic retinoid is capable of inducing the differentiation of a proportion of TERA2.cl.SP12 cells down an epithelial route. This is also in agreement with the score plot results from EC19 vs.

EC23 at day 7, see supplementary notes. In LDA 1 Fig. 15(b), increases in spectral absorption intensity can be seen for the ATRA treated cells at day 7, these can be attributed to RNA, carbohydrate at  $\sim 1155\text{ cm}^{-1}$ , proteins-amide III at  $\sim 1204\text{ cm}^{-1}$  and proteins-amide III at  $\sim 1312\text{ cm}^{-1}$ .<sup>6,45</sup> Decreases in absorbance can be seen from absorption bands attributed to proteoglycan at  $1061\text{ cm}^{-1}$  and lipids and proteins<sup>45,50</sup> (Table 1). It's clear from LDA 1, Fig. 15 (b), that the main spectral differences between the ATRA and the EC19/EC23 treated TERA2.cl.SP12 cells is the large increases in absorption at  $\sim 1204\text{ cm}^{-1}$  and  $\sim 1084\text{ cm}^{-1}$  from the ATRA treated cells, associated with proteins-amide III and RNA respectively. Results from LDA 1 are similar to those seen in Fig. 12(a) where the ATRA and EC23 treated cells were characterised at day 7. LDA 2 separates the neuronal from the epithelial differentiating cells and not surprisingly the loading plot is very similar to the loading plot seen in Fig. 14(b); Comparisons of ATRA, EC23 and EC19 at each time period have been undertaken and can be found in the supplementary data.



**Fig. 15** (a) PC-LDA analysis of ATRA (green), EC23 (blue) and EC19 (red) treated cells where three groups are stipulated at day 7 (8 PCs used) (b) Mean spectra of ATRA, EC19 and EC23 treated TERA2.cl.SP12 cells and LDA 1 loading plot. Separation along LDA 1 separates the ATRA and EC 23 treated cells (c) Mean spectra of ATRA, EC19 and EC23 treated TERA2.cl.SP12 cells and LDA 2 loading plot. Separation along LDA 2 separates the ATRA and EC23 cells from the EC19 treated cells



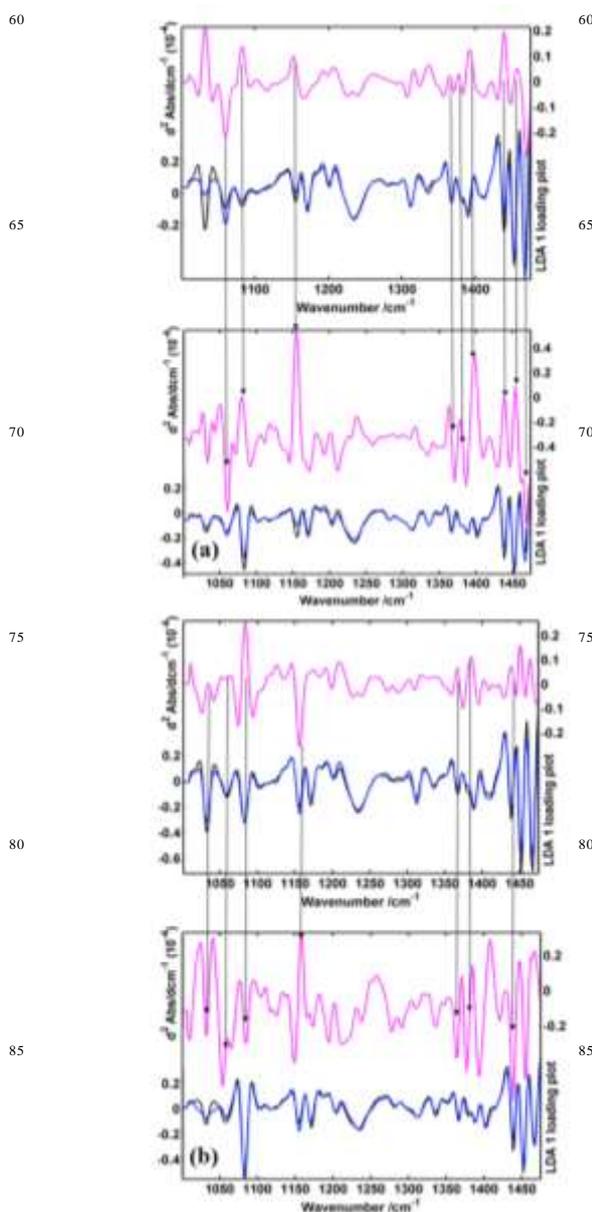
**Fig. 16** PC-LDA analysis of EC23 (blue) treated cells and the control (black) cells at days 7 (a), 5 (b) and 3 (c).

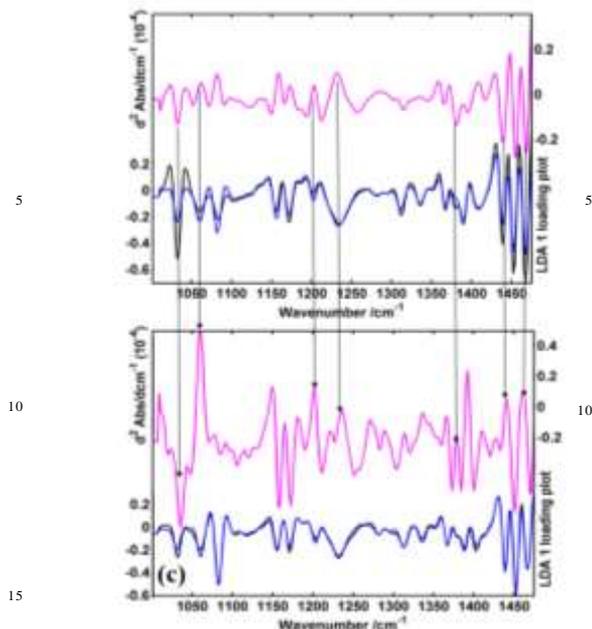
### Comparison of Control and EC23 treated cells after 7 days using Synchrotron FT-IRMS

Synchrotrons provide a highly collimated beam of light, which can be in the order of a 100-1000 times brighter than traditional sources.<sup>23,31,32</sup> The IR region of the emitted synchrotron radiation can be directed into a benchtop FT-IRMS and used as a source instead of conventional blackbody sources.<sup>58</sup> This high brightness of synchrotron sources enables much smaller regions to be probed with an acceptable Signal to Noise ratio (S/N) thus,

greater spatial resolution approaching the diffraction limit can be achieved.<sup>32,59-62</sup> To see if there is consistency between data collected using synchrotron FT-IRMS and the data collected using the benchtop FT-IR imaging instrument with an FPA detector, data was recollected from the same EC23 samples and DMSO control samples at the Diamond Light Source and subjected to the same analysis as for the array detection. Separation from control and the EC23 treated TERA2.cl.SP12 cells were seen at days 3, 5 and 7 (Fig. 16).

Separation of the data collected synchrotron facility is observable at all three time periods. To determine whether the separation observed using a synchrotron source mirrors that of the FT-IR imaging, mean spectra and loadings plots were compared and can be seen in Fig. 17.





**Fig. 17** Comparison of mean spectra and LDA 1 loading plots for synchrotron and FT-IRMS imaging generated data for EC23 treated TERA2.cl.SP12 and control cells at 7 (a), 5 (b) and 3 (c) days (Imaging 20 (top) synchrotron (bottom))

When comparing the loading plots in Fig. 17, clear similarities can be seen between the data collected using an FT-IRMS array detector and the data collected at the synchrotron facility. This 25 therefore, shows consistency between the two types of Infrared spectroscopy analytical techniques. However, some spectral differences can be seen. An explanation for this could be the number of spectra recorded using the two different types of instrumentation. At the synchrotron facility, 311 single cell 30 spectra were recorded when the EC23 treated cells were compared against the control cells, whereas 1881 cell spectra were collected using the array detector for the same comparison. This greater number of cell spectra recorded will build a better picture of the spectral differences between the two different types 35 of treated cells and a low number of spectra may not be sufficient enough to fully explain the cell cycle i.e. the spectral variations due to where the cells are within the cell cycle. Studies by both Flower<sup>29</sup> and Boydston-White have explained the spectral heterogeneity seen when collecting data from single cells. This 40 spectral heterogeneity is a product of the cell cycle with the S phase producing the most heterogeneous spectra while phases G1 and G2, show a relatively small variability.<sup>63</sup> Therefore, a low number of single cell spectra recorded could mean that there is a bias included where more cells from one phase of the cell cycle 45 are spectrally represented than another. Nevertheless, moderate

reproducible separation was observed when comparing the results produced by the two different types of instrumentation. Therefore, even though there is an S/N advantage when using synchrotron light as a source, S/N rivalling that seen at a 50 synchrotron facility can be reproduced when using an array detector.<sup>59</sup>

## Summary and Conclusions

The results of this study have shown that FT-IRMS is an effective 55 tool that can distinguish between stem cells and their differentiating derivatives without the need for fluorescent or magnetic biomarkers. However, the use of both appropriate substrates and scatter correction algorithms is essential, with the result that the spectra observed are indeed indicative of the 60 cellular biochemistry occurring and are not distorted by either scattering or electric field standing wave effects. With the addition of multivariate chemometric algorithms, pluripotent stem cells can be characterised from their induced derivatives based on the spectral differences thus, highlighting the potential 65 of FT-IRMS in finding spectral biomarkers of differentiation in a non-invasive and label free manner. The ability to characterise the retinoid treated TERA2.cl.SP12 cells at each time period has also shown the potential that FT-IRMS has in being able to monitor the progression of stem cell differentiation. Potentially, 70 this could mean that FT-IRMS can be used as a quick screening tool in future stem cell therapy for highlighting the location of undifferentiated cells before implantation thus, reducing the risk of tumour formation. The results from this study have also shown that stem cell differentiation can be seen at earlier time periods 75 than when using more traditional techniques used to monitor these effects.<sup>1,2,5</sup> In respect to these positive findings, it can be predicted that the biological response of stem cells to small molecules such as retinoids could be observed at even earlier time periods. This investigation has also confirmed findings from a 80 previous study where distinct differences in the response of the TERA2.cl.SP12 cells were seen when the retinoid ATRA and synthetic retinoid EC23 were added.<sup>5</sup> By monitoring the spectral changes of the TERA2.cl.SP12 cells over time, we hypothesise that the ATRA and EC23 treated TERA2.cl.SP12 are likely to be 85 following alternative differentiation pathways and possibly producing different types of neurons; conclusions that have not been previously drawn. However, to understand this process fully, earlier and later time periods will need to be investigated in greater detail, which is as part of our future work. Although,

significant differences in the response of the TERA2.cl.SP12 cells can be seen in the results, PC-LDA has been able to show the spectral similarities of the TERA2.cl.SP12 cells differentiating to produce neural cells when compared to the cells treated with the synthetic retinoid EC19. Flow cytometry results in Fig. 2 suggest that at day 7, the EC19 treated TERA2.cl.SP12 cells have yet to fully commit to the differentiation process. However, results in this study have been able to characterise the EC19 treated cells from the control cells at all time periods therefore, this suggests that the retinoid is having some form of biological effect on the TERA2.cl.SP12 cells, even though the majority of treated cells have yet to down-regulate the SSEA-3 stem cell marker. Spectral ratios have also been produced that look to explain the biological processes of the cells as they differentiate. The results in this study show the data for the EC23 treated cells at day 7 are in agreement with the reduced nucleus to cytoplasm ratio as differentiation progresses, previously explained by Notingher. However, this was not the case for the ATRA and EC19 treated cells, which showed more of a correlation to previous work by Heraud.<sup>33</sup>

Separation of the EC23 treated and the control TERA2.cl.SP12 samples has also been produced with data recorded at a synchrotron facility. Loading plot results from synchrotron data closely resembled the results produced from the spectral images. The data produced indicate that the two analytical techniques used are comparable even though a much larger number of cell spectra were collected when using FT-IRMS coupled with an array detector. This highlights the potential of spectral imaging first described by Krafft *et al.*<sup>31</sup> FT-IRMS spectral imaging can readily obtain much more spectra, quicker and to a comparable quality to data recorded at a synchrotron facility. This is advantageous for single cell analysis as a high number of cell spectra are needed to fully explain the cell cycle.<sup>29</sup> It should be pointed out though that for sub cellular imaging, the synchrotron still enjoys a significant advantage.

To conclude, in this study infrared microspectroscopy has been used to characterise and monitor stem cell differentiation over time. Infrared spectral imaging can be employed as an efficient rapid approach that is reproducible, does not require time consuming sample preparation methods, and is cost effective. When coupling infrared microspectroscopy with chemometric methods, the status of the stem cell phenotype and differentiation can be assessed without the need for potentially harmful biomarkers being introduced. Spectral loadings also provide an

insight into the macromolecule changes within the cells during the differentiation process. Infrared spectra can also be used to provide information on the complexes biological processes involved in the cell cycle and differentiation.

Although previous studies<sup>22,36</sup> have used infrared microspectroscopy to monitor the spectral fingerprint of stem cells during the differentiation process, in this study we have shown that we are able to monitor the differentiation process of the cells, as well as the biochemical changes of the pluripotent stem cells through time without attendant aberrations or errors. Indeed, we have also shown that by having control cells at the same period as differentiating, we are able to dissociate the biochemical changes of the pluripotent cells from spectral differences. This, therefore, means that any biochemical changes observed between differentiating cells and control stem cells at the same time period are only due to differentiation and we expect that the protocols we have employed will become standard methodology to support robust and reliable studies of this type.

#### Acknowledgements.

PG acknowledges STFC for funding of beamtime. GC acknowledges support from the School of Chemical Engineering and Analytical Science at UoM for a PhD scholarship and also Michael Brown and the Paterson Institute for Cancer Research for allowing the use of their facilities.

#### References:

- [1] V. F. Segers and R. T. Lee, *Nature.*, 2008, **451**, 937–942
- [2] R. Langer, *Tissue Engineering.*, 2007, **13**, 1–2,
- [3] L. M. Hoffman and M. K. Carpenter, *Nature Biotechnology.*, 2005, **23**, 699–708
- [4] K. Nagano, Y. Yoshida and T. Isobe, *Proteomics.*, 2008, **8**, 4025 – 4035
- [5] J. W. Chan, D. K. Lieu, T. Huser, and R. A. Li, *Analytical Chemistry.*, 2009, **81**, 1324–1331.
- [6] M. J. Walsh, A. Hammiche, T. G. Fellous, et al, *Stem Cell Research*, 2009, **3**, 15–27.
- [7] M. J. German, H. M. Pollock, B. Zhao, *Investigative Ophthalmology & Visual Science.*, 2006, **47**, 2417–2421.
- [8] H. Fukuda, J. Takahashi, K. Watanabe, *Stem Cells.*, 2006, **24**, 763–771.

- [9] J. Pruszek, K. C. Sonntag, M. H. Aung, R. Sanchez-Pernaute, and O. Isacson., *Stem Cells.*, 2007, **25**, 2257–2268
- [10] V. Doussset, T. Tourdias, B. Brochet, C. Boiziau, and K.G. Petry, *Journal of the Neurological Sciences.*, 2008, **265**, 122–126.
- [11] S. A. Przyborski, *Stem Cells.*, 2001, **19**, 500-504.
- [12] S. A. Przyborski, V. B. Christie, M. W. Hayman, R. Stewart, G. M. Horrocks, *Stem Cells Deriv.*, 2004, **13**, 400-408.
- [13] V. B. Christie, T. B. Marder, A. Whiting, S. A. Przyborski, *Mini-Rev. Med. Chem.*, 2008, **8**, 601-608.
- [14] V. B. Christie, J.H. Barnard, A. S. Batsanov, C. E. Bridgens, E. M. Cartmell, J. C. Collings, D. J. Maltman, C. P. Redfern, T. B. Marder, S. A. Przyborski, A. Whiting, *Org. Biomol. Chem.*, 2008, **6**, 3497-3507.
- [15] D. J. Maltman, V. B. Christie, J. C. Collings, J. H. Barnard, S. Fenyk, T. B. Marder, A. Whiting, S. A. Przyborski, *Mol. BioSyst.*, 2009, **5**, 458-471.
- [16] R. C. Addis, J.W. Bulte, and J. D. Gearhart, *Clinical Pharmacology and Therapeutics*, 2008, **83**, 386–389.
- [17] J. K. Pijanka, D. Kumar, T. Dale, I. Yousef, G. Parkes, V. Untereiner, Y. Yang, P. Dumas, D. Collins, M. Manfait, G. D. Sockalingum, N. R. Forsyth and J. Sule-Suso, *Analyst.*, 2010, **135**, 3126–3132
- [18] A. Downes, R. Mouras, and A. Elfick, *J. Biomed. And Biotech.*, 2010
- [19] P. Yu, *J. Synchrotron Rad.*, 2011, **18**, 790–801
- [20] E. Giorganina, C. Contia, P. Ferraris, S. Sabbatinia, G. Tosia, M. Centonze, M. Granob, G. Moric, *Vibrational Spectroscopy.*, 2011, **57**, 30–34
- [21] C. Chonanant, N. Jearanaikoon, C. Leelayuwat, T. Limpaboon, M. J. Tobin, P. Jearanaikoon and P. Heraud, *Analyst.*, 2011, **136**, 2542
- [22] D. Ami, *Biochimica et Biophysica Acta (BBA) - Molecular Cell Research*, Vol. 1783, No. 1. (January 2008), pp. 98-106
- [23] N. Jamin, P. Dumas, J. Moncuit, W.H. Fridman, J.L. Teillaud, L.G. Carr, G.P. Williams, *Proc. Natl. Acad. Sci. U.S.A.* 1997, **95**, 4837
- [24] J. Lee, E Gazi, J Dwyer, M. D. Brown, N. W. Clarke, P. Gardner, *Analyst.*, 2007, **132**, 750 - 755
- [25] P. Bassan, H. J. Byrne, F. Bonnier, J. Lee, P. Dumas, P. Gardner, *Analyst.*, 2009, **134**, 1586-1593
- [26] P. Bassan, H J Byrne, J. Lee, F Bonnier, C. Clarke, P Dumas, E Gazi, M D Brown, N W Clarke, P Gardner, *Analyst.*, 2009, **134**, 1171-1175
- [27] P. Bassan, A. Kohler, H. Martens, J. Lee, H. J. Byrne, P. Dumas, E. Gazi, M. Brown, N. Clarke, P. Gardner, *Analyst.*, 2010, **135**, 268-277
- [28] P. Bassan, A. Kohler, H. Martens, J. Lee, E. Jackson, N. Lockyer, P. Dumas, M. Brown, N. Clarke, P. Gardner, *J. Biophotonics.*, 2010, **3**, 609–620
- [29] K. R. Flower, I. Khalifa, P. Bassan, D. Demoulin, E. Jackson, N. P. Lockyer, A. T. McGown, P. Miles, L. Vaccari, P. Gardner, *Analyst.*, 2011, **136**, 498-507
- [30] C. Hughes, M. Liew, M. D. Brown, A. Sachdeva, P. Bassan, P. Dumas, C. Hart, N. W. Clarke, P. Gardner, *Analyst.*, 2010, **135**, 3133–3141
- [31] C. Krafft, R. Salzer, S. Seitz, C. Ern, M. Schieker, *Analyst.*, 2007, **132**, 647–653
- [32] A. J. Bentley, T. Nakamura, A. Hammiche, H. M. Pollock, F. L. Martin, S. Kinoshita, N.J. Fullwood, *Molecular Vision.*, 2007, **13**, 237-42
- [33] P. Heraud, et al. *Stem Cell Research.*, 2010, **4**, 140–147
- [34] J. Filik, M. D. Frogley, J. K. Pijanka, K. Wehbe and G. Cinque, *Analyst*, 2012, **137**, 853-861
- [35] P. Bassan, J. Lee, A. Sachdeva, J. Pissardini, K. Dorling, J. Fletcher, A. Henderson and P. Gardner, *Analyst accepted*, 2013, DOI: 10.1039/C2AN36090J
- [36] D. Ye, W. Tanthanuch, K. Thumanu, A. Sangmalee, R. Parnpai, P. Heraud, *Analyst*, 2012 Sep 17; **137**(20):4774-84.
- [37] G. Cinque, M. Frogley, K. Wehbe, J. Filik and J. Pijanka, *Synchrotron Radiat. News.*, 2011 **24**, 24–33
- [38] P. Bassan, A. Sachdeva, A. Kohler, C. Hughes, A. Henderson, J. Boyle, J. H. Shanks, M. Brown, N. W. Clarke and P. Gardner, *Analyst.*, 2012, **137**, 1370-1377
- [39] F. L. Martin, et al, *nature protocols.*, 2010, **5**
- [40] S. Wold et al, *Chemometrics and intelligent Laboratory Systems.*, 2001, **58**, 109-130
- [41] R. Kohavi, *International Conference on Artificial Intelligence (IJCAI).*, 1995
- [42] S. Adler, C. Pellizzer, et al, *Toxicology in Vitro.*, 2006, **20**(3), 265-271.

- [43] R. Pal, M. Mamidi, et al, *Archives of Toxicology.*, 2011, 1-11.
- [44] S. Jacob, R. Herschler, *Cryobiology.*, 1986, **23**, 14-27
- [45] Z. Movasaghi S Rehman, I U Rehman, *Applied Spectroscopy Reviews*, 2008, **43**(2), 134-179.
- [46] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts and P. Walter, *Molecular Biology Of The Cell - fourth edition*
- [47] L. Rieppo, J. Rieppo, J. S. Jurvelin, S. Saarakkala, *PLoS ONE.*, 2012, **7** (2)
- [48] M. Abaskharoun, M. Bellemare, E. Lau and R. U Margolis, *asmneuro.org.*, Volume 2 (3) / art:e00039
- [49] E. Bieberich, *Neurosignals.*, 2008, **16**, 124–139
- [50] R. K. Yu, Y. Suzuki, M. Yanagisawa, *FEBS Letters.*, 2010, **584**, 1694–1699
- [51] M. J. Walsh et al, *STEMCELLS.*, 2008, **26**, 108–118
- [52] D. Ami, P. Mereghetti, A. Natalello, S. M. Doglia, M. Zanoni, C. A. Redi, M. Monti, *Biochimica et Biophysica Acta.*, 2011, **1813**, 1220–1229
- [53] I. Notingher, I. Bisson, J. M. Polak, L. L. Hench, *Vibrational Spectroscopy.*, (2004), **35** 199–203
- [54] I. Notingher, I. Bisson, A. E. Bishop, W. L. Randle, J.M. P. Polak and L. L. Hench *Anal. Chem.*, 2004, **76**, 3185-3193
- [55] C. Chonanant et al, *Analyst.*, 2011, **136**, 2542-2551
- [56] D. Ami, T. Neri, et al, *Biochimica et Biophysica Acta.*, 2008, **1783**, 98–106
- [57] W. Tanthanuch, et al, *Journal of Molecular Structure.*, 70
- [58] P. Dumas, N. Jamin, J. L. Teillaud, L. M. Miller and B. Beccard, *Faraday Discuss.*, 2004, **126**, 289-302
- [59] M. Diem, M. Romeo, C. Matthaus, M. Miljkovic, L. Miller, P. Lasch, *Infrared Physics & Technology.*, 2004, **45**, 331–338
- [60] E. Gazi, J. Dwyer, N. P. Lockyer, J. Miyan, P. Gardner, C.A Hart, M.D Brown, N.W. Clarke, *Vibrational Spectroscopy.*, 2005, **38**, 193–201
- [61] E. Gazi, J. Dwyer, J. Miyan, P. Gardner, C. Hart, M. Brown, N.W. Clarke, *Biopolymers.*, 2005, **77**, 18 -30
- [62] E. Gazi, N. P. Lockyer, J. C. Vickerman, P. Gardner, J. Dwyer, C. A. Hart, M. B. Brown, N. W. Clarke, J. Miyan, *Applied Surface Science.*, 2004, **231 – 232**, 452 – 456
- [63] S. Boydston-White, M. Romeo, T. Chernenko, A. Regina, M. Miljković and M. Diem, *Biochim Biophys Acta.*, 2006, **1758**(7), 908–914
- [65] A. Meade, F. Lyng, P. Knief, H. J. Byrne, *Analytical and Bioanalytical Chemistry.*, 2007, **387**(5), 1717-1728
- [66] F. Severcan et al, *Talanta.*, 2000, **53**, 55–59
- [67] M. J. Walsh, et al, *Biochemical and Biophysical Research Communications.*, 2007, **352**, 213–219
- [68] P. Lasch, W. Haensch, D. Naumann, M. Diem, *Biochimica et Biophysica Acta.*, 2004, **1688**, 176–186
- [69] W. Pang, A. A. Ahmadzai, I. I. Patel, X. Qiu, M. Liles, A. J. Quantock, F. L. Martin, *IOVS.*, 2012, **53**, No. 3