NOTE

Quartz microfluidic chip for tumour cell identification by Raman spectroscopy in combination with optical traps

Sebastian Dochow · Claudia Beleites · Thomas Henkel ·
Günter Mayer · Jens Albert · Joachim Clement ·
Christoph Krafft · Jürgen Popp

Abstract Three important technical innovations are reported here towards Raman-activated cell sorting. Firstly, a microfluidic chip made of quartz is introduced which integrates injection of single cells, trapping by laser fibres and sorting of cells. Secondly, a chip holder was designed to provide simple, accurate and stable adjustment of chips, microfluidic connections and the trapping laser fibres. The new setup enables to the collection of Raman spectra of single cells at 785 nm excitation with 10 s exposure time. Lastly, a new type of modelling the various background contributions is described, improving Raman-based cell identification by the classification algorithm linear discriminant analysis. Mean sensitivity and specificity determined by iterated 10-fold cross validation were 96 and 99 %, respectively, for the distinction of leucocytes extracted from blood, breast cancer cells BT-20 and MCF-7, and leukaemia cells OCI-AML3.

Keywords Chemometrics/statistics · IR spectroscopy/ Raman spectroscopy · Microfluidics/microfabrication · Single cell analysis · Flow cytometry

S. Dochow · C. Beleites · T. Henkel · G. Mayer · J. Albert ·
C. Krafft · J. Popp
Institute of Photonic Technology, Albert Einstein Straße 9,
07745 Jena, Germany

J. Clement
Department of Internal Medicine II, University Hospital Jena,
Erlanger Allee 101,
07740 Jena, Germany

J. Popp (✉)
Institute of Physical Chemistry and Abbe Center of Photonics,
University Jena, Helmholtzweg 4,
07743 Jena, Germany
e-mail: juergen.popp@uni-jena.de

Introduction

Raman-activated cell sorting (RACS) offers prospects to complement the widely applied fluorescence-activated cell sorting. The term RACS denotes that Raman spectroscopy provides information to characterize and identify cells. Advantages of Raman spectroscopy include the obtaining of highly specific, fingerprint-like information without external labels in a non-destructive fashion, so the method is also compatible with live cells [1]. Limits of RACS are longer acquisition times and lower throughput due to weaker Raman signal relative to fluorescence emission. A key element of RACS is the microfluidic infrastructure. Our previous manuscript described two setups to trap single cells in microfluidic environments, collect Raman spectra and classify five cell types based on their spectral fingerprint [2]. The integrated setup used a microfluidic chip made of glass from which Raman spectra could only be obtained at 514 nm excitation. However, this excitation wavelength induces degradation of living cells at the required high intensities even at short exposure times. High Raman background in the fingerprint region below 1,500 cm^{-1}, which contains most of the diagnostically relevant spectral information, is generated in glass chips with 785 nm excitation. The Raman background is significantly lower for substrates made of quartz at near-infrared laser excitation. This was demonstrated in another setup using a quartz capillary as microfluidic channel from which Raman spectra could be obtained at 785 nm. Signal-to-noise ratios and classification accuracies were inferior compared with the integrated setup. Three important innovations are reported here. Firstly, the processing steps were transferred to quartz wafers to produce microfluidic chips made of quartz. Secondly, a chip holder was designed to integrate chips, microfluidic connections and the trapping laser fibres. Thirdly, a spectral data processing approach was introduced with improved

background correction. The performance of the new setup was tested for a tumour cell model that consisted of leucocytes extracted from blood, breast cancer cells BT-20 and MCF-7 and leukaemia cells OCI-AML3. Cells were identified based on Raman spectra by the classification algorithm linear discriminant analysis. Sensitivity and specificity were determined by iterated 10-fold cross validation, and the results were compared with previous data.

Experimental section

Design of quartz chip and chip holder

To enable Raman measurements of single cells in microfluidic chips at 785 nm excitation, quartz is preferred over other glass qualities, like borofloat33, due to the superior optical properties such as lower Raman background signals in the fingerprint region. Therefore, the processing procedures were optimized for amorphous quartz wafers (UV-grade fused silica; diameter, 100 mm) as described in detail elsewhere [3]. The chip comprises the same three main operation units as the previously presented borofloat33 chip (see Fig. 1).

A new chip holder enabled fluidic connections and direct fibre laser coupling at any optical setup. The holder was designed as described elsewhere [4].

Raman spectroscopy

The microfluidic setup was installed at a Raman microscopic system (RXN1, Kaiser Optical Systems, USA) equipped with a 785-nm single-mode diode laser (Xtra, Toptica, Germany) for excitation (Fig. 1, right). To reduce the refractive index gradient, the space between the objective and the chip was filled with water, and a ×60, NA 1.0 water immersion objective (Nikon, Japan) was used. Single-cell suspensions were prepared as described before [2] and injected at a flow rate of 1nl/s. Altogether, 405 cell spectra were collected at an exposure time of 10 s each: 100 leucocytes "Leuco", 104 "BT-20" and 100 "MCF-7" breast tumour cells, and 101 acute myeloid leukaemia cells "OCI-AML3". All cells were trapped manually by the 1,070-nm trapping lasers, and the acquisition of the Raman spectra was started. Afterwards, the cells were released and the next cell was trapped. In addition, 21 spectra were collected without cells and used for background compensation as described in Section 2.3.

Data analysis

Pre-processing The data were imported into R using package hyperSpec. The spectral range was cut to 650–1,800 cm^{-1}. Next, 17 spectra were excluded from further analysis (thereof, 15 with cosmic ray spikes, two grossly deviating from all other spectra). After automatic linear baseline correction, a principal component analysis (PCA) model was calculated for the 21 background spectra, and the first four principal components (without centring) were used to model these contributions. As the loadings were rather noisy while the signals are known to be rather broad, the loadings were smoothed with a fourth-order Savitzky–Golay filter (window width, 201 points or 60 cm^{-1}). The smoothed loadings together with an offset and a linear polynomial term served as reference "spectra" of background components for an extended multiplicative signal correction (EMSC) of the cell spectra using the EMSC implementation in package cbmodels. As no reference cell spectra were available, the resulting spectra have a mean intensity of 0, and the EMSC correction could not include normalization to the cell spectra intensity. Instead, the resulting spectra were offset corrected setting the 5th percentile of intensity of each spectrum to 0 and then normalized by the median intensity of the spectrum. The spectral regions below 810 and between 1,497 and
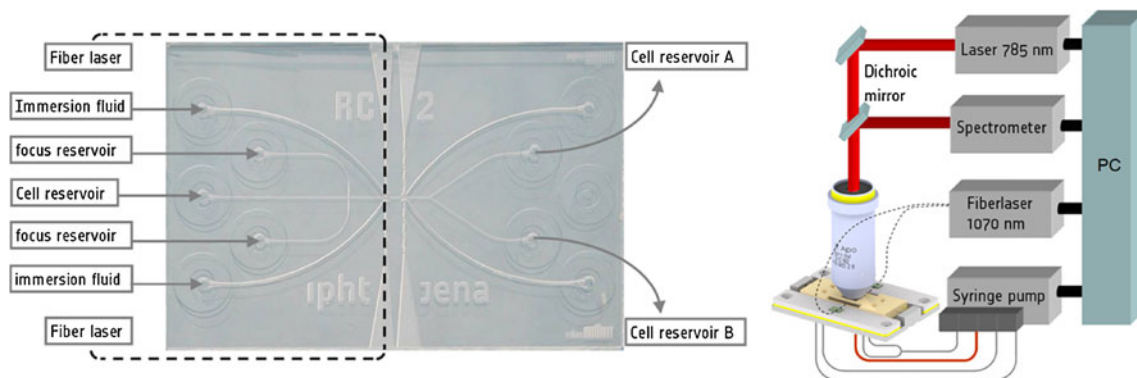

Fig. 1 *Right*: Scheme of the integrated setup with quartz chip, chip holder, micro fluidic pumps, fibre lasers and a Raman micro-spectrometer. *Left*: Raman chip 2 (*RC2*) made from quartz encompassing a channel for cell injection, two channels for hydrodynamic flow focusing, four channels for laser fibres, two channels for cell sorting and channels for immersion fluid

1,597 cm^{-1} contained residual contributions of the quartz signal and the trapping laser and were excluded from subsequent classification and validation.

Classification models Linear discriminant analysis (LDA) using 25 partial least square (PLS) latent variables as input was chosen as classifier (cbmodels, relying on MASS and pls). As a final step, only predictions were accepted that reached a posterior probability of at least 99 % according to the LDA, thus considering ca. 5 % of the predictions too uncertain.

Validation The performance of the classifiers was measured using a 100 times-iterated 10-fold cross validation, and validation folds were stratified with respect to the classes. In general, all pre-processing that depends on multiple or all spectra of a data set should be done *inside* the cross-validation loop, i.e. the respective parameters need to be calculated for each training set. In our case, this applies to
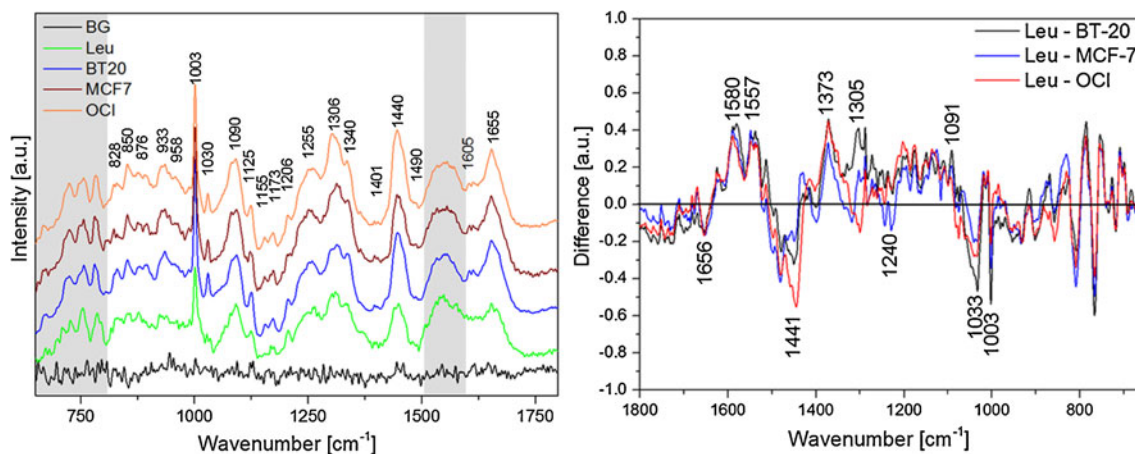
1. Estimation of PCA background components
2. Centring of the spectra matrix
3. The PLS model

Consequently, PCA estimation of the background components and all further pre-processing steps were done inside the cross-validation loop.

Results and discussion

The Raman spectra of the cells grossly agree with spectra from previous publications ([2] and references cited therein). The signal-to-noise ratio is increased by one order of magnitude, from 2.8 to 33. Spectral contributions are assigned to proteins, nucleic acids and lipids. A difference spectrum was calculated between leucocytes and all other cell line mean spectra (Fig. 2, top).

Difference bands near 1,580, 1,557, and 1,373 cm^{-1} have been observed before [2]. Although these differences are small, the reproducibility is consistent with their biological significance. Exact agreement is impaired by different signal-to-noise ratios and confounding signals of substrates and trapping lasers. An assignment of these difference bands is not straightforward. Spectral contributions of nucleic acids can be excluded because key marker bands near 1,100 and 1,092 cm^{-1} are not evident. Spectral contributions of lipids can be excluded because key marker bands near 1,440 and 1,300 cm^{-1} of the fatty acid moieties are missing. Spectral contributions of proteins are unlikely because key marker bands near 1,660 and 1,260 cm^{-1} of amide bands are missing. The observed difference bands point to the fact that other biomolecules besides nucleic acids, lipids and proteins significantly contribute to the Raman spectra and difference spectra between cells. Further Raman studies are under progress to assign these biomolecules.

Fig. 2 *Top left*: Mean spectra of the data after applying the EMSC processing (The background was multiplied by a factor of 100). Detailed band assignment can be found elsewhere [2]. *Top right*: Difference spectra from the mean spectrum of all cell lines subtracted by the mean leucocyte spectrum. *Bottom*: Average confusion matrix over 100 iterations, sensitivity and specificity if background is not independent. The values are given in percentages relative to the true number of spectra. The overall accuracy for cell identification is 98 %

The table (Fig. 2, bottom) shows the obtained confusion matrix for the trained classifier. The cell classes are recognized with uniform high sensitivities between 97 and 100 % (average over the 100 iterations; observed variation, 5th to 95th percentile deviates max.±2 spectra). The lowest classification rate of 87 % was observed for the background spectra which constitute the smallest class of 21 spectra. One incorrect classification of the background decreases the sensitivity already from 100 to 95 % whereas one incorrect classification out of 100 cell spectra decreases the sensitivity by only 1 %. Improvement is straightforward as more background spectra need to be collected. Compared with the Raman results obtained using the microfluidic glass chip at 514 nm excitation [2], the overall cell identification accuracy could be improved from 94.9 to 98 %. This is due to the higher signal-to-noise ratio of the acquired spectra and the additional group "background". The additional group offers the advantage of assigning spectra with low signal to the background group instead of to an incorrect cell group. Furthermore, during real-time Raman-activated cell sorting, acquisition of a Raman spectrum of a cell which is assigned to background can be repeated and a reclassification may be performed.

Conclusions

The progress and novelty of the current work are a microfluidic quartz chip; a chip holder, which accommodates syringe pump connections and laser fibres for optical trapping; and data processing tools. The devices simplify the coupling and accomplish high stability during long-term data acquisition. We demonstrated that Raman spectra from single cells could be collected, and four cell types could be identified based on these spectra with a new type of classification algorithm. In a few cases, background spectra were assigned to cell classes, which in practise can easily be avoided using visible CCD camera information. More importantly, no cell spectrum was confused with a background spectrum.

One still challenging problem of the upright microscopic setup is the continuous injection of single cells into the trap. First, formalin-fixed tumour cells, such as BT20 and MCF7, tend to agglomerate after detachment from the surface by trypsin treatment. This is non-critical for the chips and the periphery, but for future unsupervised and automated cell sorting, this needs to be solved. A second issue is the sedimentation of cells in the bottom loop of the injection tube at low flow rates as depicted in Fig. 1 (red tube). The subsequent injection of aggregated cell clusters causes complications in optical trapping and automatic separation. Therefore, an inverse setup would be advantageous for this application.

References

1. Chan JW (2012) Recent advances in laser tweezers raman spectroscopy (LTRS) for label-free analysis of single cells. J Biophoton. doi:10.1002/jbio.201200143
2. Dochow S, Krafft C, Neugebauer U, Bocklitz T, Henkel T, Mayer G, Albert J, Popp J (2011) Tumour cell identification by means of Raman spectroscopy in combination with optical traps. Lab Chip 11 (8):1484–1490
3. Dochow S, Beleites C, Henkel T, Mayer G, Albert J, Krafft C, Popp J (2012) Quartz microfluidic chip for tumour cell identification by Raman spectroscopy in combination with optical traps. Lab Chip. doi:10.1039/C2LC41169E, Accepted
4. Dochow S, Uhlemann W, Krafft C (2012) Justierbare Aufnahmevorrichtung für mikrofluidische Chips mit einzukoppelnder optischer Faser. Germany Patent DE102010050676B3