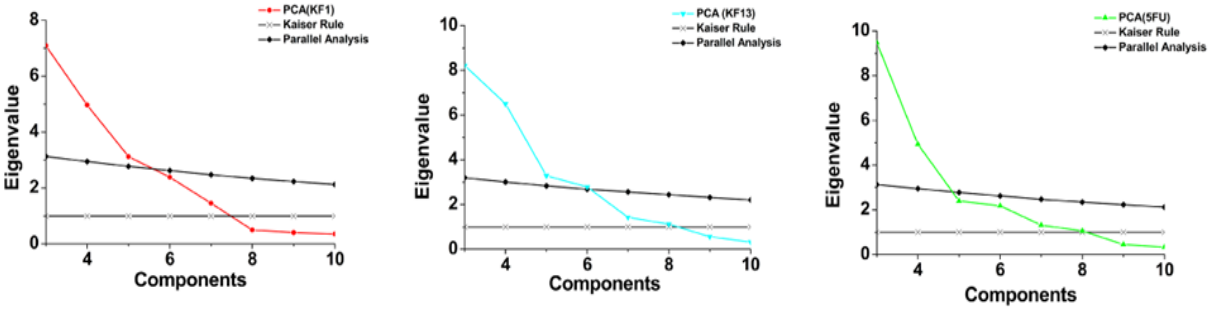
Investigating Cellular Responses to Novel Chemotherapeutics in Renal Cell Carcinoma using SR-FTIR Spectroscopy

C Hughes [a,b], M D Brown [b], N W Clarke [b], K R Flower [c] and P Gardner*[a]

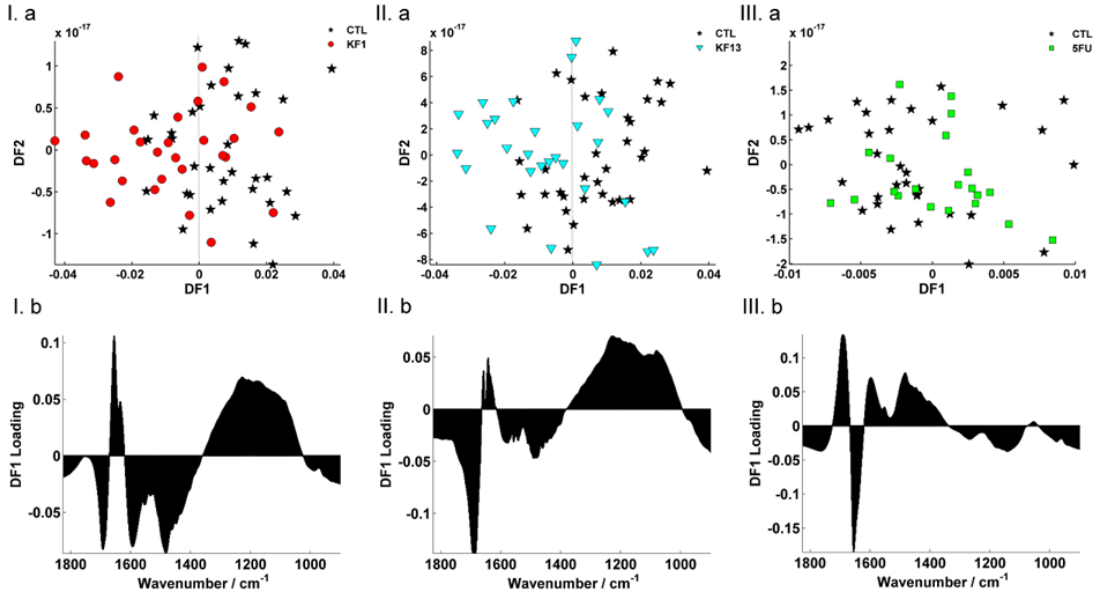
Supplementary Information

Supplementary Figure 1 | Eigenvalue scree plots for determining the number of PCs to use in LDA for (I) KF1, (II) KF13 and (III) 5FU respectively



PC-LDA was assessed as to whether overall differences could be maximised in an attempt to detect general trends. However, the 'ideal' number of principal components to use in a model can be subjective, although there are a number of ways to deduce this optimal number for a given dataset. Using theses guidelines (under Cattell/Horn/Kaiser) for the same datasets used in PCA, the highest suggested number of agreed components to input into LDA was calculated as four.

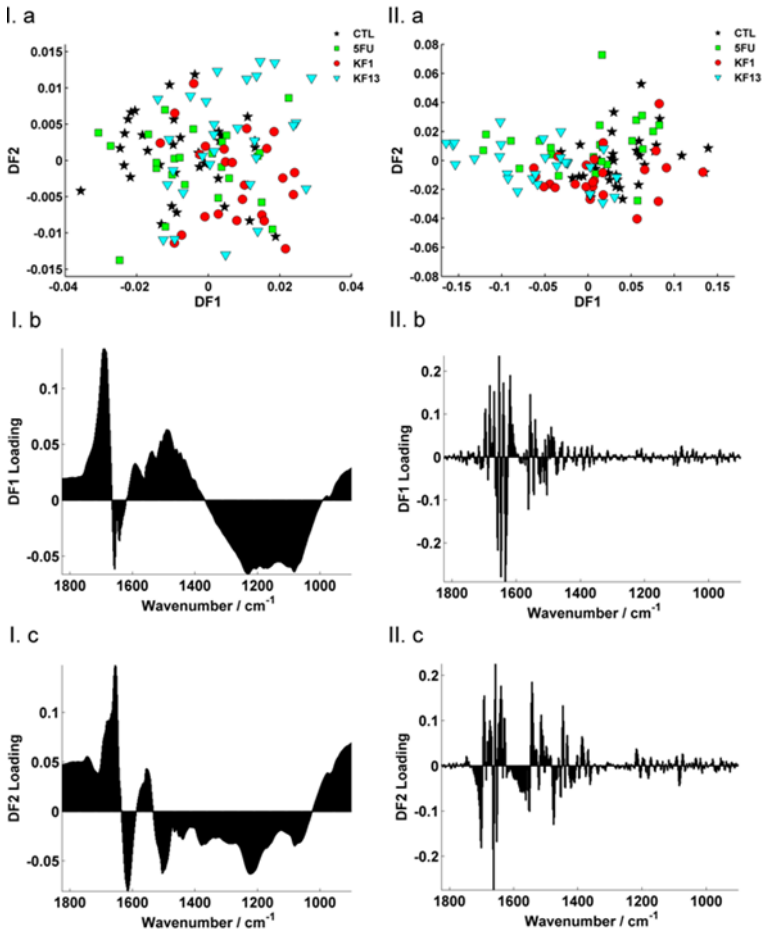
Supplementary Figure 2 | (a) LDA score plot using 4 PCs for KF1 (I), KF13 (II) and 5FU (III). (b) LDA loadings are displayed for DF1 respectively



Uisng the 4 PCs in LDA resulted in overlap of drug–treasted cell spectra and control spectra. Drug

treated cell spectra were primarily distributed in negative DF1 space and control spectra in positive DF1 space for KF1 (Fig. 6.3Ia) and KF13 (Fig. 6.3IIIa). No distinction could be seen for 5FU treated cells at this point. It could be argued that the cells are showing a greater response to the gold analogues than 5FU at this 24 hour check point, as separation in DF1 is marginally better. The weighting in the KF1 and KF13 loading plots of the 1300-900 cm^{-1} region, associated with carbohydrates and nucleic acids, appears greater than in the 5FU loading plot, where the region weighting is considerably less (Fig. 6.3b).

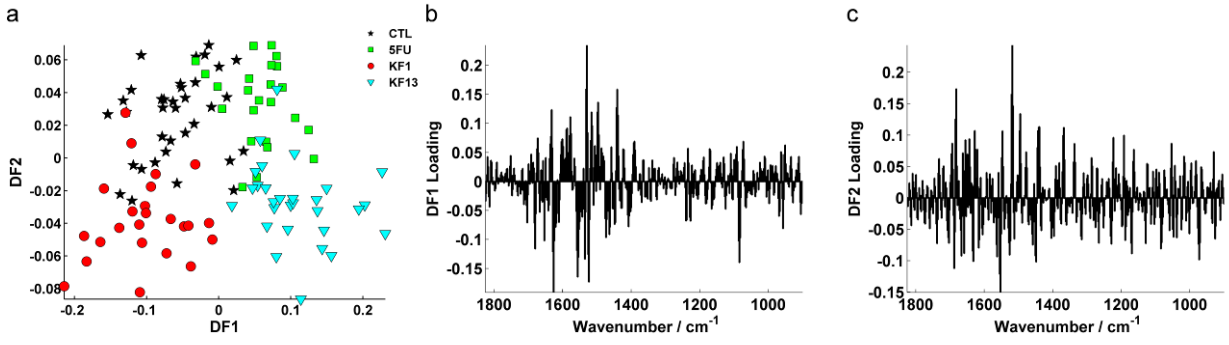
Supplementary Figure 3 | LDA score plot (a) and loadings (b, c) for non-derivative format for (I) 4 PCs (79% VE) and (II) 2nd derivative equivalent 9 PCs (90% VE)



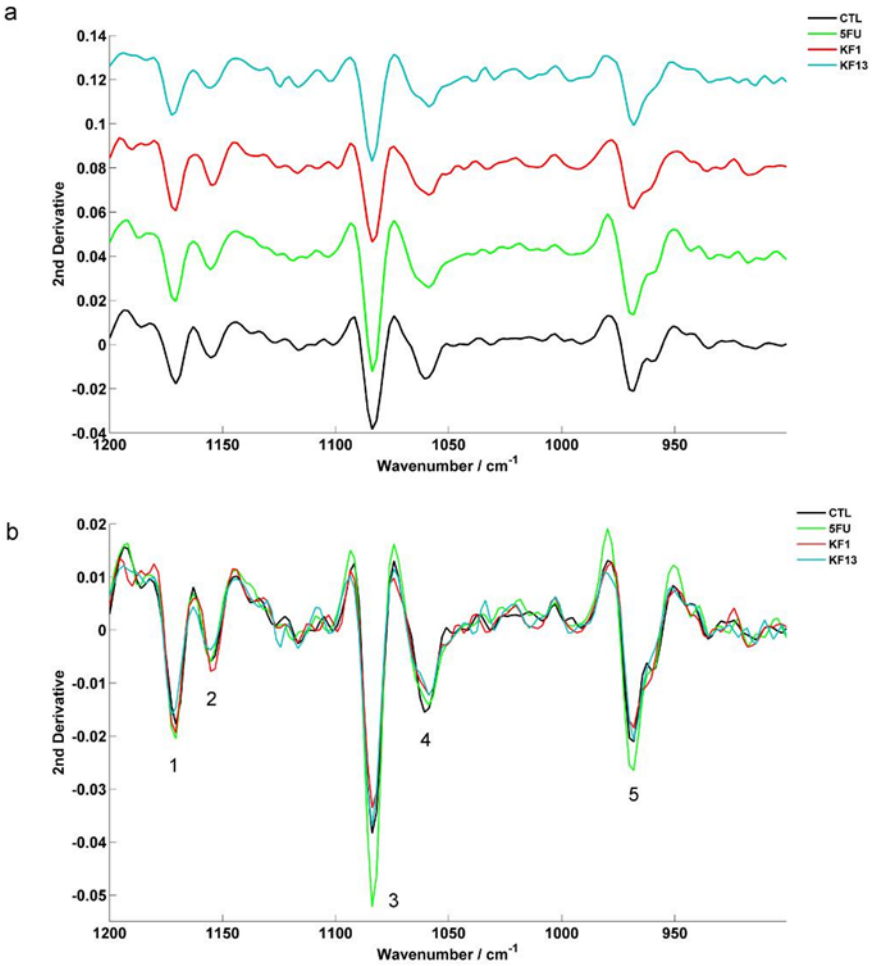
4 PCs explained 90% of variance for non-derivative data but only 79% in the 2nd derivative data (I). Therefore an equivalent 9 PCs was used for the 2nd derivative data resulting in further failure to separate the different groups (II).

Supplementary Figure 4 | Example of PC over-fitting

A comparison was made to illustrate the problems of over-fitting the data in PC-LDA. Using too many PCs (99% VE) illustrated by the 2nd derivative data LDA score plot (a) and the loadings for DF1 and DF2 (b, c) demonstrated that spectral data is non-resolvable due to the addition of noise-based PCs.



Supplementary Figure 5 | Qualitative analysis of mean group spectra



In further assessment of the mean second derivative spectra for each group (CTL, 5FU, KF1, and KF13; number of spectra, 'n', was 35, 24, 24, 29 respectively), there no apparent shift in peak minima (a). There were, however, five areas where variation was apparent when the derivative mean spectra were overlaid (b). The peak minima at these positions were at (1)1171; (2) 1155; (3) 1084; (4) 1061; (5) 968 cm^{-1}. These spectral regions are said to be associated with (1) ν_{as} (CO-O-C); (2) ν_s(C-O); (3) ν_s (PO_2^-); (4) ν_s (C-O) and (5) ν_s (C-O).