FTIR imaging analysis

FTIR spectra were recorded on three cell lines W1, W1CR and A2780 grown in the presence or not of collagen I coating. The typical FTIR cell spectrum exhibits characteristic absorption bands in the 4000-900 cm^{-1} region. In the high wavenumber region ~2800-3000 cm^{-1}, peaks corresponding to lipids and proteins (CH_2/CH_3 modes) can be observed, phospholipids at ~1740 cm^{-1} (C=O vibrations), proteins at ~1650 cm^{-1} (Amide I mode, mainly C=O vibration) and ~1540 cm^{-1} (Amide II mode, N-H deformation + C-N vibration) of the O=C-N-H moiety, the ~1370-1470 cm^{-1} range covers the symmetric and asymmetric CH_3 bending modes, and finally the 1300-900 cm^{-1} frequency range encompassing the nucleic acids absorption at ~1220-1245 cm^{-1} (asymmetric PO_2- vibration) and ~1080 cm^{-1} (symmetric PO_2- vibration) as well as the polysaccharides absorption in the 900-1200 cm^{-1} window. In the latter region, C–O stretching (alcohol groups, 1080–1000 cm^{-1}), C–O–C stretching (glycosidic ether linkage, symmetric and coupled modes appear around 1070–1030 cm^{-1} and asymmetric stretch typically near 1150–1160 cm^{-1}), and C–C skeletal stretching vibrations (1000–950 cm^{-1}) dominate.

In this study, second derivative spectra were computed and analyzed in the 1400-900 cm^{-1} spectral window representative of nucleic acids and polysaccharides fingerprint regions. It also overlaps with information from glycogen and collagen type I absorption bands. Band ratios have often been used to extract tendencies in spectral profiles. However, multivariate methods are now preferred as they grasp the variations over all the frequencies in a given region rather than a ratio of two bands. PCA is a chemometrics method that allows reducing the number of variables in a dataset while keeping as much information as possible. It focuses on variance, results in principal components that are new combined features made from the originals, and explained variance that indicates how much information each component retain.

W1 cells were compared to W1 cells grown on COL1 coating (W1COL1). The PCA score plot (PC1, 58.6% of explained variance vs PC3, 8.8% of explained variance) is shown in Fig. 1 g while the loading vector of PC3 is displayed in Supplementary figure 3a. PC3 was chosen as it discriminates best the two datasets. The positive peaks of PC3 (corresponding to W1 control cells) and its negative peaks (corresponding to W1 COL1 cells) are also indicated. A first observation showed that W1COL1 cells exhibit absorption bands in the carbohydrates, glycogen, nucleic acids and proteins, regions as was observed with A2780COL1 cells, but at different frequencies. The 17 negative peaks depicted are assigned to the below-mentioned macromolecules as follows:

(1) the carbohydrate / glycogen / glycoconjugate-rich region consisting of the bands at 926 cm^{-1} (carbohydrates / glycogen C–O ring vibration), 953 cm^{-1} (carbohydrates / glycogen + possible minor phosphate overlap), 1019 cm^{-1} (carbohydrates / glycogen C–O stretch), 1041 cm^{-1} (carbohydrates / glycogen C–O stretch), 1139 cm^{-1} (carbohydrates / glycogen C–O and C–O–C vibrations), 1155 cm^{-1} (carbohydrates / glycosidic C–O–C vibration), 1178 cm^{-1} (carbohydrates / glycoconjugates with minor phosphate contribution).

2) the phosphate / nucleic-acid / phospholipid region encompassing the bands at 1068 cm^{-1} (mainly symmetric (PO_2^-) vibrations of nucleic acids / phospholipids, 1241 cm^{-1} (mainly asymmetric (PO_2^-) of nucleic acids / phospholipids). This region can also reflect overlapping phosphate contribution at 980 cm^{-1} (phosphate-containing biomolecules + carbohydrate overlap), 1112 cm^{-1} (mixed phosphate + carbohydrate band), 1196–1200 cm^{-1} (phosphate / amide III mixed band), and 1270 cm^{-1} (phosphate / amide III mixed band).

In this region, the strong contribution of the 1068 cm^{-1} and 1241 cm^{-1} bands suggests probable DNA modifications for W1COL1 cells. However, it must be noted that the 1241 cm^{-1} my also contain information from sulfated GAGs (asymmetric sulfate vibrations).

3) the protein / lipid deformation / amide III region comprising of the bands at 1270 cm^{-1} (phosphate / amide III mixed band), 1289 cm^{-1} (amide III / protein + CH deformation), 1318 cm^{-1} (CH_2 twisting/wagging; protein/lipid contribution), 1349 cm^{-1} (CH bending / protein side-chain and lipid contribution), and 1375 cm^{-1} (CH_3 deformation / protein-lipid contribution). These bands are more consistent with amide III / CH bending / CH_2 wagging contributions.

To summarize, W1COL1 cells FTIR spectra in the 1400-900 cm^{-1} region are characterized by a strong carbohydrate/glycogen/glycoconjugate spectral profile in the 926–1178 cm^{-1} window, a phosphate/nucleic-acid component at 1068 cm^{-1} and 1241 cm^{-1} with probably some contribution from cell surface GAGs, and a protein/lipid amide III + CH deformation region in the spectral range 1270–1375 cm^{-1}.

W1COL1 cells exhibit differences in the asymmetric phosphate PO_2- vibration at 1241 cm^{-1} compared to 1220 cm^{-1} for the W1 control cells. The position of the asymmetric phosphate PO_2- vibration is more sensitive to conformation changes than the symmetric counterpart. This upshift can be due to several situations such as membrane phospholipid perturbation like a change in the environment of membrane phospholipid phosphate groups and/or nucleic-acid phosphodiester groups toward lower hydration and altered intermolecular interactions. DNA methylation marker bands are also visible in W1COL1 cells at 1041 cm^{-1} and 1155 cm^{-1} but not in the control W1 cells. The same observations were found for A2780COL1 cells indicating plausible DNA conformation changes for cells grown on COL1 coating.

A comparison between W1CR control cells and W1CR cells grown on COL1 coating was also conducted. The PCA score plot (PC1, 64.3% of explained variance vs PC3, 5.3% of explained variance) is shown in Fig. 1g, while the loading vector of PC1 is shown in Supplementary Figure 3 b. As previously, the negative peaks (characteristic of W1CR cells) and the positive peaks (characteristic of W1CRCOL1) are depicted. Sixteen positive peaks are identified characteristic of W1CRCOL1 cells.

(1) the carbohydrate / glycogen / glycoconjugate-rich region: Ten discriminating bands at 934, 980, 1004, 1024, 1069, 1100, 1108, 1132, 1148, and 1161 cm^{-1} were identified and are consistent with vibrational modes of glycogen and related carbohydrate/glycoconjugate structures, indicating that the observed spectral differences are likely driven primarily by changes in cellular carbohydrate-associated

biochemical composition. It should be noted that the 1004 cm^{-1} band can have some overlapping contribution from the phenylalanine ring mode.

2) the phosphate / nucleic-acid / phospholipid region: The bands observed at 1186 cm^{-1} and 1253 cm^{-1} are typically associated to phosphate-containing functional groups, especially PO_2^- stretching vibrations (symmetric and asymmetric, respectively) from nucleic acids (DNA/RNA phosphate backbone), phospholipids in cellular membranes or more generally, phosphorylated cellular components. The asymmetric PO_2^- stretching, more sensitive to DNA conformational changes, was seen to be upshifted compared to the W1CR cells where two lower frequency bands were observed at 1221 and 1240 cm^{-1}. This is consistent with what was observed above for cells grown on collagen I coating. Furthermore, the 1240 cm^{-1} DNA peak was found to be very weak in the control cells and absent in W1CRCOL1 cells, where the 1253 cm^{-1} is stronger. DNA methylation marker bands were found at 1148 (weak intensity) and 1365 cm^{-1} in W1CRCOL1 cells. It is possible that some DNA conformational changes could be the explanation.

3) the protein / lipid deformation / amide III region: this comprises bands at 1281, 1305, 1340 and 1365 cm^{-1} generally associated with amide III vibrations, CH_2/CH_3 deformation, and related protein/lipid modes. The 1340 cm^{-1} (and to a lesser extent the 1281 cm^{-1}) band could contain contributions from side-chain or collagen/cytoskeletal protein modes, indicating measurements with the presence of the COL1 coating, as observed in the case of A2780COL1 cells, but not W1COL1 cells.

A preliminary analysis by (Hierarchical Cluster Analysis) HCA of A2780 control cells and A2780 seeded on collagen I coating (A2780COL1) in the above-mentioned region, showed a clear delineation of the two cell populations with a high heterogeneity between the two clusters (Supplement Fig. 4), indicating large spectral differences. However, although HCA gives a good partition, it does not allow explaining the underlying origins of the discrimination. Therefore, PCA analysis was performed on the same spectra, first to confirm the results and then to identify the molecular changes leading to this difference.

Fig. 1i shows the score plot of the principal component 1 (PC1, 51.9% of explained variance) versus principal component 3 (PC3, 9.2% of explained variance). It shows a clear discrimination between A2780 cells (open circles) and A2780COL1 cells (crosses), confirming the results obtained by the HCA. In order to characterize the molecular vibrations that account for this difference, Supplement Figure 3C shows the loadings of PC1 as it is the most discriminant. Since second derivatives were used for this computation, the negative score values of PC1 correspond to its positive peaks, and are thus characteristic of the A2780 cell line. Inversely, the positive scores of PC1 correspond to its negative peaks and are characteristic of A2780COL1 cells. Focusing on the negative peaks can reveal the peaks characteristic of the A2780 cells grown on collagen compared to their control counterpart. These peaks can be grouped in three regions:

(1) the carbohydrate / glycogen / glycoconjugate-rich region comprising the bands at 936 cm^{-1} (carbohydrates / glycogen (C–O, ring vibration)), 974 cm^{-1} (carbohydrates / glycogen with possible minor phosphate overlap), 984 cm^{-1} (phosphate-containing biomolecules with a carbohydrate overlap), 1037 cm^{-1} (carbohydrates / glycogen (C–O stretch)), 1059 cm^{-1} (phosphate-associated band with carbohydrate overlap), 1151 cm^{-1} and 1161 cm^{-1} (carbohydrates / glycosidic C–O–C vibrations).

2) the phosphate / nucleic-acid / phospholipid region consisting of the bands at 1059 cm^{-1} (phosphate-associated / lower side of the symmetric (PO_2^-) region), 1082 cm^{-1} (PO_2^- symmetric vibration of nucleic acids / phospholipids), 1123 cm^{-1} (mixed phosphate and carbohydrate band), 1222 cm^{-1} (phosphate ester / amide III mixed band), 1236 cm^{-1} (PO_2-asymmetric vibration of nucleic acids / phospholipids), and 1246 cm^{-1} (phosphate / amide III mixed band). These bands are the strongest candidates for DNA/RNA phosphodiester and membrane phospholipid phosphate contributions.

3) the protein / lipid deformation / amide III region encompassing bands at 1281 cm^{-1} (amide III / protein + CH deformation), 1303 cm^{-1} (CH_2 twisting/wagging; protein/lipid contribution), 1340 cm^{-1} (CH bending / protein side-chain and lipid contribution), 1366 cm^{-1} (CH_3 deformation / protein-lipid contribution), and 1380 cm^{-1} (CH_3 deformation, possible COO^- contribution). These reported bands are more consistent with amide III / CH bending / CH_2 wagging contributions. As indicated above, the 1222 cm^{-1} and 1246 cm^{-1} can also contain some overlapping vibrations due to protein contribution. It must be pointed out that the 1340 cm^{-1} band (to a lesser extent the 1281 cm^{-1} band) is quite characteristic of collagen, suggesting that some traces of collagen 1 coating could be detected.

These results tend to indicate that the A2780COL1 cells exhibit a stronger phosphate-rich structure around 1059–1246 cm^{-1} that can be assigned to nucleic acid changes. Other variations concern the carbohydrate/glycogen/glycoconjugate envelope in the 936–1161 cm^{-1} range and protein/lipid deformation and amide III region in the 1281–1380 cm^{-1} range.

An important observation concerns the presence of peaks that are assigned to glycogen and nucleic acids (asymmetric phosphate PO_2- vibration between 1200-1250 cm^{-1}. The DNA asymmetric vibration is upshifted to 1236-1246 cm^{-1} in A2780COL1 cells compared to A2780 cells where it is observed at 1204 cm^{-1}. The latter is quite low and assigned to phosphodiester groups of nucleic acids and/or phospholipid phosphate head groups. It suggests a strong hydrogen bonding or a high hydration of phosphate groups or a different phosphate microenvironment. This consequent shift of more than 30 cm^{-1} can be generally interpreted as a change in the phosphate head group environment toward weaker hydrogen bonding and lower hydration, thus leading to altered intermolecular interactions. Other observations concern the peaks at 1037, 1151 and 1366 cm^{-1} where it can be suggested DNA methylation on cytosine. The peaks assigned to DNA methylation are not present in the control cells spectra. This could reflect changes in DNA conformation when grown on COL1 coating.

To conclude, the set of discriminating FTIR bands spans three principal biochemical domains within the cellular fingerprint region. Peaks between 934 and 1148 cm^{-1} are predominantly associated with carbohydrate-related vibrations, consistent with glycogen and other glycoconjugate components, indicating that carbohydrate-associated biochemical variation is a major contributor to the observed spectral profile. Bands at 1186 and 1253 cm^{-1} fall within the phosphate-associated region and are consistent with contributions from nucleic acids and phospholipid head groups. The higher wavenumber features at 1281, 1305, 1340, and 1365 cm^{-1} are attributable to protein amide III and lipid deformation modes, suggesting additional changes in protein conformation and/or membrane lipid structure.