Journal of Cellular Biochemistry 104:1427–1438 (2008)

Non-Invasive Analysis of Cell Cycle Dynamics in Single Living Cells With Raman Micro-Spectroscopy

Robin J. Swain,[1] Gavin Jell,[1] and Molly M. Stevens[1,2]*

[1]Department of Materials, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom
[2]Institute of Biomedical Engineering, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom

Abstract Raman micro-spectroscopy is a laser-based technique which enables rapid and non-invasive biochemical analysis of cells and tissues without the need for labels, markers or stains. Previous characterization of the mammalian cell cycle using Raman micro-spectroscopy involved the analysis of suspensions of viable cells and individual fixed and/or dried cells. Cell suspensions do not provide cell-specific information, and fixing/drying can introduce artefacts which distort Raman spectra, potentially obscuring both qualitative and quantitative analytical results. In this article, we present Raman spectral characterization of biochemical changes related to cell cycle dynamics within single living cells in vitro. Raman spectra of human osteosarcoma cells synchronized in G_0/G_1, S, and G_2/M phases of the cell cycle were obtained and multivariate statistics applied to analyze the changes in cell spectra as a function of cell cycle phase. Principal components analysis identified spectral differences between cells in different phases, indicating a decrease in relative cellular lipid contribution to Raman spectral signatures from G_0/G_1 to G_2/M, with a concurrent relative increase in signal from nucleic acids and proteins. Supervised linear discriminant analysis of spectra was used to classify cells according to cell cycle phase, and exhibited 97% discrimination between G_0/G_1-phase cells and G_2/M-phase cells. The non-invasive analysis of live cell cycle dynamics with Raman micro-spectroscopy demonstrates the potential of this approach to monitoring biochemical cellular reactions and processes in live cells in the absence of fixatives or labels. J. Cell. Biochem. 104: 1427–1438, 2008.

Key words: Raman micro-spectroscopy; live cell analysis; cell cycle; cell synchronization; flow cytometry; spectral discrimination

Cell cycle dynamics are intimately involved in all aspects of health and disease, and have justly received significant interest in biomedical research. Changes in cell cycle mechanics induced by drug administration in vitro are routinely investigated to aid in pharmaceutical development, including identification of anti-cancer agents which disrupt cell cycle machinery [Wang and El-Deiry, 2004]. Recently, scientists have explored cell cycle changes induced by cell–biomaterial interactions to develop biocompatible and bioactive tissue engineering scaffolds which support and enhance cell adhesion, proliferation, and differentiation [Lopes et al., 1998; Xynos et al., 2000a,b]. Current histological, cellular and molecular techniques available to study cell cycle biochemistry in culture are invasive, requiring fixation and/or exogenous fluorescent labeling (e.g., flow cytometry) [Wolthuis et al., 1999]. Many research fields would benefit from the development of new techniques which permit non-invasive, in situ monitoring of living cells.

Raman micro-spectroscopy is a laser-based analytical technique which enables rapid and non-invasive biochemical analysis of cells and tissues in the absence of fixatives or labels. The Raman effect is based on the inelastic scattering of light, whereby energy differences

Grant sponsor: Rothermere Foundation; Grant sponsor: Canadian Centennial Scholarship Fund; Grant sponsor: National Science and Engineering Research Council of Canada.

*Correspondence to: Dr. Molly M. Stevens, Department of Materials, Imperial College London, Exhibition Road, London SW7 2AZ, United Kingdom.
E-mail: m.stevens@imperial.ac.uk

Received 5 November 2007; Accepted 7 January 2008
DOI 10.1002/jcb.21720

between incident and scattered photons correspond to the specific vibrational energies of chemical bonds of the scattering molecules [Raman and Krishnan, 1928]. The Raman spectrum of a cell represents a unique biochemical fingerprint, providing information about all cellular biopolymers, and can be used to characterize cell phenotype [Omberg et al., 2002; Notingher et al., 2004b; Crow et al., 2005; Chan et al., 2006], monitor cell death [Verrier et al., 2004], differentiation [Notingher et al., 2004a; Jell et al., 2007], and cellular response to pharmacological treatments [Krafft et al., 2006; Owen et al., 2006].

The ability of Raman micro-spectroscopy to resolve molecular detail on the micron scale makes it a suitable method to elucidate the intricate dynamics of the cell cycle. The eukaryotic cell cycle is a highly regulated, complex process divided into a number of phases: G_0 (resting state), G_1 (gap 1), S (DNA synthesis), G_2 (gap 2), and M (mitosis). Cells can be synchronized in G_0 by serum starvation, and chemical cell cycle inhibitors can be used to target other phases. Chemical inhibitors may be administered to adherent cells in culture, and their effects are reversible upon washout and re-incubation with serum-supplemented medium. Analysis of DNA and cyclins by flow cytometry is the current gold-standard for assessing cell cycle synchronization. However, unlike Raman micro-spectroscopy, flow cytometry requires suspensions of (normally fixed) cells stained with fluorescent tags, and cannot be applied in situ on adherent cells growing in culture. Both infrared (IR) absorption spectroscopy and Raman micro-spectroscopy have been used to analyze pre-synchronized viable cell suspensions [Mourant et al., 2003b; Short et al., 2005] and single fixed and dried cells [Boydston-White et al., 1999; Matthäus et al., 2006] to characterize spectral patterns related to the cell division cycle. Strong water absorption can interfere with IR spectra, requiring potentially artefact-inducing [Mourant et al., 2003a] fixation and dehydration of samples [Boydston-White et al., 2005], or omission of distorted spectral regions [Mourant et al., 2003b]. The weak inelastic scattering of aqueous media makes Raman micro-spectroscopy amenable and uniquely suited to cell cycle analysis of living cells in situ. However, to date this advantage has not been exploited.

Here, we present Raman spectral characterization of biochemical changes related to cell cycle dynamics within single living cells. We collected Raman spectra of human osteosarcoma (MG63 cell line) cells synchronized in G_0/G_1, S and G_2/M phases of the cell cycle and applied multivariate statistics to analyze the changes in cell spectra as a function of cell cycle phase. Specifically, we apply principal components analysis (PCA) to examine cell cycle dependent variations in Raman spectral signatures, and linear discriminant analysis (LDA) as a supervised classification algorithm to differentiate between cells synchronized in different cell cycle phases. The results presented here illustrate the potential of Raman micro-spectroscopy as a sensitive and powerful tool for non-invasive biochemical analysis of live cellular systems.

MATERIALS AND METHODS

Cell Culture

MG63 (human osteosarcoma cell line) cells (ECACC, UK) were grown to 70% confluence before seeding onto sterilized magnesium fluoride (MgF_2) coverslips (Global Optics, UK) at a density of 1×10^4 cells/cm^2 for 24 h. Cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 1% antibiotic/antimycotic, 1% L-glutamine (all from Invitrogen, UK) in a 5% CO_2 incubator at 37°C. MgF_2 was selected because of its low Raman signal and low solubility, and has been shown to have no adverse effects on cell attachment, morphology, and viability [Notingher et al., 2004b].

Cell Synchronization

Cell cycle inhibitors were used to synchronize MG63 cells in G_0/G_1, S, and G_2/M phases of the cell cycle as described previously [Carbonaro-Hall et al., 1993]. At 70% confluence, cells were synchronized in G_0/G_1 by serum starvation (0.1% FBS serum) for 48 h. For S-phase block, G_0/G_1-synchronized cells were incubated with 10% FBS supplemented DMEM containing 2 μg/ml aphidicolin (Sigma-Aldrich, UK), an inhibitor of DNA-polymerase [Pedrali-Noy et al., 1980], for 24 h to synchronize cells at the G_1/S transition point, after which the aphidicolin block was washed off and the cells incubated for 5 h with 10% FBS medium to stimulate synchronized entry into S phase of the cell cycle. G_2/M block was accomplished by treating G_0/G_1-synchronized cells with 10% FBS

medium containing 0.2 μg/ml nocodazole (Sigma), an agent which disrupts tubulin polymerization, for 27 h [Jordan et al., 1992]. Flow cytometry was used to assess cell cycle synchronization. Optimal G_2/M synchronization efficiency was achieved using a 27 h incubation period with nocodazole, rather than a 30 h time point used in previous studies [Yang et al., 1999].

Flow Cytometry

Flow cytometric analysis of DNA content was used to assess cell cycle synchronization as described previously [Krishan, 1975]. Briefly, cell suspensions (1×10^6 cells) were washed three times in phosphate-buffered saline (PBS, Invitrogen, UK), fixed in 70% ice-cold ethanol whilst vortexing, and stored at 4°C. The cells were centrifuged for 5 min at 400*g* and the pellet re-suspended in 0.5 ml PBS, and 300 μl propidium iodide (PI, Sigma, 50 μg/ml). To ensure a fluorescence signal proportional to DNA content, 50 μl RNase-A (Sigma, 100 μg/ml) was added to remove cellular RNA. The samples were then covered and incubated for 30 min at room temperature. DNA content histograms were collected using a FACSCalibur flow cytometer (Becton Dickenson, Franklin Lakes, NJ), which was calibrated using the DNA histogram–profile of asynchronous control cells. Cell cycle markers were added to DNA content histograms using the CellQuest Pro software to determine the percentage of cells in G_0/G_1, S, and G_2/M phases of the cell cycle. Flow cytometry measurements were performed on at least three independently synchronized cell populations for each treatment group.

Raman Micro-Spectroscopy

Raman spectra were measured with a Renishaw InVia spectrometer connected to a Leica microscope, as described previously [Notingher et al., 2004b]. Before Raman analysis, the cell culture medium was removed and the cells rinsed twice with PBS. Spectra of live cells were measured in PBS maintained at 37°C. A high power 785 nm diode line focus laser (~120 mW power before objective) was used for excitation. This laser wavelength and power have previously been reported to have no effect on cell viability or phenotype [Notingher et al., 2002]. Other researchers have used similar laser wavelength and power to examine live cells in suspension [Short et al., 2005]. The laser was focused on individual cells by a 63× (NA = 0.9) long working distance (2 mm) water immersion objective. Under these conditions, the laser illuminates an elliptical region at the focal plane, with approximately 10 μm × 20 μm lateral spatial resolution. The spectrum of each cell was calculated as the average of four or five spectra (depending on cell size and morphology) measured at different lateral positions to ensure a complete mapping of the cell. Each point spectrum was integrated for 40 s, giving a total spectral integration time of 160–200 s per cell.

Fifteen cells per cell cycle group were analyzed with Raman micro-spectroscopy. To reduce the influence of system variability, and to limit the potential of artificial data clustering [Notingher et al., 2004b], cell spectra were measured on three separate days, with five cells per treatment group analyzed per day. Spectra were collected from a glass reference sample each day to correct for instrument response [Notingher et al., 2004b]. The spectrum of PBS buffer was measured each day, with signal integration time of 300 s.

Reference cellular components, phosphatidyl choline (lipid), DNA (nucleic acid), and actin (protein) were purchased from Sigma Ltd. (UK) and used without further preparation. The spectra of the dry samples were measured for comparison with the principal component loading vectors generated by PCA.

Data Analysis

Raman spectra were pre-processed and analyzed in Matlab (The MathWorks, Natick, MA) using in-house software previously developed by Notingher et al. [2004b]. New programs were compiled for fluorescence background subtraction, standardization and alignment of the wavenumber axis, and LDA. The PBS contribution to the signal was subtracted using the projection method described by Maquelin et al. [2000], and the background was corrected using the Modpoly algorithm described by Lieber and Mahadevan-Jansen (5th order polynomial, 1,000 iterations) [Lieber and Mahadevan-Jansen, 2003]. The wavenumber axis of each spectrum was then standardized and aligned to the sharp phenylalanine peak at 1,003 cm^{-1}, and the mean spectrum for each cell was calculated from the aligned group of 4–5 spectra per cell. Average spectra for each treatment group were calculated from the 15 mean cell

spectra per group, which were individually smoothed using a Savitsky-Golay filter (5 points, 2nd order polynomial), and then normalized using the standard normal variate (SNV) transformation [Wolthuis et al., 1999]. For PCA, spectral differentiation of the aligned, PBS-corrected spectra was used to remove the background, instead of the Lieber and Mahadevan-Jansen method. Second derivative spectra were computed using a Savitsky-Golay filter (21 points, 2nd order polynomial) and then normalized using the SNV transformation.

PCA identifies combinations of variables that account for major sources of variance within a data set [Wold et al., 1987]. PCA is often used to compress high-dimensional data by expressing the data in terms of a small number of principal components (PCs). The PCs contain no overlapping information, with the greatest variance captured by PC1, the second greatest variance by PC2, and so on. Data compression is achieved by retaining only the PCs that describe the most significant variance, while ignoring higher-order PCs which mostly account for random noise. The mean-centered, second derivative Raman spectra were analyzed using a singular value decomposition PCA algorithm.

PCA is a mathematical transformation which does not incorporate class membership of data. LDA is a supervised classification algorithm for discrimination of sample groups. The algorithm computes linear discriminant functions, which are directions in the spectral space that maximize the ratio of between-class variance to within-class variance according to Fisher's criterion [Hair, 1998]. Projecting the data along these functions produces maximal group separation. The significant PCs were used to generate the LDA model, and the cell synchronization results obtained by flow cytometry were used to calculate the prior probabilities of each cell cycle group. Classification accuracy was assessed using leave-one-out cross-validation, whereby the class of a spectrum is predicted using an LDA model built from the full data set excluding the spectrum in question. This method is repeated, with each spectrum left out in turn, so that each spectrum is predicted once.

RESULTS

Flow Cytometry

We conducted flow cytometry experiments to assess cell cycle synchronization induced by serum starvation and cell cycle inhibitors aphidicolin and nocodazole. DNA content histograms which describe the percentage distribution of cells in G_0/G_1, S, and G_2/M phases of the cell cycle are included in Figure 1. Serum starvation rendered cells quiescent in G_0/G_1 with 94% accuracy, and mitotic block with nocodazole synchronized 77% of cells in G_2/M. Synchronization at the G_1/S-transition point with aphidicolin synchronized 94% of cells in G_0/G_1 ($n = 5$, data not shown), and subsequent washout with re-stimulated growth successfully synchronized 68% of cells in S phase. Given the difficulty associated with S-phase synchro-

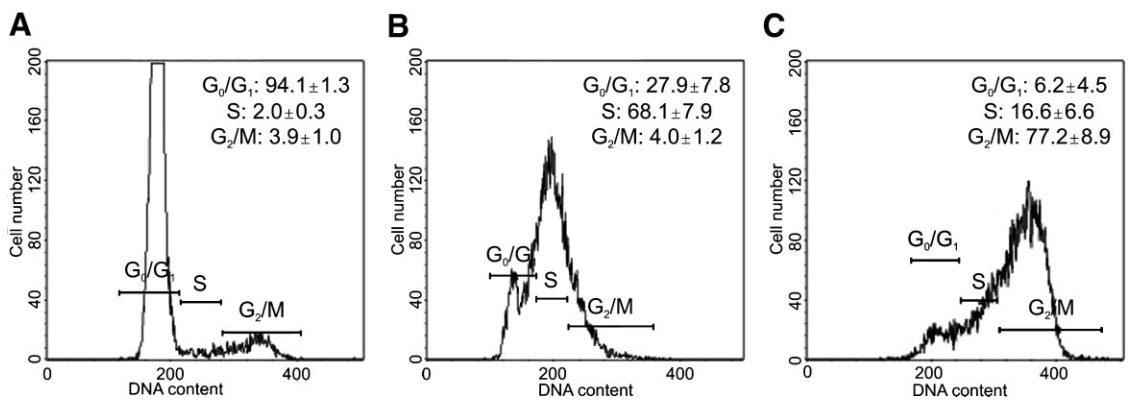

Fig. 1. Flow cytometry analysis of MG63 cell synchronization. Representative DNA content histograms are shown for (**A**) cells blocked in G_0/G_1 by serum starvation ($n = 4$), (**B**) cells synchronized in S phase with aphidicolin, washout and re-stimulated growth ($n = 3$), and (**C**) cells arrested in G_2/M with nocodazole ($n = 3$). Cell cycle markers used to estimate the percentage distribution of cells in each phase are also included.

nization, these results compare favorably with MG63 synchronization results reported in the literature [Carbonaro-Hall et al., 1993; Kawabata et al., 2000], and are higher than contact-inhibition synchronization achieved in previous Raman cell cycle studies [Short et al., 2005]. The distribution of asynchronous cells was approximately 60% G_0/G_1, 15% S phase, and 25% G_2/M ($n = 3$).

Cell Morphology

MG63 cell morphology varied at different points of the cell cycle. G_0/G_1 cells were elongated, whereas S and G_2/M cells showed a gradual progression to a more rounded morphology. Cells round up during mitosis as the cytoskeleton rearranges, but G_2/M arrest by nocodazole, an anti-microtubule agent which disrupts the cellular cytoskeleton, may contribute to the rounding of G_2/M cells [Jordan et al., 1992]. Pictures representative of cell morphology for each cell cycle group are included in the insets of Figure 2.

Raman Cell Spectra

The average normalized Raman spectra of MG63 cells synchronized in G_0/G_1, S, and G_2/M phases of the cell cycle (after pre-processing) are shown in Figure 2. The spectra of MG63 cells reported here are similar to previously published Raman spectra of MG63 cells [Notingher et al., 2004b], and consist of peaks corresponding to molecular vibrations of all cellular components (i.e., nucleic acids, proteins, lipids, and carbohydrates). Detailed tables of peak assignments can be found elsewhere [Notingher et al., 2004a].

Unsupervised PCA

The scree plot of singular values generated by the PCA indicated the first six PCs accounted for the main sources of variance, a conclusion supported by the low signal-to-noise ratio (SNR) observed in the higher PCs.

PCA was used to compress the high-dimensional spectral data to scores on the first

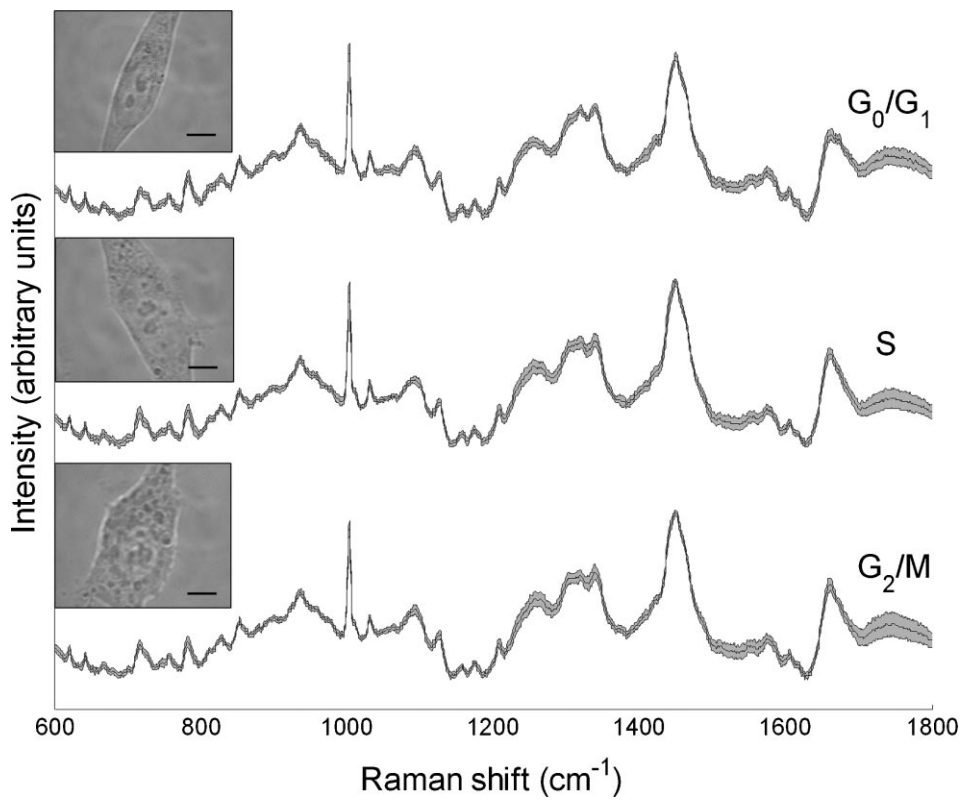

Fig. 2. Average processed Raman spectra of MG63 cells (middle black lines represent averages, shaded areas ± standard deviation): spectra are labeled as G_0/G_1, S phase, and G_2/M, based on cell cycle treatment. The spectra have been shifted vertically for clarity. Representative pictures of synchronized cells are included in the insets (scale bar = 10 μm).

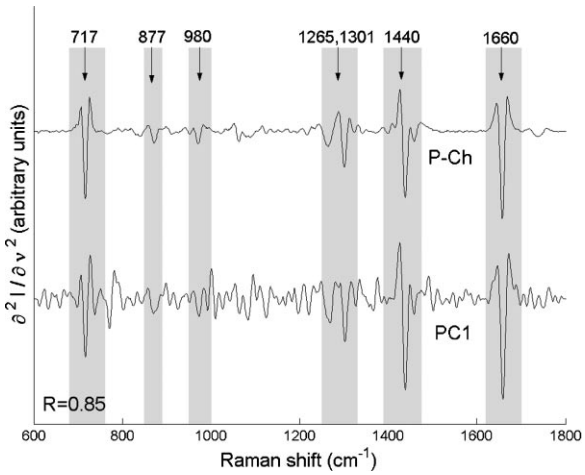

Fig. 3. Comparison of PC1 with the second derivative spectrum of phosphatidyl choline (P-Ch). Common peaks are indicated, and the spectra have been shifted vertically for clarity.

six PCs. The PCs contain spectral features corresponding to the main molecular species responsible for the statistical variation between spectra. Principal component 1 (PC1) was dominated by lipid spectral features, with peaks at 717, 877, 980, 1,265, 1,301, 1,440, 1,660 cm^{-1}, all of which describe molecular vibrations in lipids. These features are highlighted in Figure 3, where PC1 is compared to the second derivative Raman spectrum of phosphatidyl choline (P-Ch), a model lipid. A statistically significant correlation ($R_{PC1} = 0.85$) between PC1 and P-Ch was obtained by calculating

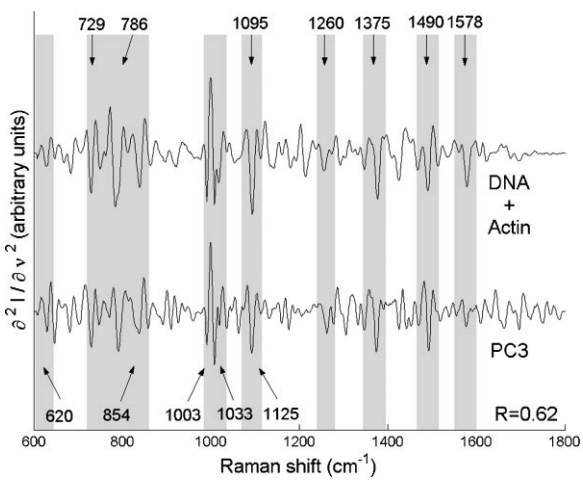

Fig. 4. Comparison of PC3 with the vector sum of the second derivative spectra of DNA and actin. Common peaks are indicated, and the spectra have been shifted vertically for clarity.

the correlation coefficient (R), which provides a statistical measure of the similarity between spectra (identical spectral profiles exhibit a correlation coefficient of unity, i.e., $R = 1$). PC3 displayed prominent features found in spectra of nucleic acids (e.g., DNA, with peaks at 729, 786, 1,095, 1,260, 1,375, 1,490, and 1,578 cm^{-1}) and cellular proteins (e.g., actin, with peaks at 620, 854, 1,003, 1,033, and 1,125 cm^{-1}). These marker peaks are indicated in Figure 4, where PC3 is compared to the sum of the second derivative Raman spectra of DNA and actin (to model nucleic acid and protein, respectively). Statistical comparison of these two spectral vectors yielded a correlation of $R_{PC3} = 0.62$. PC2 and PC4-PC6 had lower SNR and contained little information. Furthermore, the scores on PC2 revealed clustering according to the day of spectral acquisition: spectra measured on the same day tended to cluster together, with sub-clusters of segregated cell cycle groups. Despite every effort taken to minimize spectral variation, day-to-day spectral variation can occur due to instrumental factors or variation in sample preparation. Previous Raman cell biology studies have also reported significant day-to-day spectral variation [Maquelin et al., 2000; Short et al., 2005]. The segregation of the effects of day-to-day spectral variation into a single PC makes it easier to filter out these influences on the data, which otherwise might affect data interpretation. It also justifies our approach to collect spectra of cells from all phases on each experimental day in order to compensate for day-to-day variation, and thereby prevent artefactual clustering of cell spectra (i.e., clustering of spectra based on noise or daily system variation) [Notingher et al., 2004b].

In generating PCs, the PCA algorithm assigns score values on each PC to each spectrum, where a score value represents the contribution of a PC to a spectrum. PC score plots can be used to identify clustering of data within the PC model. Clustering of cell cycle groups was observed by plotting the scores on PC1 versus scores on PC3 (Fig. 5). G_0/G_1 and G_2/M cells were separated with only a few G_2/M-blocked cells grouping near the G_0/G_1 cell cluster (Fig. 5A). Plotting the scores of S phase synchronized cells (Fig. 5B) revealed that they lie in a broad band between the G_0/G_1 and G_2/M cell clusters, with much larger variance than the other cell groups.

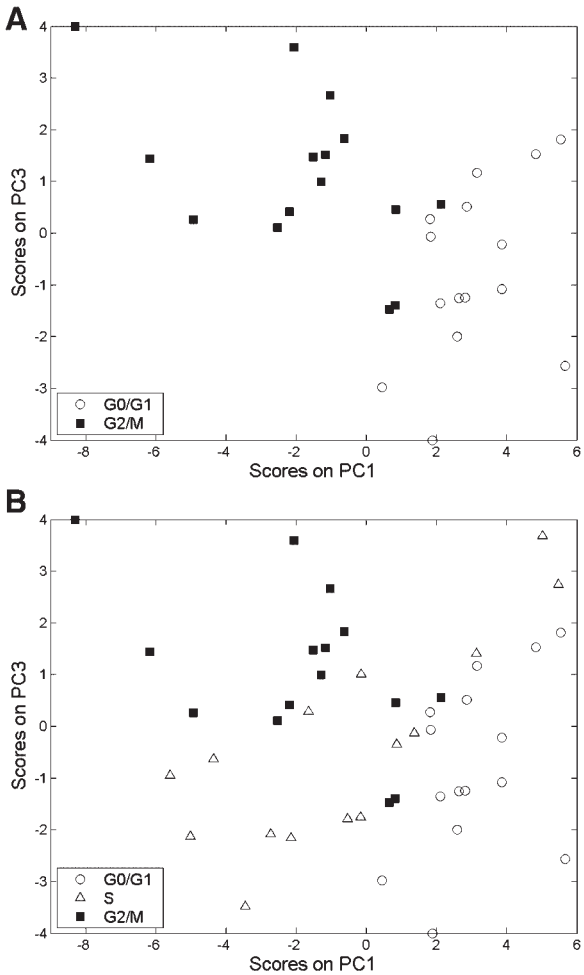

Fig. 5. PC1–PC3 score plot. **A**: Scores of G_0/G_1 and G_2/M phase cell populations. **B**: Scores of G_0/G_1, S, and G_2/M phase cell populations.

Linear Discriminant Analysis

We generated a supervised LDA classification model using the significant PC scores in order to differentiate between cells synchronized in different cell cycle phases. The scores on PC2, which described day-to-day spectral variation, were not included in the LDA model. The LDA prior probabilities of each cell cycle group were calculated using the cell synchronization data obtained using flow cytometry. In a standard LDA, the prior probability of class k, π (k), is found simply by calculating its frequency within the training data, that is, $15/45 = 1/3$ for our data. However, this assumes 100% synchronization efficiency of the chemical treatments, which is not the case. Instead, we incorporate the cell cycle distribution data from Figure 1 to calculate the prior probability for the G_0/G_1 phase class as:

$$\begin{aligned}\pi(G_0/G_1) &= (15/45)(G_0/G_1\%) \\ &\quad + (15/45)(S\%) + (15/45)(G_2/M\%) \\ &= (15/45)(94.1\%) + (15/45)(27.9\%) \\ &\quad + (15/45)(6.2\%) = 0.4273\end{aligned}$$

This expression describes the probability that an arbitrary spectrum belongs to a G_0/G_1-phase cell, by incorporating the probability of analyzing a G_0/G_1 phase cell within the G_0/G_1-synchronized culture (94.1%), and within the S phase-synchronized (27.9%) and G_2/M-synchronized (6.2%) cultures. Similarly, the prior probabilities for S and G_2/M phase classes were calculated as:

$$\begin{aligned}\pi(S) &= (15/45)(2.0\%) + (15/45)(68.1\%) \\ &\quad + (15/45)(16.6\%) = 0.2890 \\ \pi(G_2/M) &= (15/45)(3.9\%) + (15/45)(4.0\%) \\ &\quad + (15/45)(77.2\%) = 0.2837\end{aligned}$$

Two linear discriminant functions (LD1 and LD2) were generated from the scores on PC1–PC6 (exclusive of PC2). These scores were then projected onto LD1 and LD2 to produce an LDA score plot (Fig. 6).

The results of the LDA model cross-validation are reported in Table I. The ability of the LDA model to predict the phase of cell cycle synchronization based on cellular Raman spectra was evaluated by calculating the sensitivity and specificity for each cell cycle group. LDA model classification accuracy was highest for cells synchronized in G_0/G_1 (93%), with

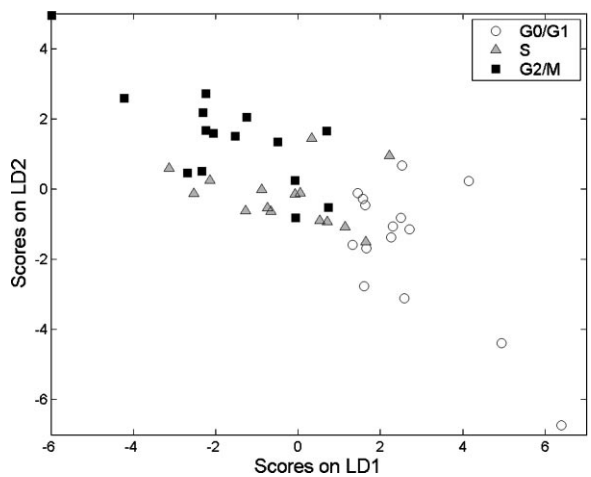

Fig. 6. LD1–LD2 score plot: LDA scores of G_0/G_1, S, and G_2/M phase cell populations.

TABLE I. Cell Cycle Classification Accuracy of Three-Group LDA Model as Assessed by Leave-One-Out Cross-Validation

	Raman-predicted cell cycle stage		
Cell cycle block	G_0/G_1	S phase	G_2/M
G_0/G_1	14	1	0
S phase	3	10	2
G_2/M	1	5	9
Sensitivity (%)	93	67	60
Specificity (%)	87	80	93

only one cell misclassified as an S phase cell. Classification accuracy was lower for cells synchronized in S phase (67%), with five cells misclassified. The predictive power was lowest for G_2/M phase cells (60%), with five cells misclassified as S phase cells and one as a G_0/G_1 phase cell. Overall, 73% of the cells were correctly classified, with most errors occurring due to misclassification of cells synchronized in G_2/M phase as S phase cells. With only one misclassification between G_0/G_1 and G_2/M cell populations, the PCA–LDA model exhibits excellent 97% spectral discrimination between G_0/G_1 and G_2/M phase cells. In addition, the model demonstrated 89% spectral discrimination between quiescent G_0/G_1 phase cells and actively cycling cells in S or G_2/M phases.

DISCUSSION

To study changes in Raman spectra of live MG63 cells as a function of cell cycle phase, we collected spectra of cells from MG63 cultures synchronized in G_0/G_1, S, and G_2/M phases of the cell cycle. The flow cytometry results (Fig. 1) revealed excellent levels of cell cycle synchronization, indicating that the synchronized populations analyzed with Raman microspectroscopy contained high-proportions of cells within the correct target phase. Once cell cultures were synchronized, the use of a line focus laser with a large spot-size (i.e., not diffraction-limited) facilitated rapid collection of spectra from single living cells within a matter of minutes (typically 5–10 min per cell to scan the entire cell volume). This is in contrast to hyperspectral imaging [Matthäus et al., 2006], where the collection of thousands of point spectra per cell requires several hours, making the use of fixatives to preserve cellular conditions a pre-requisite.

The Raman spectra of synchronized MG63 cells reported here indicate minor spectral variations between cell cycle groups. The similarity between spectra of cells in G_0/G_1, S, and G_2/M is expected, given that the cells were from the same cell line and differed only in cell cycle phase, and also because we monitored spectral changes on the level of the entire cell, rather than subjectively targeting specific sub-cellular regions (e.g., cell nuclei). The small standard deviations within each cell cycle group suggest cell populations with a high degree of homogeneity, in agreement with the flow cytometry results.

We used principal components analysis (PCA) to identify the spectral peaks which account for the variance between spectra of cells synchronized in different cell cycle phases. PC1 (Fig. 3) and PC3 (Fig. 4) described significant trends within the data, with PC1 describing changes in cellular lipid and PC3 changes in DNA and protein. The observation of biochemical-specific spectral peaks in the PCs was supported by statistically significant correlation between the PCs and the spectra of reference cellular components (PC1 and P-Ch, $R_{PC1} = 0.85$, and PC3 and DNA + actin, $R_{PC3} = 0.62$). In general, PCs do not represent pure biochemical components, but rather combinations of peaks from multiple components [Crow et al., 2005]. This explains the presence of extra peaks in the PCs and why the correlation coefficients are less than one: the significant correlation $R_{PC1} = 0.85$ indicates that PC1 mostly describes lipid features, which account for the highest variance within the dataset, but also contains some additional information about molecular vibrations of other cellular components. PC3 captures the smaller variance due to changes in nucleic acids and proteins ($R_{PC3} = 0.62$), as well as some additional biochemical components. Discrepancies between the PCs and reference spectra, and hence lower correlation coefficients, may also be due in part to the comparison of solid state spectra of the dry reference components with the PCs derived from solution phase spectra. The fact that more changes in cellular lipids were detected than in DNA and protein is not surprising, for two reasons: cells contain much more lipid than DNA, and our results represent spectral changes on the level of the entire cell; and it is

likely that it is specific types of protein rather than total protein content which changes as cells advance through the cell cycle [Boydston-White et al., 2006]. However, these findings differ from an earlier study by Notingher et al. [2003], where changes in the levels of DNA accounted for the most significant differences between the spectra of A549 adenocarcinoma cells from a confluent monolayer and spectra of rounded cells on top of the monolayer. Those findings were interpreted as spectral differences between cells resting in G_0/G_1 and those actively dividing in $S/G_2/M$ phases of the cell cycle. However, contact inhibition alone generally yields low levels of synchronization, and the degree of synchronization was not assessed, nor was the cell cycle phase of individual cells confirmed. These discrepancies may therefore be due to analysis of cell clusters and confluent cell layers rather than single individual cells.

PCA also serves as an unsupervised clustering algorithm to examine grouping of data. PCA assigns score values on each PC to each spectrum, where a score value represents the contribution of a PC to a spectrum. PC score plots can be used to determine whether or not spectra are related: clustering or grouping of spectra within these plots indicates systematic differences between the spectra [Boydston-White et al., 2006]. The clustering of cell cycle groups in the PC1–PC3 score plot (Fig. 5) suggests that cluster heterogeneity increases as synchronization efficiency decreases; the increasing variance (i.e., cluster spread) in cell cycle populations from G_0/G_1 to G_2/M to S phase reflects the lower efficiencies of G_2/M and S phase synchronization. The cluster positions ($G_0/G_1 \rightarrow S \rightarrow G_2/M$) reveal a decrease in score value on PC1 and an increase in score value on PC3, which implies a decrease in cellular lipid content and increases in cellular nucleic acid and protein content as cells progress through the cell cycle. A previous study in which suspensions of viable, synchronized cells were investigated with Raman micro-spectroscopy to characterize spectral patterns related to the cell division cycle also reported similar trends [Short et al., 2005].

The clustering in the PC1–PC3 score plot (Fig. 5) is strikingly similar to the cell cycle dependent variations in Fourier transform IR (FTIR) micro-spectra of single HeLa cells recently reported by Boydston-White et al. [2006]. In that study, good spectral discrimination between G_0/G_1, S, and G_2/M phase cells was achieved with PCA applied only to the amide I region (1,598–1,702 cm^{-1}) of cell nucleus spectra. The PCA presented here was not confined to a narrow spectral band, and so discrimination is based on spectral features from multiple biochemical components over the entire spectral fingerprint range (600–1,800 cm^{-1}). The fact that this cell cycle trend is preserved despite the use of different mammalian cell lines (MG63 and HeLa), vibrational spectroscopic techniques (Raman micro-spectroscopy versus FTIR), and sample preparation conditions (live cells versus formalin-fixed and dried cells) illustrates the potential of vibrational spectroscopy as a robust and powerful tool for non-invasive cellular biochemical analysis.

We used the significant PC scores from PCA to generate a supervised LDA model to classify cells according to cell cycle phase. The LD1–LD2 plane (Fig. 6) represents the optimal spectral subspace for group discrimination. The LDA separated G_0/G_1 phase cells from S and G_2/M phase cells along the direction of LD1, reducing the overlap between clusters of G_0/G_1 and G_2/M phase cell populations observed in the PC1–PC3 score plot (Fig. 5A). The LDA also decreased the spread of the diffuse cluster of S-phase synchronized cells. These results can be attributed to the fact that LDA model discrimination is based on score values on five PCs (PC1, PC3–PC6), whereas the PC1–PC3 plane only represents a small subset of the larger PC-space from which the functions LD1 and LD2 were computed.

The PCA–LDA model exhibited excellent 97% spectral discrimination between G_0/G_1 and G_2/M phase cell populations. The high efficiency of G_0/G_1 (94%) and G_2/M (77%) synchronization combined with the high sensitivity of Raman spectral characterization enables reproducible detection of cell cycle specific differences between these extremes of the cell cycle. Furthermore, the model also demonstrated a high-level of spectral discrimination (89%) between G_0/G_1 phase cells and actively cycling cells in either S or G_2/M phases. These results reflect the high synchronization efficiency of serum starvation and hence increased homogeneity of the G_0/G_1 cell population over the S and G_2/M phase groups. This is also evident when all three cell cycle groups are

considered separately: classification accuracy for G_0/G_1 group (93%) is much higher than that observed for either S-phase (67%) or G_2/M (60%) cells. A decrease in synchronization efficiency leads to a loss in predictive power, due to increased cluster overlap and higher levels of data label uncertainty. While the PCA results indicate spectra of S-phase cells display features intermediate to those of G_0/G_1 and G_2/M cell spectra, the LDA projection (Fig. 6) appears to merge the S and G_2/M phase clusters, obscuring class distinction which leads to lower prediction accuracy.

The LDA model classification accuracy assessed using cross-validation (Table I) did not decrease with decreasing synchronization efficiency, that is, $G_0/G_1 > G_2/M > S$, as expected. We expected to observe this trend on consideration that cluster heterogeneity in the PC1–PC3 plane appeared to increase with decreasing synchronization efficiency, and given that the accuracy of classification algorithms generally decreases as cluster heterogeneity increases. However, the LDA model prior probabilities explain the observed trend: although the synchronization efficiency for G_2/M phase arrest (77%) is higher than that for S phase synchronization (68%), the prior probability for the S phase class, $\pi(S) = 0.2890$, is slightly greater than the prior probability for the G_2/M phase class, $\pi(G_2/M) = 0.2837$. Thus LDA classification accuracy decreased with prior probability, rather than synchronization efficiency. Also, it is possible that the population of S phase cells within the five-dimensional PC-space (PC1–PC6 exclusive of PC2) may in fact be less heterogeneous than the G_2/M phase cell population, at least within our model. In this case, examining the data in the two-dimensional PC1–PC3 plane could produce a distorted representation of the overall structure of the data clusters. This interpretation is supported by the LDA score plot (Fig. 6), where the S phase cell population is actually less heterogeneous than the G_2/M phase cluster, if we estimate cluster heterogeneity by the average Euclidean distance of the data from the cluster centroid in the LD1–LD2 plane.

One limitation with our LDA model is that it requires unequivocal initial labeling of spectra as G_0/G_1, S, or G_2/M phase cell spectra. Initial labeling does not incorporate class prior probabilities, which account for imperfect cell synchronization. As the exact cell cycle phase of each individual cell was not determined, the initial labeling is based on the cell synchronization treatment administered, rather than on true/verified cell cycle phase. Thus the supervisor for the LDA model is imperfect, and has to contend with some degree of "label noise." In this case, it is possible that some misclassified G_2/M phase cells were actually in S phase of the cell cycle, and hence correctly classified, but are considered misclassified because they were initially labeled as G_2/M phase cells. A similar argument holds for misclassified G_0/G_1 and S phase cells. Thus, the use of data which does not permit unambiguous initial labeling is likely to be a contributing factor to LDA model misclassifications. We hope to improve cell cycle-related spectral discrimination and classification in a larger sample of synchronized cells, by developing fully supervised classification models which can incorporate label noise in the training data, or by removing initial label noise by using fluorescence staining to confirm the cell cycle status of individual cells [Boydston-White et al., 2006].

In this study we have taken full advantage of the high sensitivity and non-invasive sensing capabilities of Raman micro-spectroscopy to characterize biochemical dynamics related to the cell division cycle within live, synchronized cells. By analyzing live cells in situ, we have avoided the possibility of artefacts arising from fixation and desiccation. The high levels of chemical synchronization enabled spectra of individual cells from the target phases to be collected and subsequently analyzed using multivariate statistics. Small biochemical differences between cells were detected and used to classify cells based on their Raman spectral signatures. This research demonstrates the potential for Raman micro-spectroscopy to non-invasively assess cell-cycle related changes (e.g., retardation, impairment, apoptosis) induced by pharmacological treatments of cells grown in vitro, or by cell interactions (attachment, differentiation, and proliferation) with biomaterial substrates.

ACKNOWLEDGMENTS

The authors wish to thank Chris Owen for assistance with Raman spectral measurements and proof-reading, and Aaron Rae for assistance with the flow cytometry experiments. We are

also grateful to Ioan Notingher for helpful discussions concerning data analysis. RJS acknowledges the funding of the Rothermere Foundation, Canadian Centennial Scholarship Fund, and the National Science and Engineering Research Council of Canada.

REFERENCES

Boydston-White S, Gopen T, Houser S, Bargonetti J, Diem M. 1999. Infrared Spectroscopy of Human Tissue. V. Infrared Spectroscopic Studies of Myeloid Leukemia (ML-1) Cells at Different Phases of the Cell Cycle. Biospectroscopy 5:219–227.

Boydston-White S, Chernenko T, Regina A, Miljkovic M, Matthäus C, Diem M. 2005. Microspectroscopy of single proliferating HeLa cells. Vibrational Spectrosc 38:169–177.

Boydston-White S, Romeo M, Chernenko T, Regina A, Miljkovic M, Diem M. 2006. Cell-cycle-dependent variations in FTIR micro-spectra of single proliferating HeLa cells: Principal component and artificial neural network analysis. Biochim Biophys Acta 1758:908–914.

Carbonaro-Hall D, Williams R, Wu L, Warburton D, Zeichner-David M, MacDougall M, Tolo V, Hall F. 1993. G1 expression and multistage dynamics of cyclin A in human osteosarcoma cells. Oncogene 8:1649–1659.

Chan JW, Taylor DS, Zwerdling T, Lane SM, Ihara K, Huser T. 2006. Micro-Raman spectroscopy detects individual neoplastic and normal hematopoietic cells. Biophys J 90:648–656.

Crow P, Barrass B, Kendall C, Hart-Prieto M, Wright M, Persad R, Stone N. 2005. The use of Raman spectroscopy to differentiate between different prostatic adenocarcinoma cell lines. Br J Cancer 92:2166–2170.

Hair JF. 1998. Multivariate data analysis. London: Upper Saddle River.

Jell G, Notingher I, Tsigkou O, Notingher P, Polak JM, Hench LL, Stevens MM. 2007. Bioactive glass-induced osteoblast differentiation: A noninvasive spectroscopic study. J Biomed Mater Res A DOI: 10.1002/jbm.a.31542.

Jordan MA, Thrower D, Wilson L. 1992. Effects of vinblastine, podophyllotoxin and nocodazole on mitotic spindles. Implications for the role of microtubule dynamics in mitosis. J Cell Sci 102(Pt 3):401–416.

Kawabata H, Germain RS, Vuong PT, Nakamaki T, Said JW, Koeffler HP. 2000. Transferrin receptor 2-alpha supports cell growth both in iron-chelated cultured cells and in vivo. J Biol Chem 275:16618–16625.

Krafft C, Knetschke T, Funk RH, Salzer R. 2006. Studies on stress-induced changes at the subcellular level by Raman microspectroscopic mapping. Anal Chem 78:4424–4429.

Krishan A. 1975. Rapid flow cytofluorometric analysis of mammalian cell cycle by propidium iodide staining. J Cell Biol 66:188–193.

Lieber CA, Mahadevan-Jansen A. 2003. Automated method for subtraction of fluorescence from biological Raman spectra. Appl Spectrosc 57:1363–1367.

Lopes MA, Knowles JC, Kuru L, Santos JD, Monteiro FJ, Olsen I. 1998. Flow cytometry for assessing biocompatibility. J Biomed Mater Res 41:649–656.

Maquelin K, Choo-Smith LP, van Vreeswijk T, Endtz HP, Smith B, Bennett R, Bruining HA, Puppels GJ. 2000. Raman spectroscopic method for identification of clinically relevant microorganisms growing on solid culture medium. Anal Chem 72:12–19.

Matthäus C, Boydston-White S, Miljkovic M, Romeo M, Diem M. 2006. Raman and infrared microspectral imaging of mitotic cells. Appl Spectrosc 60:1–8.

Mourant JR, Gibson RR, Johnson TM, Carpenter S, Short KW, Yamada YR, Freyer JP. 2003a. Methods for measuring the infrared spectra of biological cells. Phys Med Biol 48:243–257.

Mourant JR, Yamada YR, Carpenter S, Dominique LR, Freyer JP. 2003b. FTIR spectroscopy demonstrates biochemical differences in mammalian cell cultures at different growth stages. Biophys J 85:1938–1947.

Notingher I, Verrier S, Romanska H, Bishop AE, Polak JM, Hench LL. 2002. In situ characterization of living cells by Raman spectroscopy. Spectrosc Int J 16:43–51.

Notingher I, Verrier S, Haque S, Polak JM, Hench LL. 2003. Spectroscopic study of human lung epithelial cells (A549) in culture: Living cells versus dead cells. Biopolymers 72:230–240.

Notingher I, Bisson I, Bishop AE, Randle WL, Polak JM, Hench LL. 2004a. In situ spectral monitoring of mRNA translation in embryonic stem cells during differentiation in vitro. Anal Chem 76:3185–3193.

Notingher I, Jell G, Lohbauer U, Salih V, Hench LL. 2004b. In situ non-invasive spectral discrimination between bone cell phenotypes used in tissue engineering. J Cell Biochem 92:1180–1192.

Omberg KM, Osborn JC, Zhang SL, Freyer JP, Mourant JR, Schoonover JR. 2002. Raman spectroscopy and factor analysis of tumorigenic and non-tumorigenic cells. Appl spectrosc 56:813–819.

Owen CA, Selvakumaran J, Notingher I, Jell G, Hench LL, Stevens MM. 2006. In vitro toxicology evaluation of pharmaceuticals using Raman micro-spectroscopy. J Cell Biochem 99:178–186.

Pedrali-Noy G, Spadari S, Miller-Faures A, Miller AO, Kruppa J, Koch G. 1980. Synchronization of HeLa cell cultures by inhibition of DNA polymerase alpha with aphidicolin. Nucleic Acids Res 8:377–387.

Raman CV, Krishnan KS. 1928. A new type of secondary radiation. Nature 121:501–502.

Short KW, Carpenter S, Freyer JP, Mourant JR. 2005. Raman spectroscopy detects biochemical changes due to proliferation in mammalian cell cultures. Biophys J 88: 4274–4288.

Verrier S, Notingher I, Polak JM, Hench LL. 2004. In situ monitoring of cell death using Raman microspectroscopy. Biopolymers 74:157–162.

Wang S, El-Deiry WS. 2004. Apoptosis Signaling in Normal and Cancer Cells. In: Stein GS, Pardee AB, editors. Cell cycle and growth control: Biomolecular regulation and cancer. New Jersey: John Wiley & Sons. pp. 497–524.

Wold S, Esbensen K, Geladi P. 1987. Principal components analysis. Chemo Intell Lab Sys 2:37–52.

Wolthuis RT, Bakker Schut TC, Caspers PJ, Buschman HPJ, Römer TJ, Bruining HA, Puppels GJ. 1999. Raman spectroscopic methods for in vitro and in vivo tissue characterization. In: Mason WT, editor. Fluorescent and luminescent probes for biological activity: A practical guide to technology for quantitative real-time analysis. London: Academic Press. pp. 433–455.

Xynos ID, Edgar AJ, Buttery LD, Hench LL, Polak JM. 2000a. Ionic products of bioactive glass dissolution increase proliferation of human osteoblasts and induce insulin-like growth factor II mRNA expression and protein synthesis. Biochem Biophys Res Commun 276: 461–465.

Xynos ID, Hukkanen MV, Batten JJ, Buttery LD, Hench LL, Polak JM. 2000b. Bioglass 45S5 stimulates osteoblast turnover and enhances bone formation in vitro: Implications and applications for bone tissue engineering. Calcif Tissue Int 67:321–329.

Yang R, Muller C, Huynh V, Fung YK, Yee AS, Koeffler HP. 1999. Functions of cyclin A1 in the cell cycle and its interactions with transcription factor E2F-1 and the Rb family of proteins. Mol Cell Biol 19:2400–2407.