

Rapid Quantification of Tryptophan and Tyrosine in Chemically Defined Cell Culture Media using Fluorescence Spectroscopy.

Amandine Calvet, Boyan Li, and Alan G. Ryder.*

Nanoscale Biophotonics Laboratory, School of Chemistry, National University of Ireland, Galway, Galway, Ireland.

* Corresponding author: Email: alan.ryder@nuigalway.ie Phone: +353-91-492943.

Supplemental Information:

Contents

1) Spectral analysis of the EEM data: identification of fluorophores.....	1
2) Additional figures and data.....	2
3) Real-world sample variation of an eRDF used in industrial cell culture.....	6
4) Identification of samples affected by fluorescent contaminants.....	9
5) Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) analysis of dataset1.....	10
6) References.....	12

1) Spectral analysis of the EEM data: identification of fluorophores.

In the eRDF spectra, Tyr emission appears as a pair of bands: at the (excitation, emission pair of 230/305 nm and at 275/305 nm. The second-order fluorescence Tyr emission appears at the upper limit of the emission range of the EEMs collected (600 nm). The band present at the lowest excitation wavelength, 225 nm is due to fluorescence of the peptide moiety. Trp emission appears in the 300–475 nm range, which is comprised of a broad peak at 285/355 nm with a shoulder at 260/355 nm. From the excitation spectrum of pure Trp (Figure S-2a, supplemental information) we expect a single peak centered at 280 nm, but in these concentrated solutions, inner-filter effects (IFE) distort the excitation band generating a saddle at 275 nm. The effect is more obvious when one examines the spectral changes due to the total eRDF concentration variation on the model generated using PARAFAC2 (Figure 2, main manuscript).

Pyridoxine emission appeared at 325/395 nm and was, as expected, much weaker due to a combination of its low concentration (4.9 μM) and moderate fluorescence quantum yield of 0.11 ± 0.01 at $\text{pH} = 7/25^\circ\text{C}$ [1]. Pyridoxal ($\sim 2.5 \mu\text{M}$) is also present but, it emits at a shorter wavelength (370 nm), and has a very low quantum yield (0.048 at 25°C in buffer) and we expect that the emission is obscured by the much stronger Trp emission.[1] The very weak band at 375/520 nm corresponds to riboflavin fluorescence which is present at a

concentration of 0.5 μM (literature values give an emission at 520 nm for excitation at 270, 370 and 445 nm).[2] The emission is visible because the fluorescence quantum yield in water at pH 7 is relatively high (0.27) compared to the other fluorophores.[3] The very weak folic acid fluorescence band appears 355/440 nm and this agrees with literature values (see also Figure S-2 and 4).[4]

2) Additional figures and data.

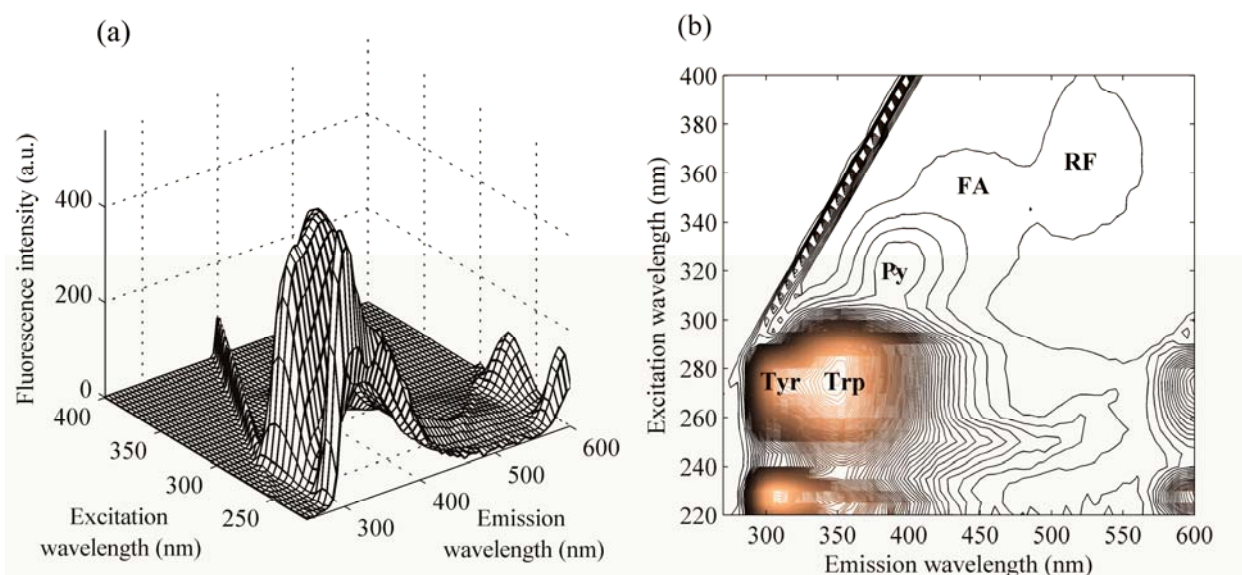


Figure S-1: EEM spectra of a 17.7g/L aqueous solution of eRDF plotted as: (a) a landscape plot, and (b) a contour map showing the location of the main fluorophores.

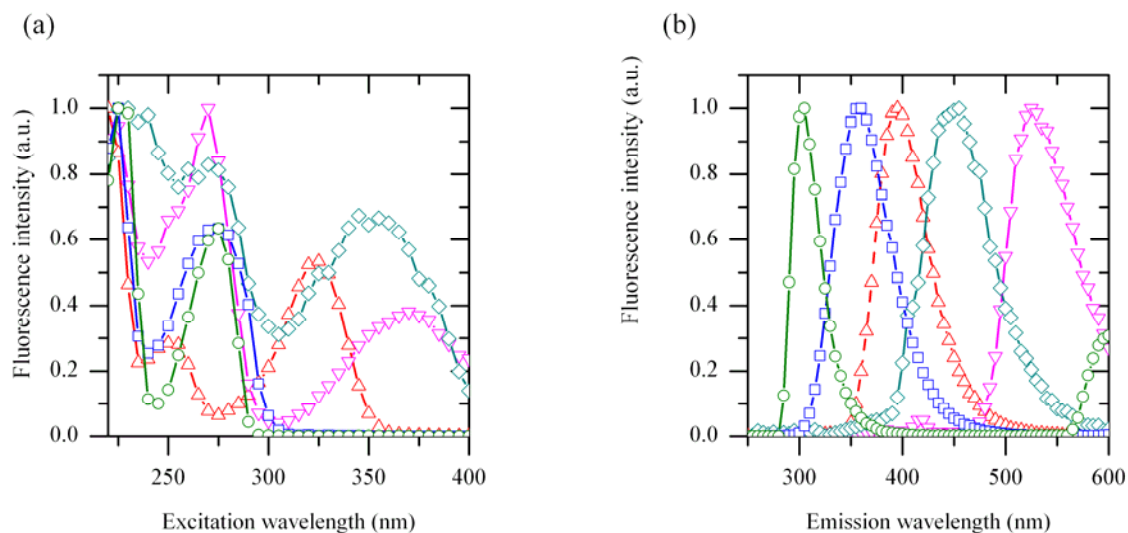


Figure S-2: Excitation and emission spectra extracted from the normalized EEM's of the pure fluorophores collected at 25°C in pH7 buffer. Tryptophan (\square , 90 μM), tyrosine (\circ , 480 μM), pyridoxine (Δ , 4.5 μM), riboflavin (∇ , 0.5 μM) in water and folic acid (\diamond , 19.9 μM) in 1.13 g/L NaHCO_3 . Rayleigh scattering was removed by replacing the data by a linear approximation.

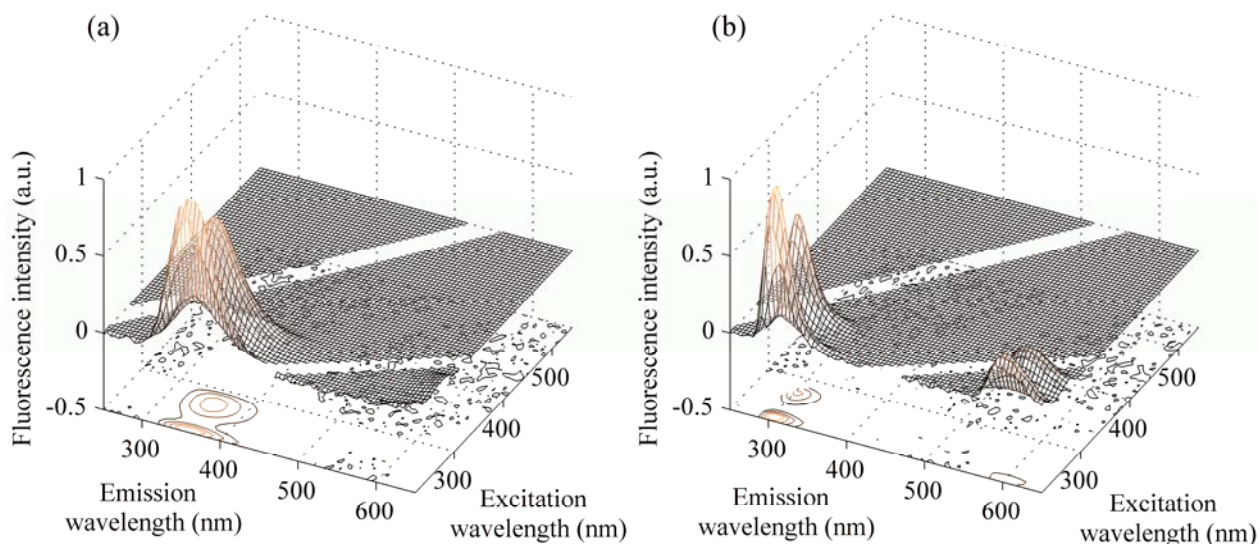


Figure S-3: Normalized EEMs of pure primary fluorophores: a) 90 μM Tryptophan, and b) 480 μM Tyrosine, in pH7 buffer at 25 $^{\circ}\text{C}$. Note that plot has a different emission axis orientation to Figure S-1.

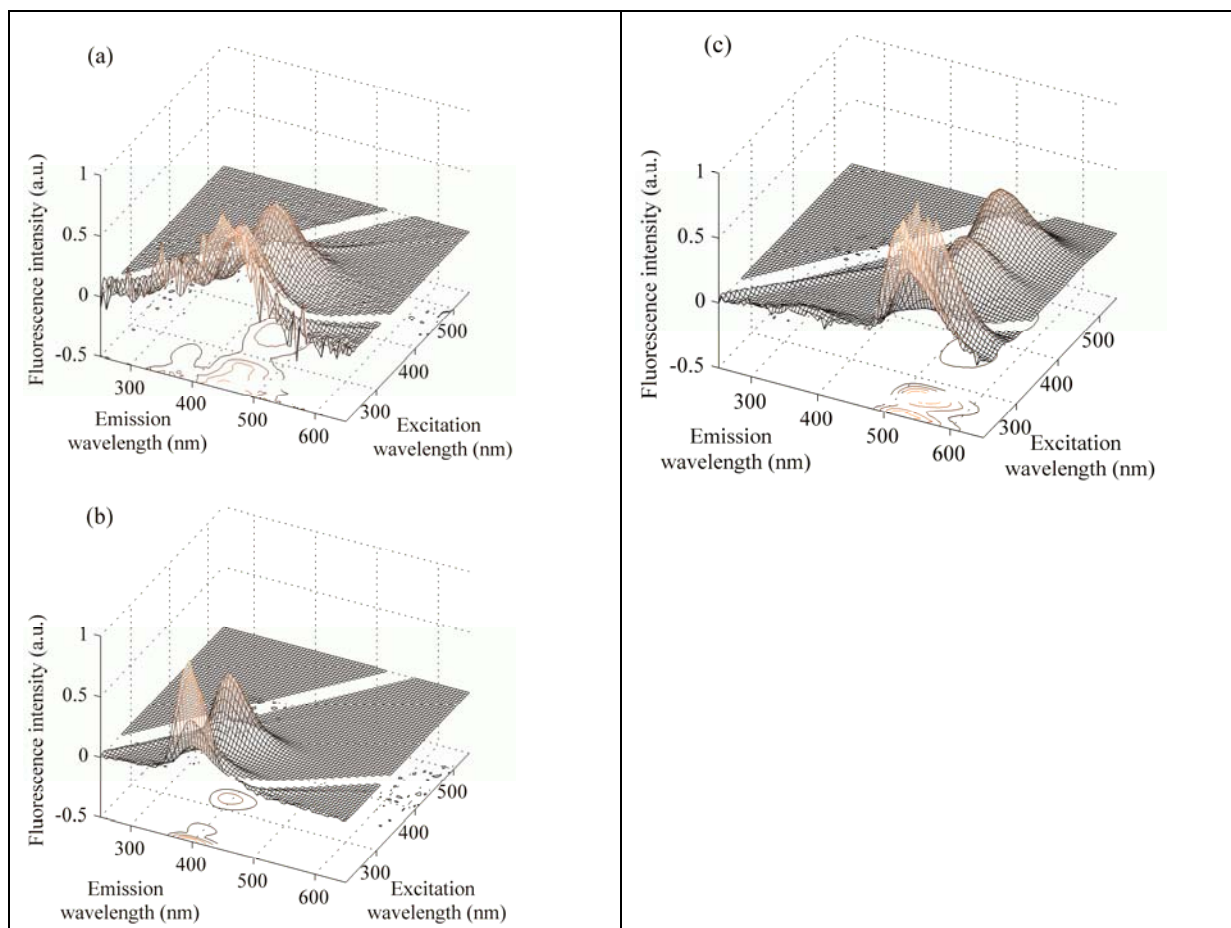


Figure S-4: Normalized EEMs of pure secondary fluorophores: a) Folic acid at 19.9 μM , b) Pyridoxine at 4.5 μM and c) Riboflavin at 0.5 μM in pH7 buffer at 25 $^{\circ}\text{C}$. Note that plots have a different emission axis orientation to Figure S-1.

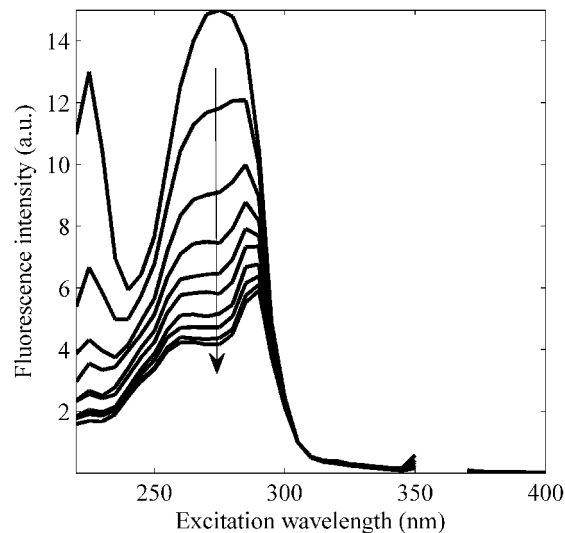


Figure S-5: Excitation profile of eRDF sample from dataset 1 measured at 360 nm emission and normalized based on the intensity at 335 nm. The arrow indicates the direction of increasing eRDF concentration.

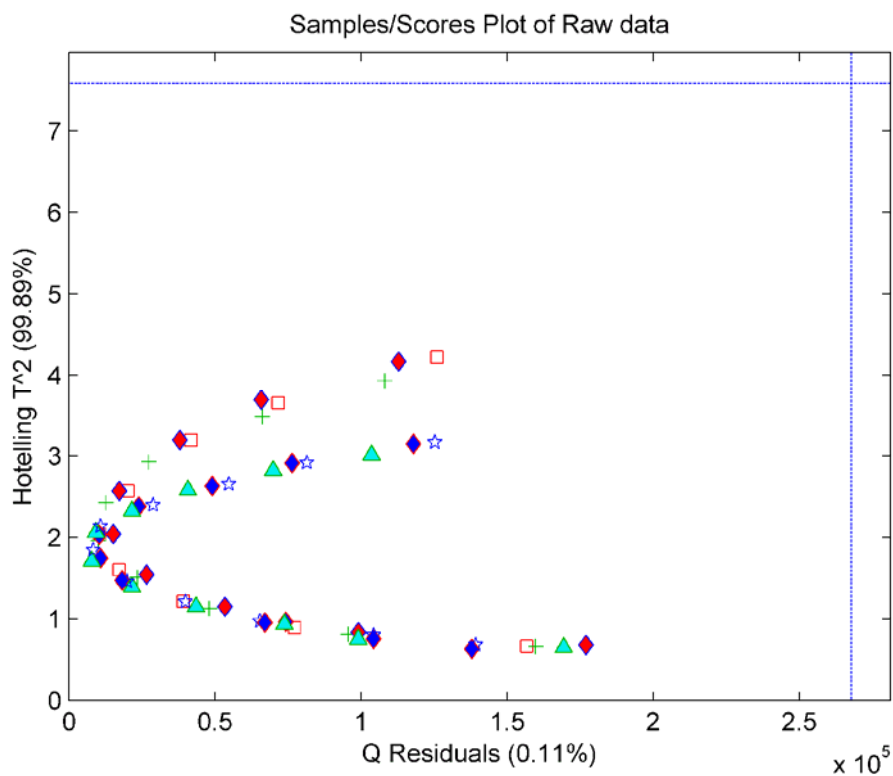


Figure S-6: Q residuals versus Hotelling T^2 plot from the Multi-way principal components analysis (MPCA) of all calibration samples (blue lines are the threshold limits at a 97% confidence limit). Trp spiked samples: \square , 1st series, $+$, 2nd series, \blacklozenge , 3rd series. Tyr spiked samples \blacklozenge , 1st series, \blacktriangle , 2nd series, \blackstar , 3rd series.

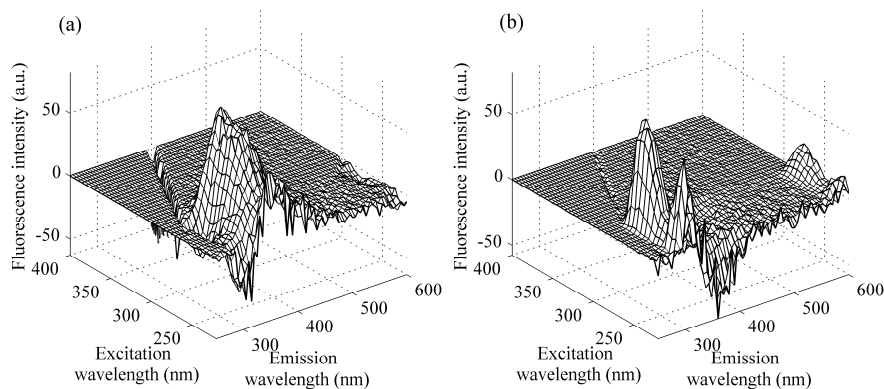


Figure S-7: Landscape plots showing the test matrix subtracted EEM (TMS-EEM) spectra used in the calibration modeling. eRDF media (17.7 g/L): spiked with: (a) Trp (to 1.2 times nominal concentration, 108.0 μM) and (b) Tyr (to 1.2 times nominal concentration, 575.6 μM).

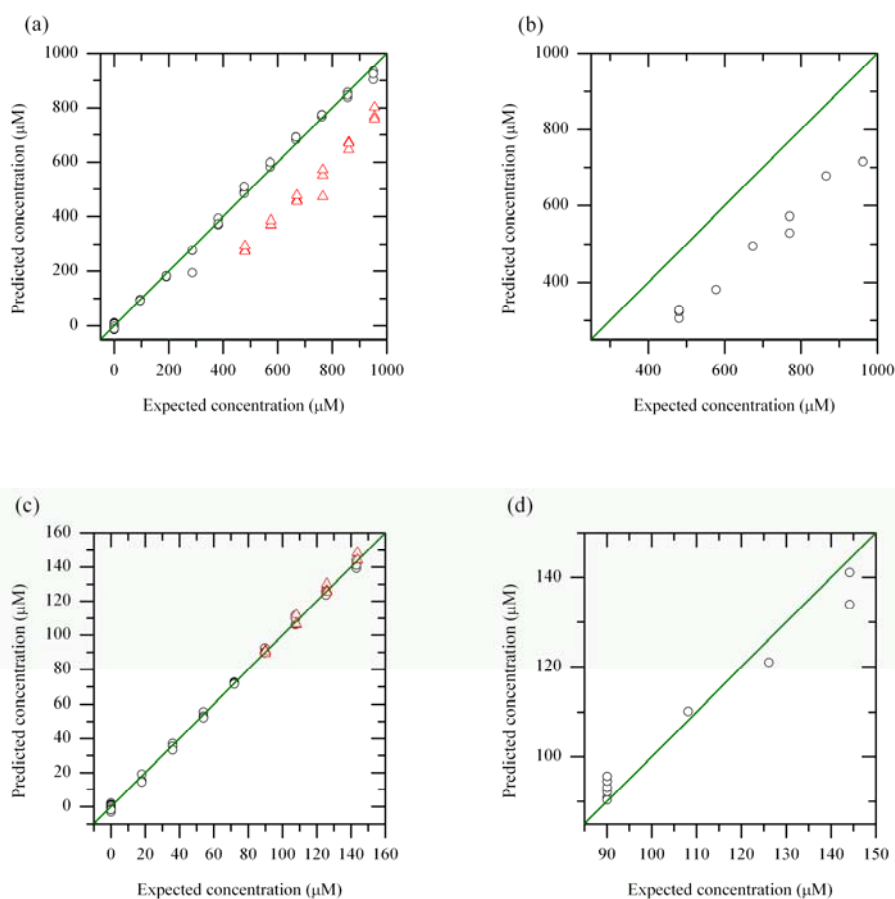


Figure S-8: Plots of predicted versus expected concentrations plots of the models calculated on the full EEM and described in table 4 for the Tyr calibration (\circ) and validation (Δ) samples (a), and prediction samples (b); similarly for the Trp calibration (\circ) and validation (Δ) samples (c), and prediction samples (d). The line represents the diagonal where the predicted concentration is equal to the expected concentration (formulation).

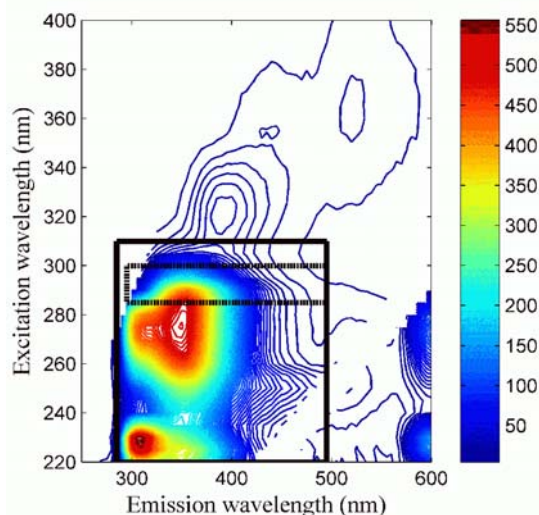


Figure S-9: Sub-sampled EEM spectral regions selected for Tyr and Trp prediction respectively, R1 (dashed) 220-310/285-495 and R2 (solid) 220-275/285-325 nm (excitation/emission).

Tryptophan						
Model type	LV	Pre-P ^(a)	RMSEC (μM)	RMSEV (μM)	RMSEP (μM)	RMSEPT (μM)
CalC1	3	C1S2,C	0.93	16.51	10.81	30.76
CalC3	10	C1S2,AS	0.73	4.03	4.50	2.96
CalC4	3	C1,C	2.16	4.23	5.59	1.91
Tyrosine						
Model type	LV	Pre-P (a)	RMSEC (μM)	RMSEV (μM)	RMSEP (μM)	RMSEPT (μM)
CalC2	14	C1S2,C	1.66	79.5	42.2	61.0
CalC3	8	C1S3,C	12.91	20.7	39.6	20.4
CalC5	6	C1S2,C	14.2	36.2	24.6	22.1

(a) PreP, Pre-processing on X, Y. Centering (C or C1), Scaling on mode 2 and/or 3 (S2 and/or 3), Autoscaling (AS).

Table S-1: Summary of results from NPLS modeling of restricted EEM data for the prediction of Trp/Tyr concentration.

3) Real-world sample variation of an eRDF used in industrial cell culture

We have previously demonstrated the use of EEM and chemometrics for the analysis of chemically defined cell culture media [5, 6] and discussed the typical variation that can be seen in industrially source media.

While the eRDF samples used here (eRDF-K samples) are not identical to the eRDF media used in the 2011 study (eRDF-AS11 samples), the formulation and EEM spectra (Figure S-10) are similar enough so as to be comparable in the context of discussing lot-to-lot variation.

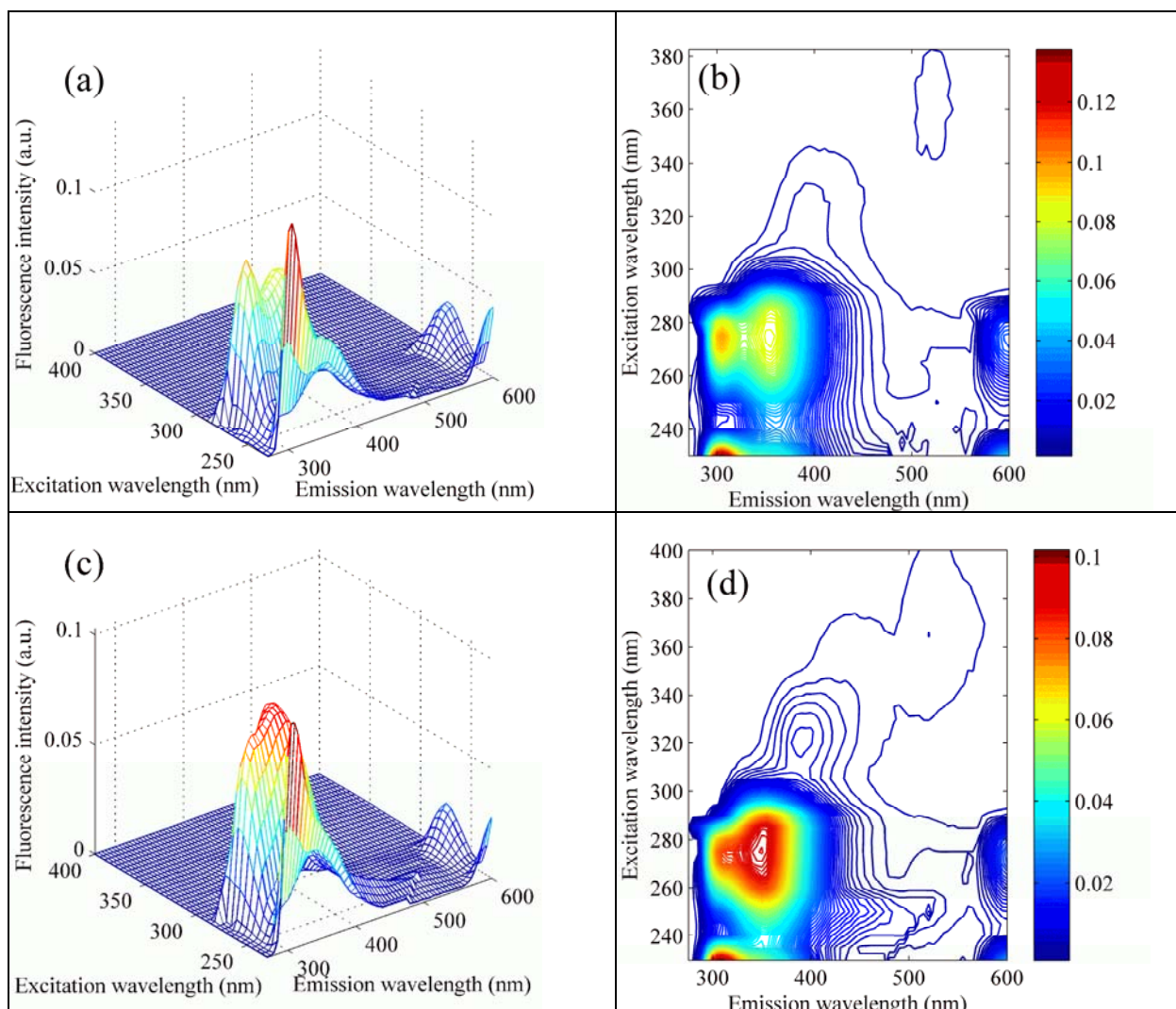


Figure S-10: Normalised (by area) EEM spectra of a) & b) eRDF-AS11, c) & d) eRDF-K samples.

U-PCA analysis

Undertaking an UPCA study allows one to compare the spectral variances obtained for the different sample sets. When we use the normalized EEM data we can get a clear picture of the true variance as shown in the scores plots below. The principal components scores plots show that the spectral variability induced by the Trp or Tyr spiking is similar in magnitude to the observed lot-to-lot variability from 21 samples of a commercially used eRDF.

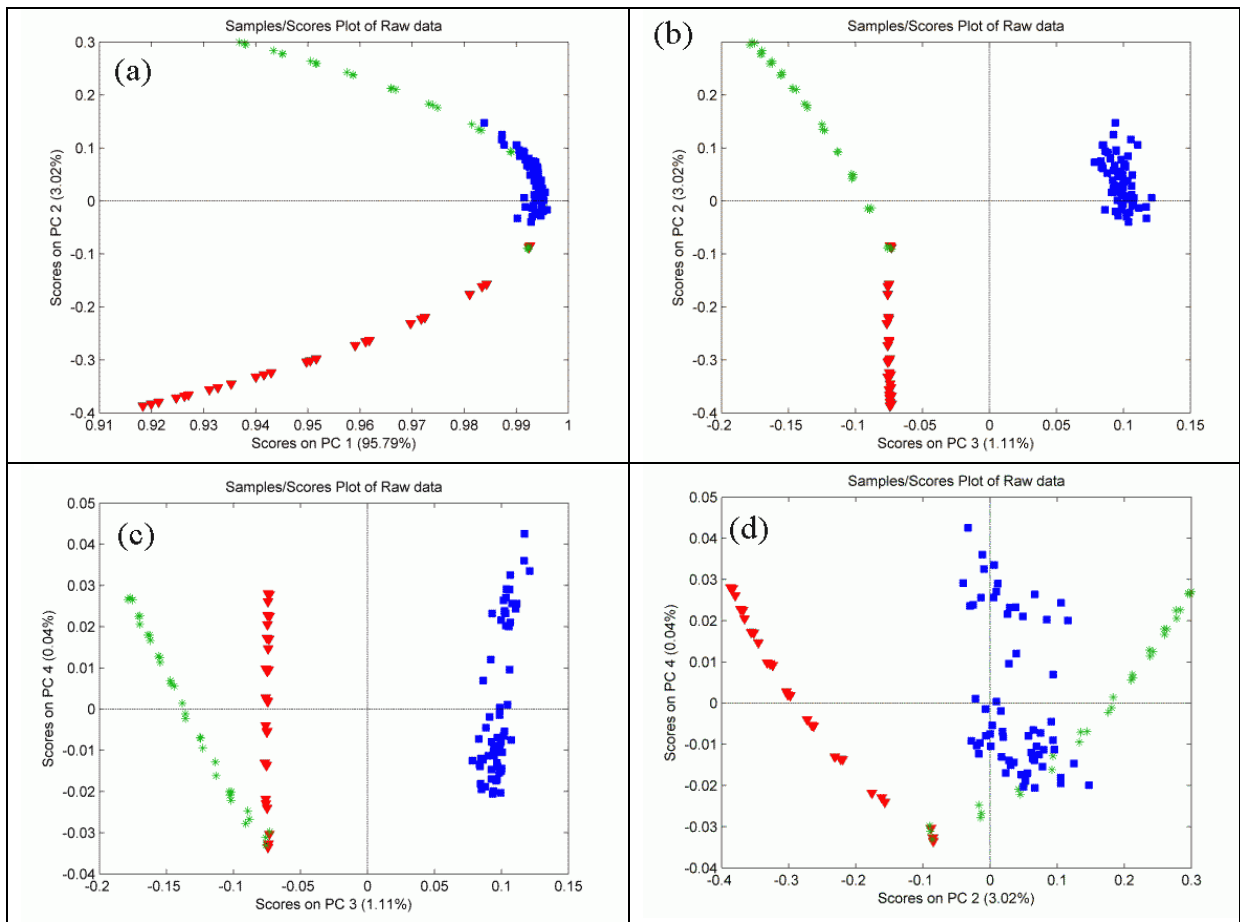


Figure S-11: Summary of UPCA study of eRDF samples from two sources: Scores plots showing sample variation: a) PC1 vs. PC2, b) PC1 vs. PC3, c) PC1 vs. PC4, d) PC2 vs. PC4, Sample #. Legend: Red (triangle): Trp spiked eRDF-K samples, Green (*): Tyr spiked eRDF-K samples, Blue (square): eRDF-AS11 samples.

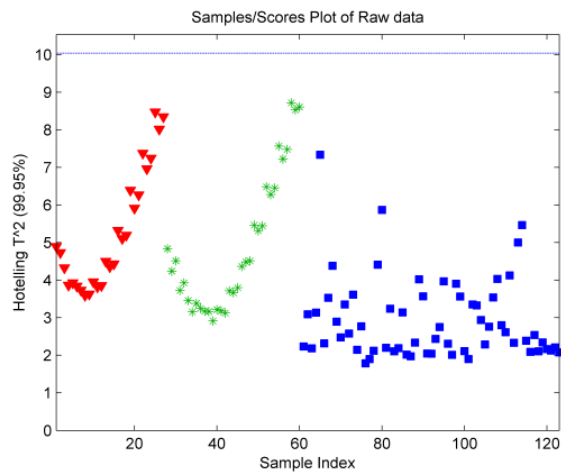


Figure S-12: Hotelling T^2 value vs. Sample # plot showing the magnitude of the sample variance.. Legend: Red (triangle): Trp spiked eRDF-K samples, Green (*): Tyr spiked eRDF-K samples, Blue (square): eRDF-AS11 samples.

The Hotelling T^2 statistic is a good indicator of sample variability (about a mean value) and when we look at Figure S-12 we see that the lot-to-lot variability is comparable to the variability induced by Trp or Tyr spiking.

Relative variation within group values:

To try and simplify further the comparison of the EEM data sets we calculated the overall relative variation within group values using a previously described method [7].

Code	Samples	Non normalised (%)	Normalised (%)
AS11	eRDF samples analysed in manuscript [6].	0.6	0.3
K	Kyokuto eRDF, spiked with both Trp and Tyr	9.8	6.2
Kw	Kyokuto eRDF only spiked with Trp	3.9	1.0
Ky	Kyokuto eRDF only spiked with Tyr	2.2	1.7

Table S-2: Summary of the relative variation within group (RVWG) values for the raw and normalized EEM spectra collected from the various samples sets.

The first thing to note is that the RVWG values are larger for the raw data as collected where there are large intensity variations. However, once we normalized the EEM data (as is widely used in chemometric modeling) we saw that the RVWG values decrease significantly. The RVWG for the multi-lot eRDF sample set is ~0.3% which is ~three times the RVWG for the Trp spiked sample set.

4) Identification of samples affected by fluorescent contaminants.

When using NPLS for quantitative analysis it is important to ensure that the samples being predicted are similar to those used to develop the calibration model. For these complex CD-media there is always a chance (despite rigorous GMP controls) of inadvertent preparation and or contamination of the sample by a fluorescent interferant. Thus in practical operation one would need to validate the sample quality prior to implementation of the NPLS model. This can be done simply using Unfolded Principal Components Analysis (UPCA) or equivalent chemometric analysis tools. In the example below we show how eRDF samples that have been changed via the deliberate irradiation using visible light can be easily discriminated from the calibration samples. Figure S-13 shows the Q residuals *versus* Hotelling T^2 plot for the UPCA model that included the irradiated eRDF samples and all of the calibration samples used to construct the NPLS models for Trp/Tyr prediction. It was evident that after irradiation for ~ one week, sufficient spectral changes were induced that could be easily discriminated by the Q- T^2 plot [8]. It is important to note that discrimination occurs mainly along the Q-residual axis which is a measure of the unexplained variance in these samples with respect to the UPCA model. In practice the boundaries of the Q- T^2 plot can be adjusted to suit the specific media being analyzed and the required accuracy of the Trp/Tyr prediction.

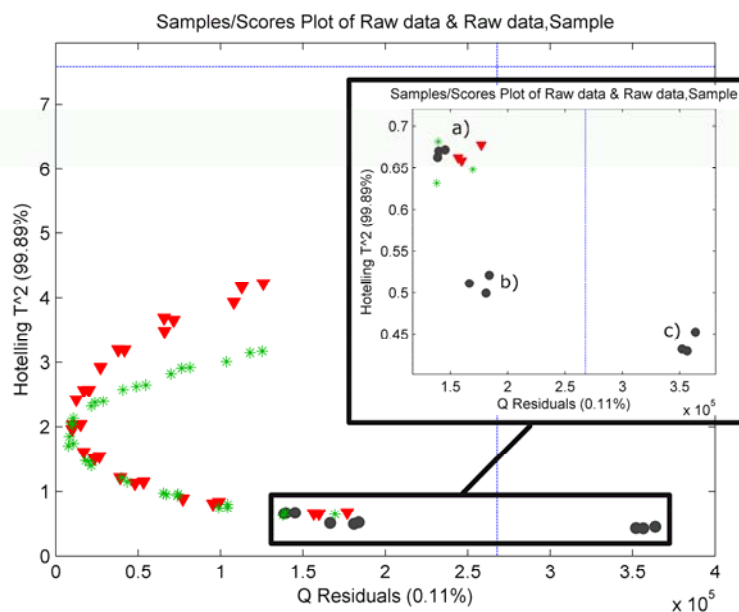


Figure S-13: Q residuals *versus* Hotelling T^2 plot from the UPCA model on a sample set containing 'normal' and irradiated eRDF samples (blue dashed lines are the threshold limits at a 97 % confidence limit) MPCA. The red triangles and green stars are the calibration samples: eRDF samples spiked with Trp and Tyr respectively. The black circles are the predicted samples: eRDF samples which have been exposed to visible light for 0 (a), 4 (b) or 7 days (c). The insert graph gives a clearer view of the a) non-irradiated and non-spiked samples and b) & c), irradiated samples (4 & 7 days).

5) Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) analysis of dataset1.

Following a reviewer's suggestion, we undertook MCR-ALS for comparison with the PARAFAC2 model. An augmented matrix was built by unfolding **dataset1** keeping the emission axis in common and stacking one on top of the others the excitation profiles of all samples at each emission wavelength (mode 2). The new matrix $D1u$ (71×1110) was then decomposed using MCR-ALS with non-negativity constraint on the spectra and concentrations. Up to 5 components could be resolved, and a 99.87 % of cumulative variance was explained. The emission profiles corresponding to each of the components are shown in Figure S-14. It is clear that components 1 to 3 can be clearly associated with Trp, Tyr and Py. The last component, as with PARAFAC2 is a composite signal of FA and RF. We believe that the two components could not be resolved for the same reasons described in the manuscript for PARAFAC2.

Note that herein the Rayleigh scattering was removed by interpolation because in this case Py and Trp could not be resolved using missing values (Figure S-17). The fourth component could be associated to the biased artificial signal introduced by the scattering removal. Figure S-16 represents the excitation profiles (average of three replicates) for each dilution sample, which were extracted from the spectral loadings (Figure S-15).

These results are in good agreement with the results obtained from implementing PARAFAC2, which does substantiate that the PARAFAC2 model in the manuscript is able to

deal with the spectral changes present in the data; and that these results indicate that the use of MCR-ALS did not allow for further resolution of the composite component.

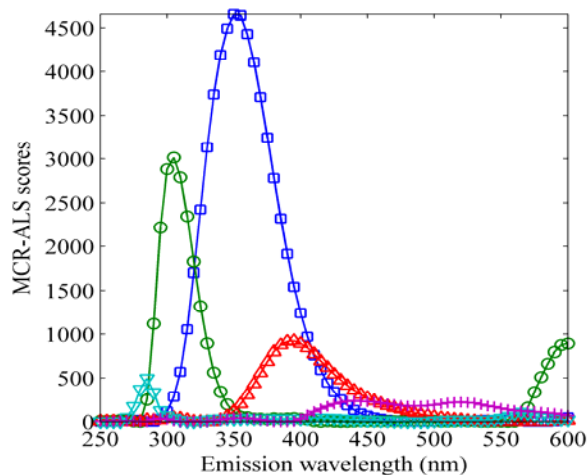


Figure S-14: Emission spectra obtained by MCR-ALS decomposition of dataset1 with non-negativity constraint on spectra and concentration. The first (\square), second (\circ), third (Δ), fourth (∇) and fifth (+) components are shown.

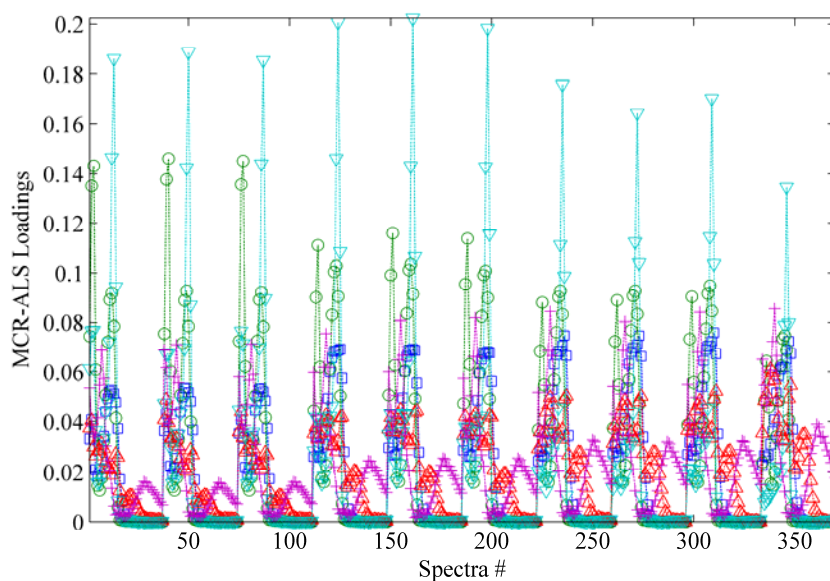


Figure S-15: Dilution profiles. Only the profiles of the first replicate (of three) are shown. The first (\square), second (\circ), third (Δ), fourth (∇) and fifth (+) components are represented.

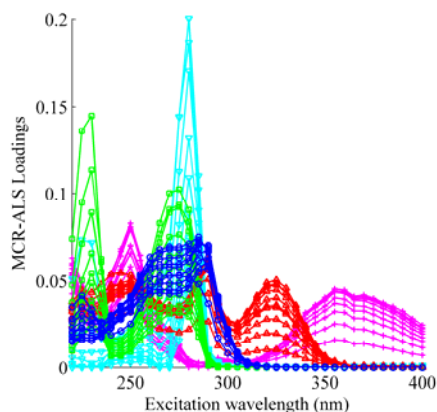


Figure S-16: Excitation profiles at the 10 eRDF concentrations. These profiles were average of the triplicates. The first (\square), second (\circ), third (Δ), fourth (∇) and fifth ($+$) components are shown.

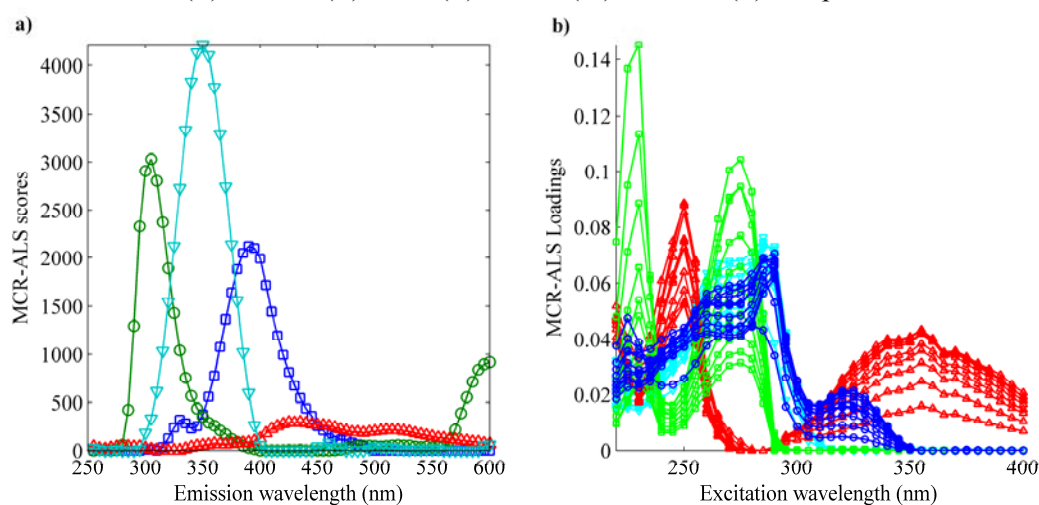


Figure S-17: Emission (a) and excitation (b) profiles of the MCR-ALS decomposition of D1u.

6) References

- [1] R. Chen, Fluorescence quantum yield measurements: vitamin B6 compounds, *Science*, 150 (1965) 1593-1595.
- [2] E. Miquel Becker, J. Christensen, C. Frederiksen, V. Haugaard, Front-face fluorescence spectroscopy and chemometrics in analysis of yogurt: rapid analysis of riboflavin, *Journal of Dairy Science*, 86 (2003) 2508-2515.
- [3] P. Drössler, W. Holzer, A. Penzkofer, P. Hegemann, Fluorescence quenching of riboflavin in aqueous solution by methionin and cystein, *Chemical Physics*, 286 (2003) 409-420.
- [4] A. Thomas, C. Lorente, A. Capparelli, M. Pokhrel, A. Braun, E. Oliveros, Fluorescence of pterin, 6-formylpterin, 6-carboxypterin and folic acid in aqueous solution: pH effects, *Photochemical & Photobiological Sciences*, 1 (2002) 421-426.
- [5] P. W. Ryan, B. Li, M. Shanahan, K. J. Leister, A. G. Ryder, Prediction of Cell Culture Media Performance Using Fluorescence Spectroscopy, *Analytical Chemistry*, 82 (2010) 1311-1317.
- [6] B. Li, P. W. Ryan, M. Shanahan, K. J. Leister, A. G. Ryder, Fluorescence EEM Spectroscopy for Rapid Identification and Quality Evaluation of Cell Culture Media Components., *Applied Spectroscopy*, 65 (2011) 1240-1249.
- [7] A. G. Ryder, J. De Vincentis, B. Y. Li, P. W. Ryan, N. M. S. Sirimuthu, K. J. Leister, A Stainless Steel Multi-Well Plate (SS-MWP) for High-Throughput Raman Analysis of Dilute Solutions, *J. Raman Spectrosc.*, 41 (2010) 1266-1275.
- [8] A. Calvet, B. Li, and A.G. Ryder *in preparation*.