SUPPORTING INFORMATION:

Low-content quantification in powders using Raman spectroscopy:

a facile chemometric approach to sub 0.1% limits of detection.

Boyan Li, Amandine Calvet, Yannick Casamayou-Boucau, Cheryl Morris, and Alan G. Ryder*

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

* To whom all correspondence should be addressed.

Tel: +353 91 49 2943 **Fax:** +353 91 49 4596 **Email:** alan.ryder@nuigalway.ie

TABLE OF CONTENTS

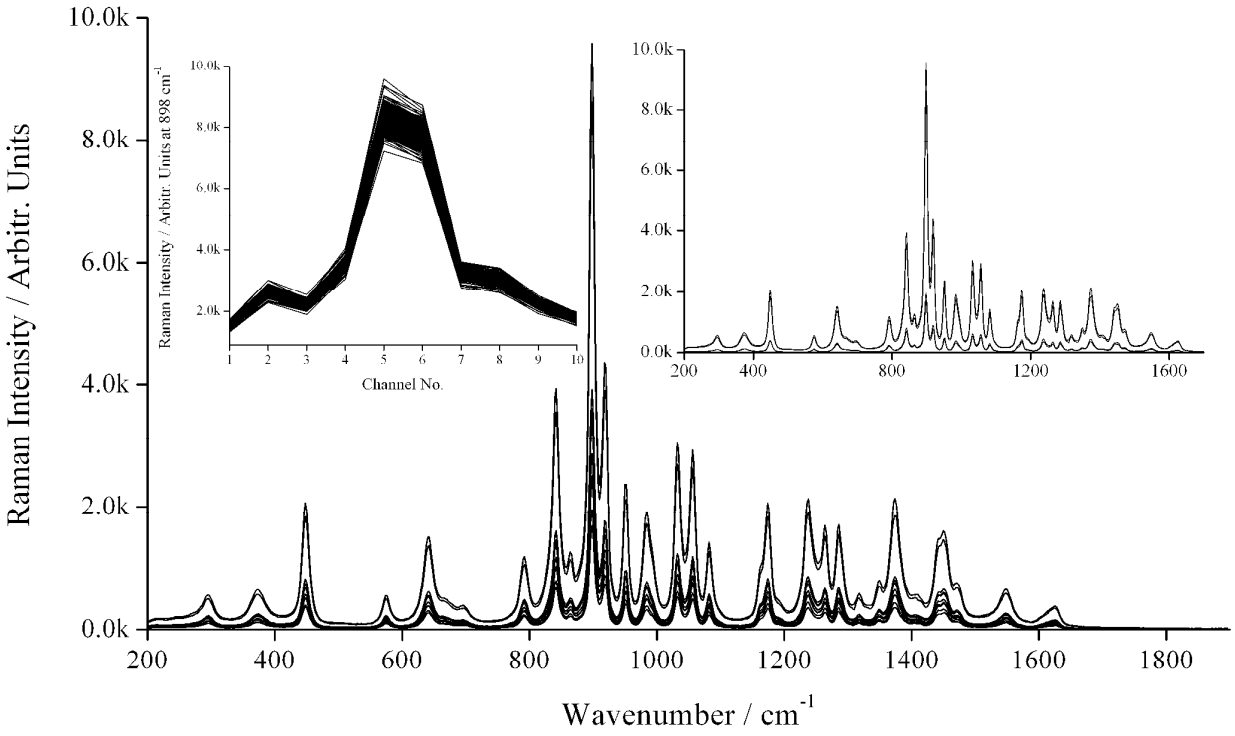

Figure S-1: Raman spectra obtained from the 10 channels at one sampling point (0.10% piracetam-proline mixture). Left inset shows 898 cm^{-1} band intensity variation for the 10 different channels (250 sampling points). Right inset shows overlaid individual spectra from channels 5, 6 (strong) and 1, 10 (weak) for a single sampling point.

S1. Proline Hydration

Dry proline powder (3 g) was exposed to 70% relative humidity, and every hour a Raman mapping measurement was made. After 4 hours, the proline powder surface had formed a hardened crust, and overnight the proline completely dissolved. Figure S-2a shows the raw spectra of dry proline powder (0 time) and hydrated proline (4 hrs). Hydration gave rise to significant changes in the Raman spectrum and induced a large baseline offset especially in the 250~1650 cm^{-1} range. Figure S-2b shows PC2 described in the paper (from kPCA of the 0.10% piracetam powder mixture Raman data) and the spectrum of hydrated proline after baseline correction and normalization. The similarity in terms of correlation coefficient between the spectrum and PC2 was ~0.5 indicating a reasonable fit. Thus, we can conclude that a significant proportion of the 0.01% data variance explained by PC2 was generated by hydration of the proline matrix. Although efforts were made to control sample exposure to humidity during Raman data collection it was clearly not 100% effective. During the extended data collection (several weeks) the relative humidity in the laboratory varied from a minimum of 30% to a maximum of 71%. This was the likely cause of the proline hydration.

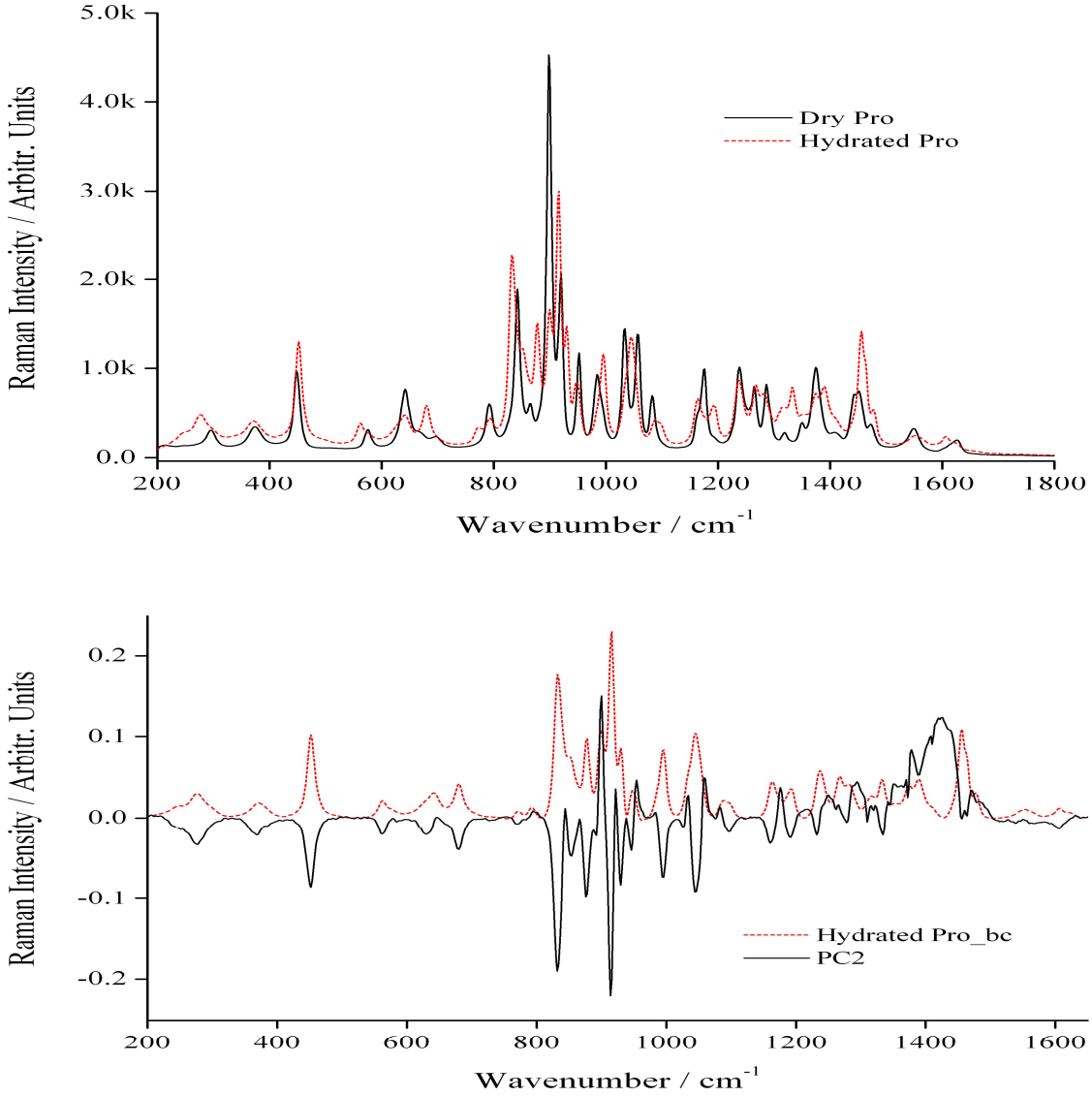

Figure S-2: (**a**) Spectra of dry proline powder and 4 hr hydrated proline; (**b**) comparison of the spectrum (baseline being corrected) of 4 hr hydrated proline and PC2 obtained by kPCA of the Raman mapping spectra of 0.10% piracetam powder mixture described in the manuscript. Spectra respectively collected at 1, 2, 3 hrs. were similar to the 4 hr. spectrum (data not shown).

S2. Piracetam/Proline Model Selection

The piracetam/proline powder mixture model was selected to develop the analytical methodology, based on the following criteria:

- Piracetam and proline have approximately equal Raman scattering coefficients. When we compared the integrated spectra (200~1896 cm^{-1} range) the Piracetam to Proline to Hydrated Proline ratio was 100:94:13. Obviously, LOD will be limited by this ratio, with LOD decreasing as the relative scattering efficiency of the target analyte increases compared to the matrix component.
- The ratio of the spectrum overlap integral to the total spectral area of a constituent was 0.59 for piracetam and 0.63 for proline; in essence, both were close to 50%. The smaller this ratio (which can vary from 0 to 1), the easier it should be to quantify a low-content analyte in mixtures. We selected this combination, as it was in the middle of the range and possibly more representative of what may be encountered in real world applications.
- For the HPLC validation study, piracetam and proline could be easily separated and the low-content piracetam produced a quite strong peak facilitating accurate quantification.
- The piracetam polymorph was stable while the proline matrix was sensitive to environmental factors *e.g.* water absorption leading to hydrate formation. This introduced another variable, which made the quantification of low-level analyte more complicated than a simple binary mixture model. Hydration is a common issue with solid-state matrix/formulation analysis, and this method needed to be able to identify samples that were compromised.

Table S-1: Macro (0–100% piracetam) PLS models generated from the Raman spectra of piracetam/proline powder mixtures using different pre-processing procedures.

PLS model	**Model-A**	**Model-B**	**Model-1**
Spectra	Raw	Baseline-Corrected	Baseline, & CRA-Corrected
Pre-processing	MSC/SNV/MC	MSC/SNV/MC	ACO/ MSC/SNV/MC
	RMSEC/RMSECV	**RMSEC/RMSECV**	**RMSEC/RMSECV**
Channel1	2.30/2.61	2.84/3.20	1.63/1.86
Channel2	2.37/2.65	2.93/3.27	1.65/1.86
Channel3	2.29/2.57	2.86/3.20	1.59/1.81
Channel4	2.32/2.61	2.88/3.23	1.63/1.85

Channel5	2.32/2.62	2.89/3.24	1.63/1.85
Channel6	2.32/2.61	2.90/3.25	1.60/1.81
Channel7	2.32/2.61	2.90/3.25	1.62/1.83
Channel8	2.38/2.67	2.94/3.29	1.70/1.91
Channel9	2.37/2.65	2.93/3.28	1.67/1.88
Channel10	2.34/2.62	2.91/3.25	1.66/1.88
Mean value	**2.33/2.62**	**2.90/3.25**	**1.64/1.85**
Standard Dev.	0.030/0.028	0.033/0.031	0.033/0.032
REC%/RECV%	**8.33/9.36**	**10.35/11.59**	**5.86/6.61**

S3. Quantification of Piracetam using Raman Mapping Data

In order to prove that multiplicative scatter correction (MSC), standard normal variate (SNV) and ant colony optimization (ACO) were necessary to improve model performance, four scenarios were evaluated, using the same spectral range (200~1896 cm^{-1}):

(1) Raw Raman spectra, MSC and SNV pre-processing, ACO variable selection;
(2) the data *after* baseline correction but *before* cosmic ray spike removal, with MSC and SNV pre-processing, ACO variable selection;
(3) Raw and baseline-corrected spectra with no MSC/SNV pre-processing;
(4) Raw and baseline-corrected spectra with MSC and SNV pre-processing.

For each scenario, a pool of PLS calibration models were built for each of the 10 channels.

S3.1--Scenario 1: ***Raw Raman mapping spectra*****:** Table S-2 shows the model statistical parameters generated from the raw Raman mapping spectra with no baseline correction but with MSC and SNV pre-processing, and by using ACO-selected variables. Compared to Table 1 (main manuscript), the RMSEC/RMSECV values are larger. When these models were applied for piracetam concentration prediction in the mixtures, over the 0.05~1.0% concentration range a relative prediction accuracy of 3.70% was obtained (Figure S-3a).

Table S-2: RMSEC/RMSECV values (in %) obtained from the PLS models built using the raw Raman spectra (10 channels) when different piracetam content ranges used (*i.e.*, concentration-segmented). The model accuracy was assessed by relative REC% and RECV% for calibration and cross-validation respectively.

PLS model	**Model1**	**Model2**	**Model3**	**Model4**	**Model5**
Piracetam in %	0~100	0~2.5	2.5~21.5	21.5~85.0	85.0~100
Channel1	1.98/2.26	0.039/0.044	0.15/0.18	1.14/1.54	0.31/0.66
Channel2	1.99/2.23	0.039/0.043	0.14/0.17	1.09/1.44	0.33/0.73

Channel3	1.96/2.22	0.035/0.038	0.18/0.23	1.03/1.37	0.25/0.52
Channel4	1.98/2.24	0.038/0.042	0.16/0.20	1.11/1.47	0.26/0.56
Channel5	2.00/2.26	0.036/0.039	0.16/0.19	1.12/1.49	0.25/0.52
Channel6	1.99/2.25	0.036/0.040	0.17/0.20	1.10/1.45	0.24/0.49
Channel7	1.98/2.23	0.036/0.040	0.17/0.20	1.06/1.41	0.25/0.50
Channel8	2.02/2.27	0.040/0.044	0.16/0.19	1.15/1.51	0.27/0.61
Channel9	2.01/2.26	0.038/0.041	0.17/0.18	1.08/1.42	0.24/0.49
Channel10	2.02/2.27	0.033/0.037	0.16/0.19	1.08/1.43	0.24/0.57
Mean value	**1.99/2.25**	**0.037/0.041**	**0.16/0.19**	**1.10/1.45**	**0.26/0.56**
Standard Dev.	0.020/0.019	0.002/0.002	0.009/0.016	0.038/0.051	0.033/0.079
REC%/RECV%	**7.12/8.03**	7.34/8.10	2.17/2.59	2.31/3.06	0.28/0.60

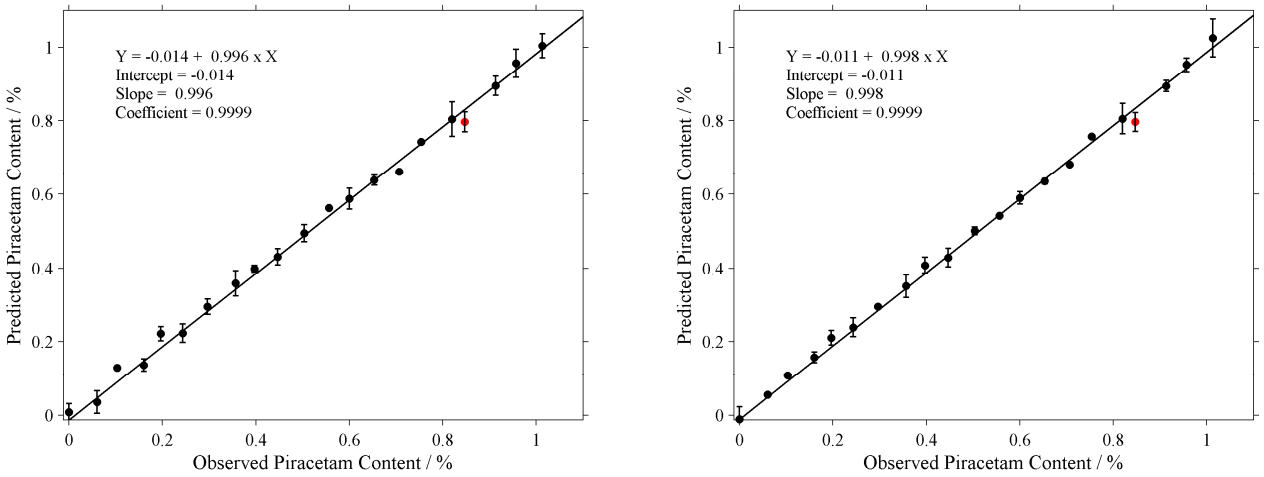

Figure S-3: Predictions of piracetam content in the powder mixtures, obtained by PLS model with ACO-selected variables using: (**a**) raw Raman spectra and (**b**) baseline-corrected spectra. Error bars represent the standard error for *n*=3 (triplicate measurements made for each powder mixture).

S3.2--Scenario 2: *Baseline-corrected Raman spectra*: Table S-3 shows the statistical parameters obtained from the models that used baseline-corrected Raman spectra with MSC and SNV pre-processing and ACO variable selection. The resulting RMSEC/RMSECV values were similar to those in Table 1 (main manuscript). When these models were applied for piracetam concentration prediction in the mixtures, over the 0.05~1.0% concentration range a relative prediction accuracy of 2.45% was obtained (Figure S-3b).

Table S-3: RMSEC/RMSECV values (in %) obtained from the PLS models using the baseline-corrected Raman spectra of piracetam/proline powder mixtures using different piracetam content ranges used. Model accuracy was assessed by relative REC% and RECV% for calibration and cross-validation respectively.

PLS model	**Model1**	**Model2**	**Model3**	**Model4**	**Model5**
Piracetam in %	0~100	0~2.5	2.5~21.5	21.5~85.0	85.0~100
Channel1	1.63/1.86	0.028/0.030	0.14/0.16	0.84/0.99	0.20/0.30
Channel2	1.65/1.86	0.039/0.042	0.14/0.16	0.90/1.04	0.18/0.27
Channel3	1.59/1.81	0.036/0.038	0.17/0.22	0.78/0.89	0.20/0.26
Channel4	1.63/1.85	0.036/0.039	0.15/0.18	0.86/0.99	0.19/0.26
Channel5	1.63/1.85	0.036/0.039	0.15/0.18	0.85/0.98	0.20/0.27
Channel6	1.60/1.82	0.034/0.037	0.15/0.18	0.84/0.96	0.22/0.36
Channel7	1.62/1.83	0.036/0.039	0.16/0.19	0.84/0.95	0.20/0.31
Channel8	1.70/1.91	0.041/0.045	0.15/0.17	0.95/1.10	0.19/0.27
Channel9	1.67/1.88	0.037/0.040	0.15/0.18	0.87/0.99	0.19/0.25
Channel10	1.66/1.88	0.031/0.034	0.16/0.18	0.86/0.99	0.16/0.31
Mean value	**1.64/1.86**	**0.035/0.038**	**0.15/0.18**	**0.86/0.99**	**0.19/0.29**
Standard Dev.	0.032/0.031	0.004/0.004	0.010/0.015	0.045/0.054	0.016/0.036
REC%/RECV%	**5.86/6.63**	7.02/7.62	2.06/2.44	1.81/2.09	0.20/0.30

S3.3--Scenario 3: ***Neither MSC nor SNV applied to Raman mapping spectra*****:** If neither MSC nor SNV pre-processing was used, then both the raw and baseline corrected Raman spectra generated much poorer calibration models (Table S-4), clearly indicating the need for these pre-processing steps.

S3.4 Scenario 4: ***Need for ACO variable selection*****:** When MSC and SNV were applied to pre-process the raw and baseline-corrected spectra, another two sets of PLS models were generated (Table S-4). Model quality improved in terms of the RESEC/RMSECV values, compared to the non-processed data. However, the values were still relatively large, because all spectral variables were used for modelling. This is why ACO was implemented to select the informative variables.

Table S-4: Comparison of PLS models (RMSEC/RMSECV values in %) generated for different pre-processing conditions. All models were the Model-1 (0-100% Piracetam) and used the full spectral range of 200~1896 cm^{-1} for each of the 10 channels. Model accuracy was assessed by relative REC% and RECV% for calibration and cross-validation respectively.

Details	**Raw spectra**	**Raw spectra**	**Baseline-corrected spectra**	**Baseline-corrected spectra**
Pre-proc.	**No MSC or SNV**	**MSC & SNV**	**No MSC or SNV**	**MSC & SNV**
Piracetam in %	0~100	0~100	0~100	0~100
Channel1	2.71/3.01	2.30/2.61	3.58/3.87	2.84/3.20
Channel2	2.81/3.10	2.37/2.65	3.61/3.90	2.93/3.27
Channel3	2.79/3.09	2.29/2.57	3.64/3.92	2.86/3.20
Channel4	2.69/2.99	2.32/2.61	3.55/3.83	2.88/3.23
Channel5	2.64/2.93	2.32/2.62	3.51/3.79	2.89/3.24
Channel6	2.69/2.98	2.32/2.61	3.57/3.85	2.90/3.25
Channel7	2.70/2.99	2.32/2.61	3.57/3.85	2.90/3.25
Channel8	2.76/3.05	2.38/2.67	3.58/3.86	2.94/3.29
Channel9	2.76/3.05	2.37/2.65	3.56/3.85	2.93/3.28
Channel10	2.76/3.06	2.34/2.62	3.59/3.88	2.91/3.25
Mean value	***2.73/3.02***	**2.33/2.62**	***3.57/3.86***	**2.90/3.25**
Standard Dev.	0.052/0.053	0.030/0.028	0.035/0.036	0.033/0.031
REC%/RECV%	9.75/10.80	8.33/9.36	12.77/13.79	10.35/11.59

S4. Prediction of Piracetam in Mixtures

High-Content prediction: The high-content piracetam mixture samples (2.0~100%) were also well predicted with a relative prediction accuracy of 0.9% by the final calibration model (Figure S-4). Only a single sample (the 60% piracetam mixture) showed a significant deviation between the triplicate predictions.

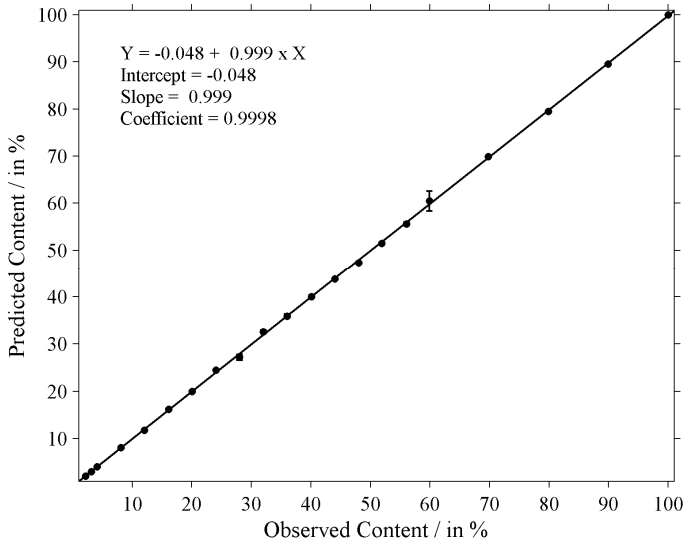

Figure S-4: Raman based PLS quantification of piracetam (2.0~100% concentration). Error bars represent standard error for *n*=3 (triplicate measurements made for each powder mixture).

Low Content prediction: The concentration of piracetam in the low concentration powder mixtures was estimated from the spectral data using the following sequence of operations:

(1) All the Raman spectra of each Raman mapping measurement (841 sampling points × 10 channels) was assessed for gross outliers and spectra with an average intensity of <70% of the mean were excluded.
(2) This spectral dataset which comprised of >8000 spectra of 849 variables were subjected to the baseline correction, cosmic ray artefact removal, and the various data pre-treatment (MSC, SNV).
(3) The corrected spectra were then subjected to the same variable selection as was performed on the calibration spectra. This produced a reduced spectral variable set (138 instead of 849 variables) for each spectrum.
(4) The resultant spectral data for each of the 10 channels were then input into the appropriate 10-channel PLS calibration model (Model-1, covering the 0–100% range).
(5) An initial estimate of piracetam content (*Ini_Est*) was made for each of the >800 spectra for each channel.
(6) According to the magnitude of the obtained *Ini_Est* values, the piracetam content was then re-predicted for each of the >8000 spectra:

a. if *Ini_Est* ≤2.5%, **Model-2** was used for prediction for each channel spectrum,
b. if 2.5%<*Ini_Est* ≤21.5%, **Model-3** was used for prediction for each channel spectrum,
c. if 21.5%<*Ini_Est* ≤85.0%, **Model-4** was used for prediction for each channel spectrum,
d. if 85.0%<*Ini_Est* ≤100%, **Model-5** was used for prediction for each channel spectrum.

(7) Eventually for each sample, a total of >8000 piracetam content predictions were obtained.
(8) Finally, all of these local piracetam content predictions were averaged to give the true concentration of piracetam (a single number) in the sample mixture.

S5. Outlier in prediction model: a hydration event

The PLS model failed to precisely predict the 0.85 piracetam mixture. To understand why data kernel PCA was run on the triplicate mapping data[1] from the 0.80 (25190 spectra) and 0.85 (25,170 spectra)[2] piracetam mixtures the (**X**) comprising. The percentage of data variance captured by each PC (Table S-5) showed significant differences between the two samples. Next the obtained PCs were compared to the spectra of piracetam, proline, and hydrated proline, and the correlation coefficient similarity was calculated (Table S-6). If we assume that the spectra of proline, piracetam, hydrated proline were orthogonal, and that there was no unwanted water absorption, then their content (as expressed by the % data variance explained by kPCA) in the low-content piracetam mixtures should follow in decreasing order: proline, piracetam, hydrated proline. These showed that:

(1) PC1 from the 0.85% piracetam mixture (denoted PC1_85) and PC1 from the 0.80% piracetam mixture (PC1_80) mostly represented proline, because the similarity coefficients with proline were 0.9973 and 0.9965 respectively. We would have expected that the PC1_85 value should have been smaller because proline content increased from 99.15% to 99.20%. When we look at the correlation coefficients with the hydrated proline we see that PC1_85 (0.6278) is marginally higher than PC1_80 (0.6261). This suggests that the 0.85% sample contains some more hydrated proline, however this is not conclusive because the spectral correlation between proline and hydrated proline is relatively large (0.6239).
(2) PC3 from the 0.85% piracetam mixture (denoted PC3_85) and PC2 from the 0.80% piracetam mixture (PC2_80) most likely represented piracetam, because the similarities of PC3_85 and PC2_80 to piracetam were –0.89 and 0.92 respectively. The correlation coefficient between PC3_85 and PC2_80 was also very high –0.95, (Figure S-5a).

[1] After baseline correction and cosmic ray spike removal.
[2] Excluded 60 spectra with <30% of the average intensity.

(3) PC3_85 explained much less data variance (0.0055%) compared to PC2_80 (0.0098%). This was opposite to what was expected.
(4) PC2_85 captured a data variance of 0.0071%, which was more variance than the piracetam component (PC3_85, 0.0055%). PC2_85 and piracetam correlation was low (0.22) indicating that this represented something else other than piracetam, although some of the variance might be piracetam linked.
(5) PC2_85 and PC3_80 were very similar (Figure S-5b) and both had bands which corresponded to hydrated proline, but the coefficients with hydrated proline were very low (–0.05 for PC2_85, and –0.11 for PC3_80).
(6) PC4_85 and PC4_80 were very similar (similarity coefficient = 0.96) and were likely describing small proline solid-state differences the 841 sampling points (Figure S-5c).

Based on these analyses and the observations that proline absorbs water relatively quickly when exposed to normal laboratory humidity levels we can conclude that the root cause of the under-prediction of the 0.85% piracetam mixture was unwanted proline hydration. The significant amount of hydrated proline in this mixture obscured the variance caused by piracetam, leading to an under prediction.

Table S-5: Summary of the kernel PCA on the triplicate Raman mapping data of the individual 0.85% and 0.80% piracetam mixtures. Data baseline and CRA corrected.

Percent variance Captured by kPCA Model		
# Principal Component	0.85% piracetam mixture % Variance by this PC	0.80% piracetam mixture % Variance by this PC
1	99.9726	99.9678
2	0.0071	**0.0098**
3	**0.0055**	0.0067
4	0.0033	0.0033

Table S-6: Calculated similarity coefficients between the obtained PCs, and Raman spectra of piracetam, proline, and hydrated proline.

Similarity	Pir	Proline	HydrPro	PC1_85	PC2_85	PC3_85	PC4_85	PC1_80	PC2_80	PC3_80	PC4_80
Piracetam	1	0.1700	0.2277	0.1726	0.2211	-0.8869	-0.0125	0.1721	0.9195	-0.0524	-0.0377
Proline	0.1700	1	***0.6239***	0.9973	-0.0023	0.1233	-0.1447	0.9965	-0.1326	0.0484	-0.1411
Hydr Pro	0.2277	0.6239	1	0.6278	-0.0519	-0.0297	0.0267	0.6261	0.0048	-0.1078	-0.0949
PC1_85	0.1726	**0.9973**	***0.6278***	1	-0.0657	0.1067	-0.1186	0.9999	-0.1352	-0.0165	-0.1121
PC2_85	0.2211	-0.0023	-0.0519	-0.0657	1	0.0219	-0.0429	-0.0735	0.2854	0.9517	-0.0403
PC3_85	**-0.8869**	0.1233	-0.0297	0.1067	0.0219	1	0.0404	0.1052	-0.9493	0.3147	0.0820
PC4_85	-0.0125	-0.1447	0.0267	-0.1186	-0.0429	0.0404	1	-0.1131	-0.0116	-0.0553	0.9594
PC1_80	0.1721	**0.9965**	***0.6261***	**0.9999**	-0.0735	0.1052	-0.1131	1	-0.1360	-0.0241	-0.1061
PC2_80	**0.9195**	-0.1326	0.0048	-0.1352	0.2854	**-0.9493**	-0.0116	-0.1360	1	-0.0105	-0.0528
PC3_80	-0.0524	0.0484	-0.1078	-0.0165	**0.9517**	0.3147	-0.0553	-0.0241	-0.0105	1	-0.0252
PC4_80	-0.0377	-0.1411	-0.0949	-0.1121	-0.0403	0.0820	**0.9594**	-0.1061	-0.0528	-0.0252	1

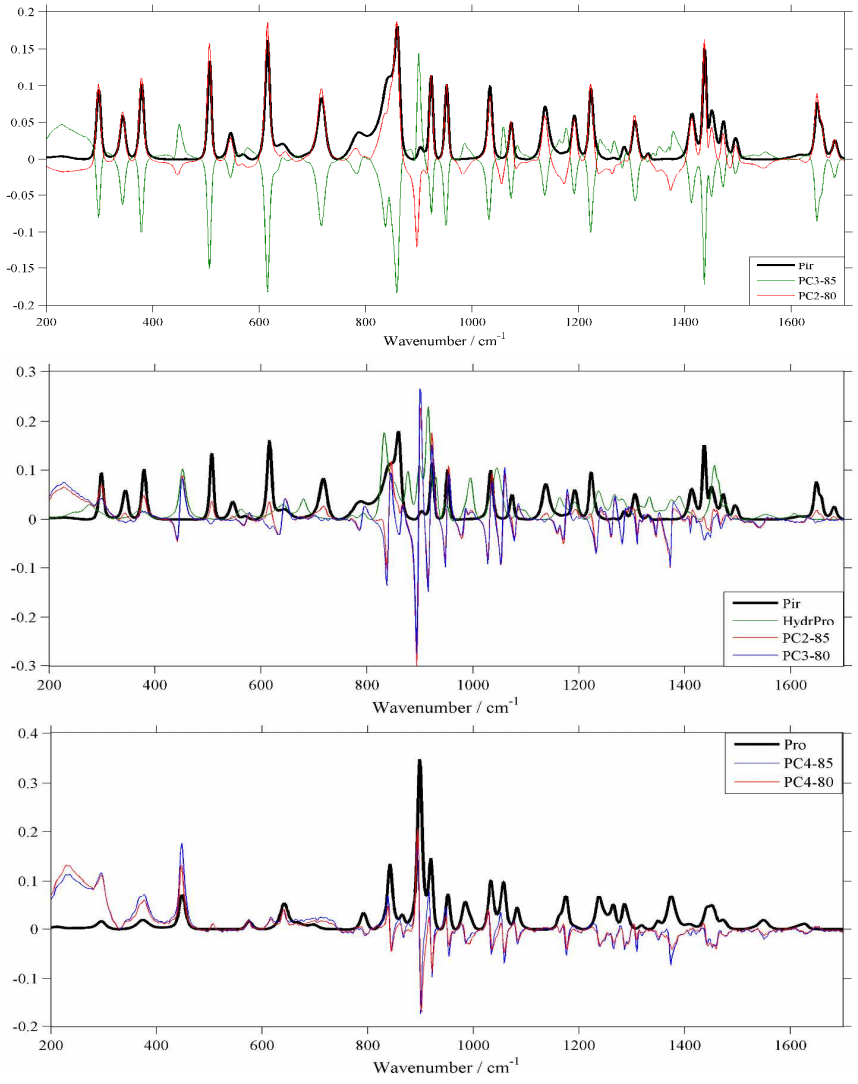

Figure S-5: Comparison of the PCs obtained from the 0.85% and 0.80% piracetam mixtures, with the normalized spectra of piracetam, proline, and hydrated proline.

S6. HPLC Quantification of Piracetam: validation

S6.1--HPLC validation analysis: The HPLC system used was a Waters Alliance 2695 separation module controlled by Empower 3. It was equipped with a SunFire C_{18} column (150×4.6 mm, 5μm) and a Waters 2487 dual wavelength absorbance detector. Piracetam has a maximum absorption 197 nm (in water) however; measurements were performed at 208 nm because of instrument limitations. The mobile phase was 10 mM phosphate buffer (KH_2PO_4) which was sonicated for 25 min, and then filtered using a 0.22 μm membrane. HPLC analyses were performed on a reversed-phase chromatographic system using isocratic elution with a flow rate of 1.0 mL/min. The injection volume for each solution was 50 μL, and runtime was ten minutes and four injections were made for each sample (total analysis time of 40-45 min).

S6.2--Calibration curve for piracetam quantification: Eleven standard solutions containing preset quantities of piracetam and proline (Table S-7) were prepared to generate the HPLC calibration curve so that the unknown mixture samples could be quantified. The concentration range and elution method employed were optimized for piracetam quantification. Figure S-6a shows the overlaid chromatograms for the 11 standard solutions (4 injections for each solution).

Table S-7: Composition of the standard solutions used in HPLC analysis.

Standard Solution	Piracetam in solution (in mg/100 mL)	Proline in solution (in mg/100 mL)
1	0.05	0.05
2	0.25	0.25
3	0.5	0.5
4	1.0	1.0
5	5	5
6	10	10
7	15	15
8	20	20
9	25	25
10	30	30
11	40	40

Piracetam produced a single peak (**peak1**) with a retention time (RT) between 4.53 and 4.68 min. (mean RT = 4.62 min.). Proline gave rise to a weak, well-resolved single peak (**peak2**) at an RT of 1.52 min. The third peak at RT 1.72 min. (**peak3**) originated from the mobile phase (present

in blank). One can see an obvious increase in intensity of both **peak1** and **peak2** with increasing amount of piracetam and proline, along with variations in **peak3**. Another observation was that peak shift occurred to **peak1** when the piracetam amount in the standard solutions increased.

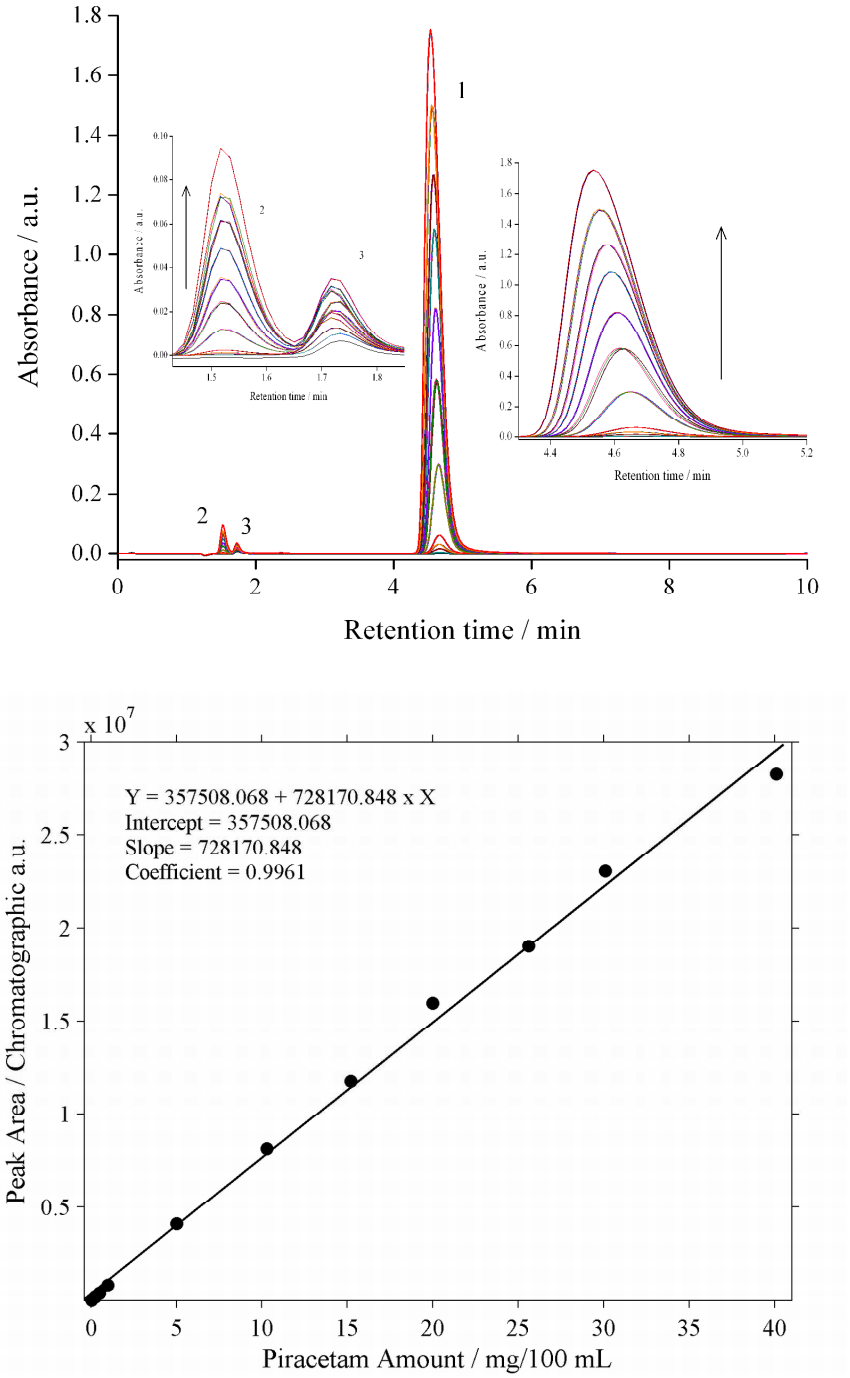

Figure S-6: (**a**) Overlay of the 11 standard solution chromatograms show peak intensity and RT changes; (**b**) Piracetam calibration curve obtained from calculating the peak area of **peak1** in each chromatogram of all the standard solutions given in Table S-7. Error bars represent standard error for *n*=4 (4 injections made for each standard solution).

A quantitative calibration curve was then built by using Empower 3 from the **peak1** area in each chromatogram. Figure S-6-Sb shows the near perfect calibration curve obtained where the standard deviations between the peak1 areas measured for each of the 4 injections for each solution were too small to be seen in this plot. The high quality of this calibration curve can be attributed to the fact that all samples were collected in a single day using the auto-sampler and as such both operator, and day-to-day variation were minimized. This calibration curve was then used to quantify the piracetam content of 23 samples (Table S-8). The LOD for the HPLC method was calculated from this calibration curve (using piracetam peak areas), and a value of 0.0123 mg/100mL was obtained which was equivalent to a piracetam content of 0.041% in the solid state.[3]

S6.3--*Quantification of the low content piracetam samples*: Three 30 mg powder portions were taken from one of the three replicate piracetam/proline mixture samples that had been analyzed by Raman. The sample was vortex mixed to ensure homogeneity between aliquots. Each 30 mg aliquot was then dissolved in 100 mL water. The HPLC sample set (Table S-8) comprised 69 solutions (23 mixtures in triplicate) and covered the full piracetam concentration range of (0~100%).

Table S-8: Piracetam concentration in the solid mixtures / solutions and the HPLC quantification (*mean value ± standard deviation*) from triplicate measurements using raw and corrected peak areas. The StdDev-to-mean ratio (%) gives an assessment of method reproducibility.

Sample ID #	Piracetam in mixture (w/w%)	Piracetam in solution (mg/100 mL)	HPLC (*n*=3) quantification (mg/100 mL)	StdDev-to-mean ratio (%)	HPLC (*n*=3) quantification (mg/100 mL) after peak area correction	StdDev-to-mean ratio (%)
1	0	0	0	n/a	0	n/a
2	1	0.3	0	n/a	0	n/a
3	2	0.6	0.23±0.07	30.97	0.25±0.08	30.94
4	3	0.9	0.78±0.26	**33.60**	0.82±0.27	32.67
5	4	1.2	0.80±0.23	28.71	0.83±0.24	28.84
6	8	2.4	2.01±0.41	20.19	2.07±0.42	20.46
7	12	3.6	3.49±0.44	12.55	3.60±0.45	12.61
8	16	4.8	4.14±1.39	33.55	4.28±1.43	33.43
9	20	6.0	6.15±0.19	3.14	6.33±0.20	3.13
10	24	7.2	6.85±1.14	16.68	7.05±1.17	16.61
11	28	8.4	8.33±0.52	6.27	8.59±0.53	6.17
12	32	9.6	10.37±0.48	4.62	10.68±0.40	3.78
13	36	10.8	11.94±0.75	6.28	12.28±0.72	5.90

[3] LOD calculations from: ICH Harmonised Tripartite Guideline. Validation of analytical procedures: text and methodology, Q2 (R1).

14	40	12.0	10.84±1.02	9.37	11.12±0.99	8.91
15	44	13.2	13.66±0.62	4.57	14.04±0.73	5.22
16	48	14.4	15.04±0.68	4.50	15.54±0.69	4.45
17	52	15.6	15.98±0.20	1.23	16.48±0.25	1.52
18	56	16.8	16.32±0.66	4.07	16.78±0.68	4.05
19	60	18	18.33±0.41	2.22	18.89±0.47	2.49
20	70	21	21.78±0.13	**0.61**	22.25±0.20	0.89
21	80	24	23.83±0.62	2.60	24.41±0.80	3.27
22	90	27	27.48±0.36	1.30	28.23±0.37	1.30
23	100	30	30.52±0.23	0.76	31.52±0.32	1.00

All 69 samples were analyzed in random order over a four-day period and each day a control sample (sample 7 from the calibration sample set, 4 injections) was run. Analysis of the raw chromatograms (*n=69*) showed that the piracetam peak RT was observed between 4.454 and 4.614 min (mean = 4.533 min) whereas the proline peak at 1.520 min, experienced very little peak shift. When the calibration curve was used on the uncorrected data (Figure S-7a) a reasonably good quantitative correlation was obtained, with an REP of 5.4% for the 1.0~100% range. However, there were significant variations in the triplicate measurements made over the 4 days. For example the standard deviation-to mean value ratios of the individual samples with respect to their triplicate measurements increased from high-content to low-content piracetam samples (from 0.61% to 33.60% respectively). This implied that the HPLC method was less accurately quantifying the low-content piracetam mixtures

The control standards showed that there were very significant day-to-day variations in the HPLC method performance, as could be expected. The 16 peak1 areas of the control samples had displayed a 12.8% variation. The variation on the 4 days was 0.03, 0.15, 0.06, and 0.19% respectively for the control measurements, indicating very good reproducibility on a daily basis. To try and improve the HPLC quantification, the raw chromatograms were corrected by normalization to the control sample area, which had been measured every day. The corrected data gave no significant improvements (Figure S-7b). The use of an internal standard was thus warranted but this then constituted an additional sample pre-treatment step, and would result in sample contamination.

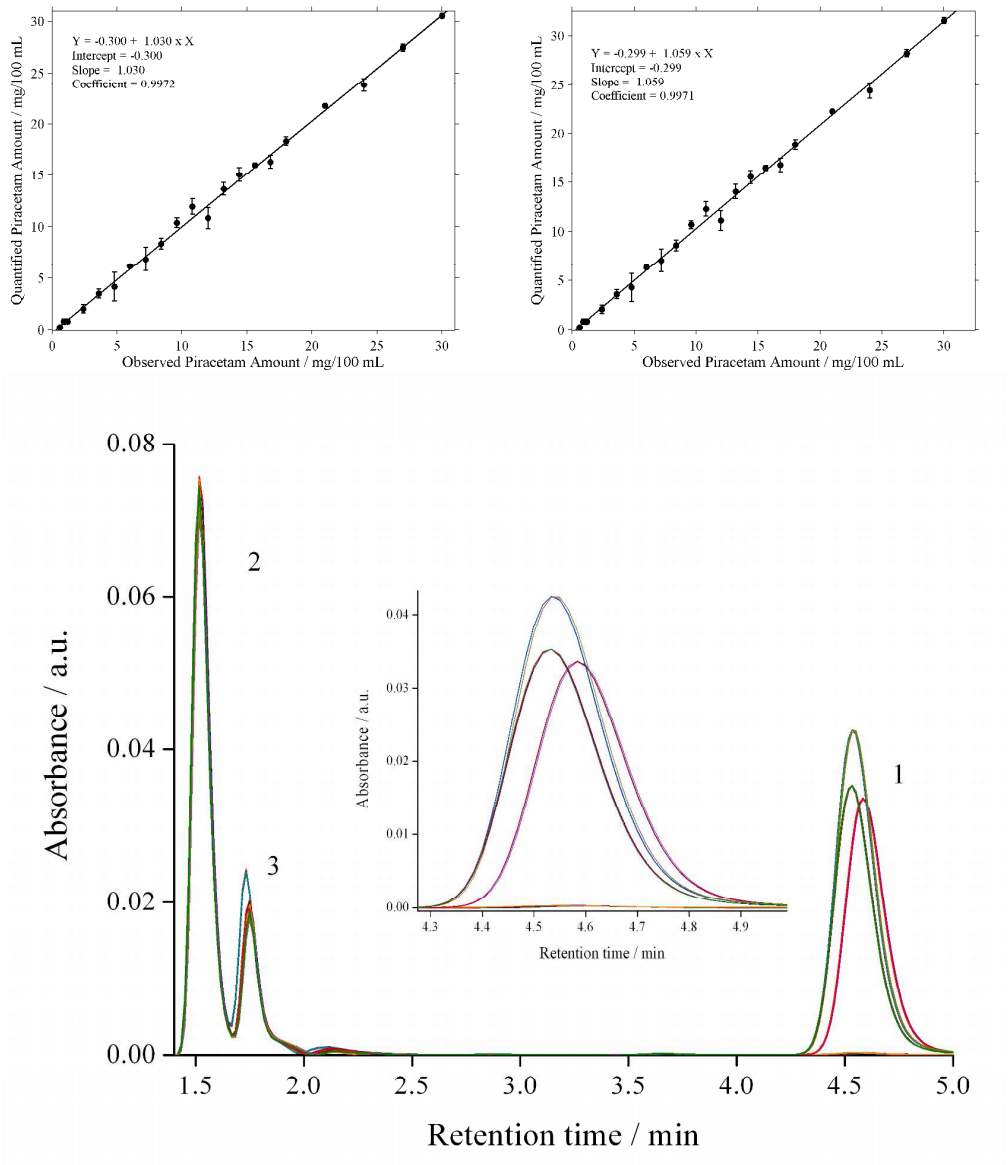

Figure S-7: HPLC based piracetam quantification of the 23 piracetam/proline mixture solutions using: **(a)** piracetam peak areas, and **(b)** the corrected chromatograms using control samples collected over 4 days; **(c)** Overlay of the uncorrected triplicate chromatograms for Sample**#1** (0 mg/100 mL piracetam) and **#3** (0.6 mg/100 mL piracetam). Each chromatogram was the average of 4 injections for each sample collected on a single day. The replicate data were collected on 3 of 4 different days. This shows the large variation between uncorrected (*i.e.* raw data) replicate HPLC measurements in terms of both peak area and retention time.