Journal of Pharmaceutical and Biomedical Analysis 56 (2011) 944–949

Contents lists available at ScienceDirect

Journal of Pharmaceutical and Biomedical Analysis

journal homepage: www.elsevier.com/locate/jpba

Data fusion of Fourier transform infrared spectra and powder X-ray diffraction patterns for pharmaceutical mixtures

Rahul V. Haware [a], Patrick R. Wright [a,1], Kenneth R. Morris [a], Mazen L. Hamad [b,*]

[a] University of Hawaii at Hilo, College of Pharmacy, 34 Rainbow Drive, Hilo, HI 96720, USA
[b] Department of Chemistry, University of Hawaii at Hilo, 200 W Kawili St., Hilo, HI 96720, USA

ARTICLE INFO

Article history:
Received 18 May 2011
Received in revised form 9 August 2011
Accepted 10 August 2011
Available online 17 August 2011

Keywords:
Data fusion
Multivariate analysis
Pharmaceutical powder mixtures
Fourier transform infrared spectroscopy
Powder X-ray diffraction

ABSTRACT

Fusing complex data from two disparate sources has been demonstrated to improve the accuracy in quantifying active ingredients in mixtures of pharmaceutical powders. A four-component simplex-centroid design was used to prepare blended powder mixtures of acetaminophen, caffeine, aspirin and ibuprofen. The blends were analyzed by Fourier transform infra-red spectroscopy (FTIR) and powder X-ray diffraction (PXRD). The FTIR and PXRD data were preprocessed and combined using two different data fusion methods: fusion of preprocessed data (FPD) and fusion of principal component scores (FPCS). A partial least square (PLS) model built on the FPD did not improve the root mean square error of prediction. However, a PLS model built on the FPCS yielded better accuracy prediction than PLS models built on individual FTIR and PXRD data sets. The improvement in prediction accuracy of the FPCS may be attributed to the removal of noise and data reduction associated with using PCA as a preprocessing tool. The present approach demonstrates the usefulness of data fusion for the information management of large data sets from disparate sources.

1. Introduction

The need to understand the critical material and process attributes on the end product quality of pharmaceutical products is now an imperative with respect to the ICH Q8 guideline issued by FDA [1]. Consequently, quantitative and qualitative applications of sophisticated high data density analytical tools like Fourier transform infrared spectroscopy (FTIR), powder X-ray diffraction (PXRD), Raman spectroscopy and near infrared spectroscopy have gained wider acceptance in characterizing pharmaceutical processes, intermediates and products [2–4]. A major obstacle associated with these analytical techniques is the generation of large data matrices which may be complex and difficult to interpret. Thus, it is critical that the end users of these tools have appropriate methods of data reconciliation in order to extract the sought after information for subsequent prediction of the process outcomes. Multivariate analysis, also called chemometrics when applied to chemical-specific applications, has been offered as one key to extracting critical information from large data sets generated by a single high data density tool.

A second obstacle associated with using high data density tools arises when these techniques are used to make measurements on samples that are considered non-ideal. An example of a non-ideal sample for FTIR is a heterogeneous, multi-component, solid state pharmaceutical mixture. Ideally, the absorbance of infrared light at a particular wavenumber will be directly proportional to the concentration of each absorbing species; however, the variation in the extent of light scattering at particulate interfaces tends to significantly increase the error in the measurement. This artefact, caused by variation in the physical aspects of the sample matrix, may prevent the use of FTIR as a technique for the quantitative characterization of multi-component solid state pharmaceutical samples.

PXRD, on the other hand, is a technique known for structural characterization, not chemical characterization, of single component solid-state material samples. Thus, neither method is ideal for the quantitative chemical characterization of multi-component, pharmaceutical samples. However, since the two techniques yield different kinds of information, their individual data sets can be combined into a single data set which provides more information than either technique by itself. What is not clear is how data from the two techniques can be combined to better perform a single task than either technique could perform on its own.

An important technique emerging from the informatics domain is data fusion. The aim of data fusion is to facilitate the faultless integration of information from various sources to develop a single model or decision [5]. It is hypothesized that data fusion may be a

* Corresponding author. Tel.: +1 808 933 2194; fax: +1 808 974 7693.
E-mail address: mazen@hawaii.edu (M.L. Hamad).
[1] Student undertaking summer internship from Albert-Ludwigs-University Freiburg, Freiburg, Germany.

0731-7085/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpba.2011.08.018

Table 1
Four component simplex-centroid experimental design in units of percentage concentration [acetaminophen (APAP), caffeine (CAF), ibuprofen (IBU) and aspirin (ASA)].

Sr. No.	Exp. Name	APAP (%)	CAF (%)	IBU (%)	ASA (%)
1	Vert001a	100.00	0.00	0.00	0.00
2	Vert002a	0.00	100.00	0.00	0.00
3	Vert003a	0.00	0.00	100.00	0.00
4	Vert004a	0.00	0.00	0.00	100.00
5	Edge001a	50.00	50.00	0.00	0.00
6	Edge002a	50.00	0.00	50.00	0.00
7	Edge003a	50.00	0.00	0.00	50.00
8	Edge004a	0.00	50.00	50.00	0.00
9	Edge005a	0.00	50.00	0.00	50.00
10	Edge006a	0.00	0.00	50.00	50.00
11	Face001a	33.33	33.33	33.33	0.00
12	Face002a	33.33	33.33	0.00	33.33
13	Face003a	33.33	0.00	33.33	33.33
14	Face004a	0.00	33.33	33.33	33.33
15	Axis001a	62.50	12.50	12.50	12.50
16	Axis002a	12.50	62.50	12.50	12.50
17	Axis003a	12.50	12.50	62.50	12.50
18	Axis004a	12.50	12.50	12.50	62.50
19	Cent-a	25.00	25.00	25.00	25.00
20	Cent-b	25.00	25.00	25.00	25.00
21	Cent-c	25.00	25.00	25.00	25.00

useful strategy to integrate data from FTIR and PXRD for the characterization of multi-component, pharmaceutical samples. Various scientific and engineering disciplines, such as robotics, remote sensing, image analysis, and analytical chemistry, are employing data fusion concepts and realizing better information management [6–9]. Furthermore, data fusion coupled with multivariate analysis (MVA) is a new and potentially very powerful approach to modeling pharmaceutical systems.

The goal of the present work was to investigate the suitability of data fusion [10] in combination with MVA methods to build more accurate predictive models. Principal component analysis (PCA) was used for exploratory data analysis and it was also used as a data reduction technique prior to data fusion. Partial least square (PLS) regression was ultimately used to build predictive models based on the FTIR data set, the PXRD data set, the data set prepared by fusion of preprocessed data (FPD) and the data set prepared by fusion of principal component scores (FPCS). The quantitative prediction accuracy of fractions of acetylsalicylic acid, caffeine, ibuprofen, and acetaminophen in blended powder samples was compared using leave-one-out cross validation. The models were also used to predict fractions of the four components in blind, unknown powder mixtures.

2. Experimental and methods

2.1. Materials

Acetylsalicylic acid (ASA) was purchased from Alfa Aesar (Ward Hill, MA). Acetaminophen (APAP) was purchased from Ortho-McNeil Pharmaceuticals (Titusville, NJ). Ibuprofen (IBU) and caffeine (CAF) were purchased from Spectrum Chemical (Gardena, CA).

2.2. Experimental design

Four active ingredients (APAP, ASA, CAF and IBU) were tested by a four-component simplex-centroid design (SCD). The four-component SCD was used to achieve better predictability with a high accuracy of unknown fractions of subjected active ingredients [11]. In total, 21 combinations of subjected active ingredients were tested by both FTIR and PXRD techniques (Table 1); 4 vertices for 4 pure components, 6 edge centers of the binary mixtures with equal proportions of the selected two components, 3 face centers of the ternary mixtures with equal proportions of the selected three components, 4 axes of the quaternary mixtures with varying proportions of the selected four components and 3 center experiments of the quaternary mixtures with equal proportions of the selected four components to check both the linearity and repeatability of the experimental results.

2.3. Fourier transform infrared spectroscopy (FTIR)

The spectra of the 21 calibration samples and 4 unknown samples were collected using a Thermo Nicolet NEXUS 670 FTIR instrument equipped with a Nicolet Smart MIRacle accessory (Thermo Fisher Scientific, Waltham, MA). The MIRacle accessory uses a glassy material known as AMTIR (Amorphous Material Transmitting Infrared Radiation – composed of Ge, As, and Se) to measure the absorbance in the attenuated total reflectance (ATR) mode. The samples were measured by inserting approximately 25 mg of mixed sample powder into the trough insert and supplying sufficient pressure using the micrometer pressure clamp to compress the sample against the AMTIR glass. For each sample, 32 scans in the wavenumber range from 650 cm^{-1} to 4000 cm^{-1} (at a resolution of 4 cm^{-1}) were averaged to produce a single spectrum. The resulting spectral data vectors contained 1738 data points. The spectral data were acquired in absorbance mode using OMNIC software (Thermo Fisher Scientific, Waltham, MA) and exported to MATLAB® (Mathworks, Natick, MA) and the Unscrambler® (Unscrambler® 10.0.1, CAMO AS, Norway) for data processing.

2.4. Powder X-ray diffraction (PXRD)

The PXRD data were collected for all experiments that were conducted on a Bruker D8 Advanced system in Bragg-Brentano geometry using a Cu Kα radiation point source (λ = 1.5406 Å) at an operating voltage and amperage of 40.0 kV and 40.0 mA, respectively. The powdered samples were analyzed in a low background cell. The samples (approx. 25 mg) were scanned at a rate of 0.005° per minute at step size of 0.01° from 5° to 35° 2θ, resulting in row vectors of 2894 data points for each sample. The obtained PXRD data was exported to Unscrambler® prior to MVA modeling and data fusion.

2.5. Data fusion

The data from FTIR and PXRD were combined using two different fusion methods: fusion of preprocessed data (FPD) and fusion of principal component scores (FPCS). It was necessary to preprocess each set of data individually prior to data fusion. Without preprocessing, the scales for each data set would have been dramatically different and this would have caused inappropriate and unequal weighting in the models. Therefore, the FTIR data set and the PXRD data set were preprocessed using the standard normal variate (SNV) function in the PLS Toolbox (Eigenvector Research, Inc., Wenatchee, WA, USA). The SNV function standardizes the row vectors to mean zero and unit variance. Additionally, the CO_2 peak, including data from 2268 cm^{-1} to 2402 cm^{-1} was removed from each FTIR spectrum and the FTIR data above 3377 cm^{-1} were removed since they did not contain any useful information. After removal of these data points, the FTIR spectra contained 1346 data points. After preprocessing, the data were considered normalized, allowing the multivariate models to apply appropriate weightings to each variable to yield the most descriptive models. For FPD, the normalized data for each sample was fused by concatenating its FTIR row vector with its PXRD row vector, resulting in row vectors with 4240 data points.

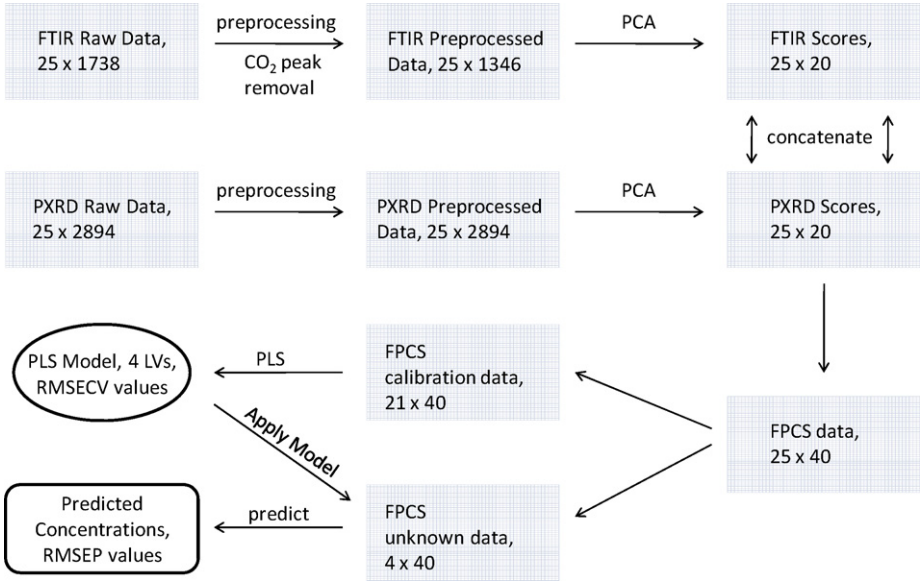

Fig. 1. Schematic flow of data processing in the PLS prediction of the fusion of principal component scores (FPCS) data set. Blocks in the diagram are not intended to represent matrix size. Matrix sizes are represented as M rows × N columns = M samples × N variables.

In FPCS, it was the PCA score values that were fused. A PCA was performed on each individual data set. A PCA was performed on the normalized FTIR spectra of the 21 calibration samples and the 4 unknown samples (a total of 25 samples). The intention of this procedure was to extract as much variation from these samples as possible so the score values of the first 20 principal components were saved. Next, a PCA was performed on the PXRD data set and the score values for its first 20 principal components were saved. The score values from each technique were concatenated into one fused matrix of 25 rows (samples) by 40 columns (score values). Finally, this matrix was separated into a calibration matrix (21 rows by 40 columns) and an unknown matrix (4 rows by 40 columns). Thus, the fused data set was condensed from the respectively large number of variables (4240) to a matrix containing only 40 variables.

2.6. Multivariate analysis (MVA)

Principal component analysis (PCA) followed by partial least square regressions (PLS-2) were performed on the individual FTIR and PXRD data, as well as on the fused data from both FTIR and PXRD analysis, to check the three dimensional spatial distribution of the score values (Unscrambler® 10.0.1, CAMO AS, Norway). Optimized PLS-2 models were used to predict the unknown concentrations of the ingredients in the mixtures. Leave-one-out cross validation was used to calculate the PLS-2 models [12]. The best PLS-2 model for the prediction of the concentrations of the ingredients was selected on the basis of yielding the lowest root mean square of cross validation (RMSECV) values.

Finally, the ability of the optimized PLS models based on an individual FTIR data, PXRD data and the fused data were tested by subjecting the data of four unknown samples mixtures of varying compositions of active ingredients. A schematic of the data processing steps involved in the PLS prediction of the FPCS is shown in Fig. 1.

3. Results and discussion

3.1. Analysis of calibration samples by FTIR and PXRD

The FTIR spectra of the 21 calibration samples were obtained, as indicated in Section 2.3. An overlay plot of the FTIR spectra of the 21 calibration samples is complex due to the large quantity of overlapping data; however, it is instructive to view an overlay plot of the fingerprint spectral region of the 4 vertices (i.e. pure components). Fig. 2 shows that each of the pure components has at least one absorption peak with little overlap from the peaks of other pure components. Vertical lines are included in Fig. 2 to indicate the locations of these unique peaks. APAP has a peak at 1562.1 cm^{-1}, CAF has a peak at 744.4 cm^{-1}, IBU has a peak at 779.1 cm^{-1}, and ASA has a peak at 1562.1 cm^{-1}. This feasibility check shows that there is sufficient variation in the FTIR absorbance spectra to enable multivariate quantitative analysis of these 4 components. A similar analysis was performed on the functional group region of the FTIR spectra (2400–3450 cm^{-1}). It was found that APAP contained a unique peak at 3223 cm^{-1} representing the N–H vibration of its secondary amine, but the rest of peaks in the functional group region were broader than those in the fingerprint region. Therefore, most of the peaks from a single component in the functional

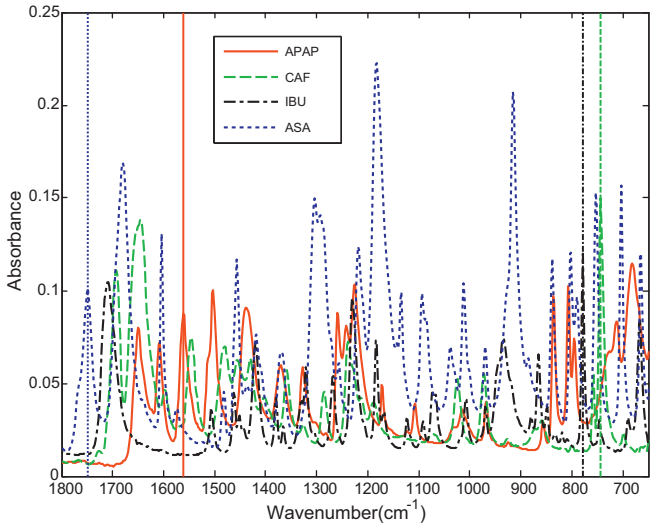

Fig. 2. Fingerprint region of the FTIR spectra of APAP, CAF, IBU, and ASA. Vertical lines show locations of peaks for single components with little overlap from other components.

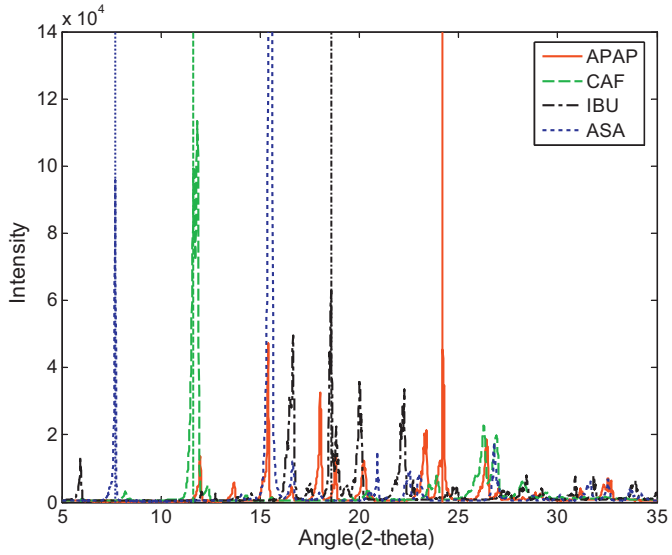

Fig. 3. Powder X-ray diffraction patterns of the 4 pure components: APAP, CAF, IBU, and ASA. Vertical lines show locations of peaks for single components with little overlap from other components. The ASA peak at $2\theta = 15.5^\circ$ was cut off to provide an enlarged view of the majority of peaks in the pattern.

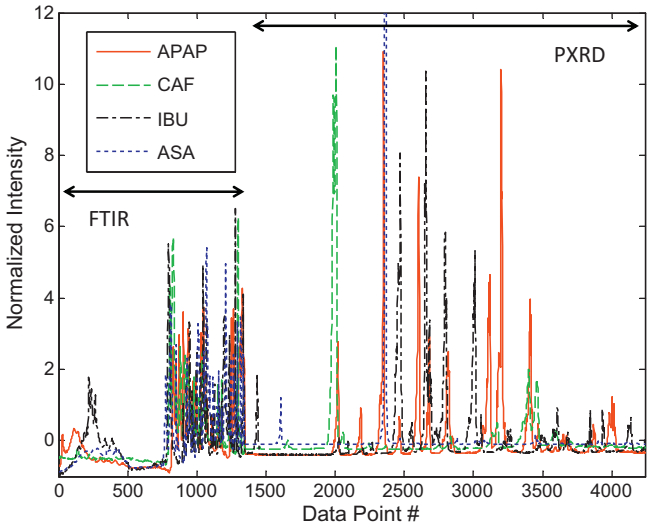

Fig. 4. The fused FTIR spectra and PXRD patterns. The first 1346 data points represent the FTIR spectra and data points 1347–4240 represent the PXRD patterns. SNV was applied separately to each data set prior to concatenation. The tallest ASA peak (18.5 normalized intensity units) was cut off to provide a better view of the remaining peaks.

group region shared overlap with peaks from at least one of the other pure components.

A similar feasibility check was performed on the PXRD patterns of the four pure components. Fig. 3 shows the PXRD patterns for each of the pure components. Vertical lines are included in Fig. 3 to point out that there is at least one peak for each component with little overlap from the peaks of other components. APAP has a peak at $2\theta = 24.2067^\circ$, CAF has a peak at $2\theta = 11.6477^\circ$, IBU has a peak at $2\theta = 18.565^\circ$, and ASA has a peak at $2\theta = 7.6964^\circ$. Again, this feasibility check shows there is sufficient variation in the PXRD pattern to enable multivariate quantitative analysis of the four pure components.

Since it was found that each technique could measure variation in each of the four components, the next step was to fuse the data. In FPD, the data sets were concatenated after preprocessing each individual data set with SNV. The resulting overlay plots of the four pure components for the normalized and fused data vectors are shown in Fig. 4. Data points 1–1346 represented the FTIR spectra and data points 1347–4240 represented the PXRD patterns. While there were more data points in the PXRD data, there was a similar degree of variation in each of the data sets due to preprocessing with SNV. This is ensured because the SNV subtracts the mean of each data set from each data vector, then divides each data vector by the standard deviation each data set. Thus, the FTIR spectra and the PXRD patterns have been placed onto similar normalized scales and the fused data is now ready for multivariate analysis.

3.2. PCA patterns and trends

The purpose of using PCA is usually to elucidate trends or classify samples within data sets. The trends in the 21 calibration samples were clear as these trends were deliberately introduced into the four-component simplex-centroid experimental design. The question is whether the analytical tools were sensitive to those trends and whether the fused data would provide better sensitivity to the trends. To examine these questions, a PCA was first performed on the y-data block (the concentration matrix) of the experimentally designed data set (i.e. the pure component percentages listed in Table 1). The resulting PCA plot is given in Fig. 5A, showing several interesting features. First, 100% of the variation in the data set is compressed into 3 principal components (PCs). Second, the pure component samples (labeled as 1-APAP, 2-CAF, 3-IBU, and 4-ASA) align themselves at the vertices of a trigonal pyramid. Third, each of the calibration samples aligns itself at the expected location within the three dimensional structure. For example, sample #8 is 50% (2-CAF) and 50% (3-IBU); thus aligning itself halfway between samples 2 and 3 while not showing any orientation with samples 1 and 4. Furthermore, the center-points (samples 19, 20, and 21) are precisely overlapping one another and located at the center of the pyramid.

Fig. 5B shows the resulting three-dimensional PCA plots for the FTIR data sets for the 21 calibrations samples. A similar trigonal pyramidal structure emerges, however, the error in the measurement is evident as the calibration samples do not align themselves at the exact locations of the pyramid that would be expected. For example, sample 7, which is 50% (1-APAP) and 50% (4-ASA), is found somewhat half way between samples 1 and 4, but it is now located off of the line connecting vertex points 1 and 4. Fig. 5C shows the corresponding PCA plot for the PXRD patterns for the 21 calibration samples. Again, the same pattern emerges and it appears there is less error in the PXRD data set than in the FTIR data set. Finally, Fig 5D shows the corresponding PCA plot for the FPD. Again, the same pattern emerges and from a qualitative point of view, it appears there is less error in the FPD set than in the FTIR data set, but approximately the same amount of error as in the PXRD data set. Fig. 5B–D shows that the FTIR, PXRD and fused data sets, respectively, track the trends in the concentration variance in the samples (as elucidated in Fig. 5A). The next step is to use PLS to determine the extent to which each of the methods can quantitatively predict the concentrations of each of the components in each of the samples.

3.3. PLS quantitative predictions for calibration samples A

PLS was performed on each of the data sets (FTIR, PXRD, and fused data) to determine the accuracy of prediction for each of the components. Fig. 6 shows an example of the PLS results for the prediction of APAP. Similar trends in results were seen for CAF, IBU, and ASA, although each of the plots is not included here. Fig. 6A shows the PLS prediction based on the FTIR data set for each of the

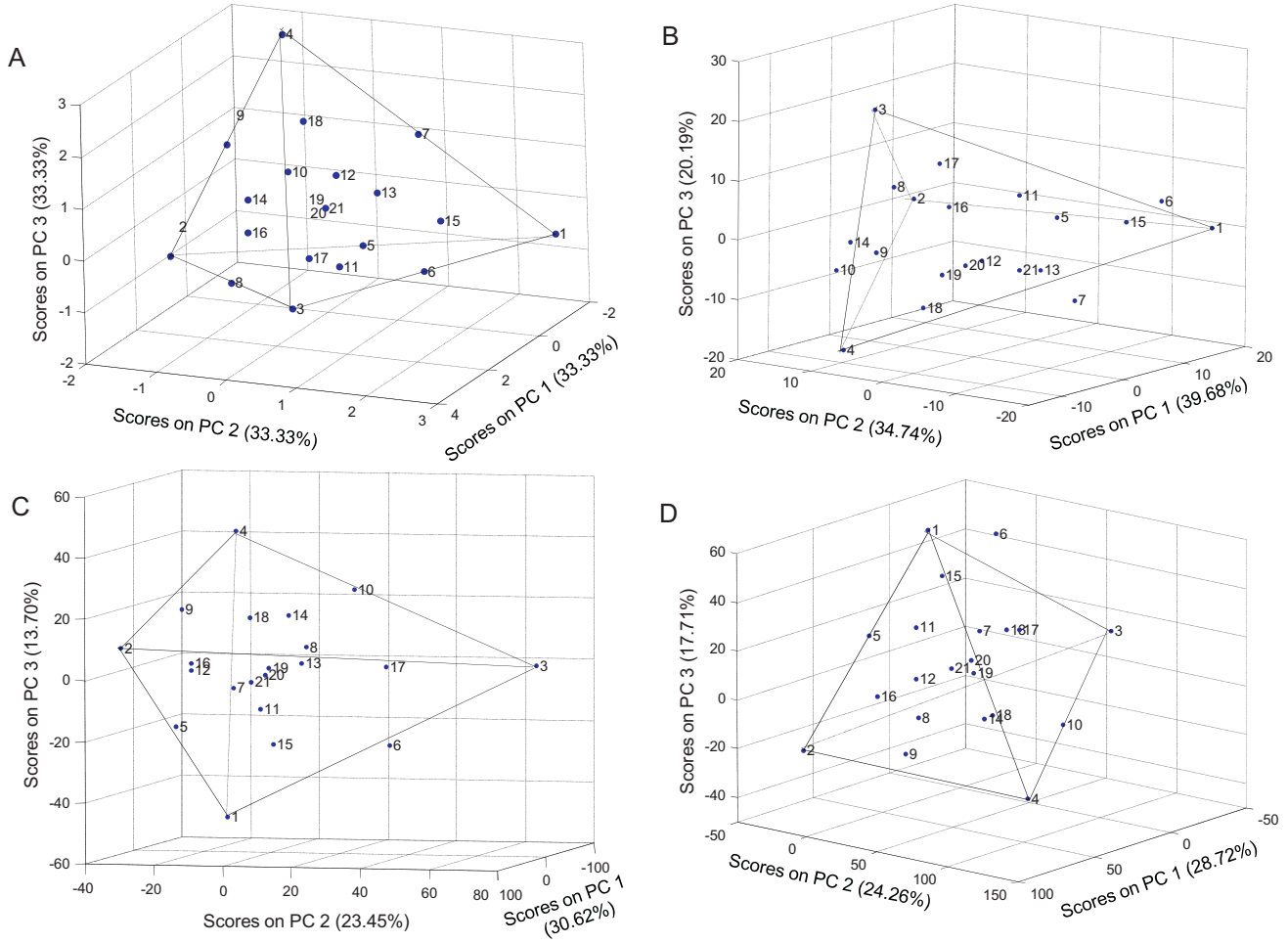

Fig. 5. PCA of (A) the percentages of the pure components in the 21 calibrations samples, (B) the preprocessed FTIR spectra of the 21 calibration samples, (C) the preprocessed PXRD patterns of the 21 calibration samples, and (D) the fusion of preprocessed data (FPD) set representing the 21 calibration samples. The percent variance captured by each PC is shown in parenthesis along each axis.

21 calibration samples. Fig. 6B shows the corresponding plot for the PXRD patterns. As can be seen from the decrease in the scatter in the plot as well as the decrease in RMSEC and RMSECV values, the PLS model for the PXRD patterns did a better job of prediction than the PLS model built on the FTIR data. A PLS model was also built on the FPD set and the results were very similar to those obtained from the PXRD PLS model. However, using the FPCS approach, there was an improvement in the accuracy of the prediction. Fig. 6C shows the corresponding plot for the PLS model that was applied to the FPCS. The RMSEC and RMSECV values included in Fig. 6 indicate that the PXRD prediction (Fig. 6B) is better than the FTIR prediction (Fig. 6A); but the FPCS prediction (Fig. 6C) is better than both the FTIR and PXRD data sets when they are used alone. The RMSECV values were calculated using the "leave one out" cross validation method. Each of the RMSEC and RMSECV values for each of the four components is detailed in Table 2. These results indicate that a FPCS model based on the concatenation of PCA scores can be used to improve prediction accuracy.

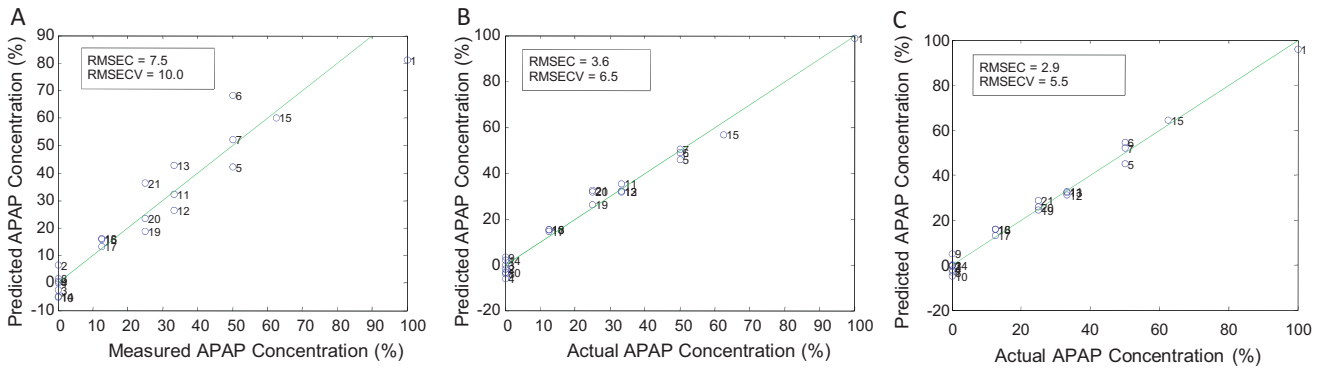

Fig. 6. Representative predicted vs. measured models obtained by partial least square regression (PLS-2) of the (A) FTIR spectra, (B) PXRD patterns, and (C) fusion of principal component scores (FPCS) data for the quantitative prediction of APAP.

Table 2
Summary of partial least square regression (PLS-2) models (RMSEC, root means square error of calibration; RMSECV, root mean square error of prediction; PCs, principal components).

Partial least square regressions				
Parameters	APAP	CAF	IBU	ASA
FTIR				
Optimum no. of PCs used	3	3	3	3
RMSEC (%)	7.50	8.1	9.8	8.1
RMSECV (%)	10.0	11.1	13.7	10.3
PXRD				
Optimum no. of PCs used	3	3	3	3
RMSEC (%)	3.6	3.2	4.2	4.0
RMSECV (%)	6.5	4.3	7.1	6.6
Fusion of principal component scores (FPCS)				
Optimum no. of PCs used	4	4	4	4
RMSEC (%)	2.9	2.6	2.2	3.0
RMSECV (%)	5.5	4.1	5.4	6.6

Table 3
Actual and predicted values of unknown samples of APAP, CAF, IBU and ASA by partial least square regressions (PLS-2) method.

Sample name	Unknown samples				RMSEP
	U1	U2	U3	U4	
Actual concentration					
APAP	33.33	33.33	0	0	
Caffeine	33.33	50	33.33	50	
Ibuprofen	0	16.6	66.66	0	
ASA	33.33	0	0	50	
Predicted concentration					
FTIR					
APAP	25.7	35.4	−2.38	−0.134	4.1
Caffeine	21.2	47.5	46	51.4	8.9
Ibuprofen	3.38	17.5	56.8	8.03	6.6
ASA	49.7	−0.396	−0.37	40.7	9.4
Mean RMSEP					7.3
PXRD					
APAP	30.8	48	3.51	1.35	7.7
Caffeine	30.8	40.7	29.2	40.3	7.1
Ibuprofen	2.16	12.1	69.1	3.68	3.3
ASA	36.2	−0.766	−1.75	54.66	2.9
Mean RMSEP					5.3
Fusion of principle component scores (FPCS)					
APAP	29.3	42.8	0.178	1.38	5.2
Caffeine	30	43	34.9	42.2	5.6
Ibuprofen	4.43	16.6	68.8	5.4	3.7
ASA	36.3	−2.44	−3.86	51	2.8
Mean RMSEP					4.3

3.4. PLS quantitative predictions for blind, unknown samples

The PLS models built in Section 3.3 (the PLS model of the FTIR data, the PLS model of the PXRD data, and the PLS model of the FPCS) were applied to unknown samples. The actual and predicted values for each of the four components in each of the four unknowns are given in Table 3. The results indicate that the PXRD PLS prediction is more accurate than the FTIR PLS prediction, but the FPCS PLS prediction is more accurate than either of the single instrument PLS predictions. Again, the FPD model did not improve the accuracy prediction over the single instrument PLS predictions. Therefore, these results indicate that FPCS, combined with appropriate preprocessing and multivariate statistical analysis, can improve the prediction outcome as compared with the single instrument PLS prediction outcomes. Furthermore, it was found that the FPCS PLS prediction outperformed the FPD PLS prediction.

4. Conclusions

The present work demonstrates the ability of data fusion to combine the information in the FTIR and PXRD data for better quantification and prediction of the amounts of APAP, CAF, IBU and ASA from blended, powder mixtures. PCA of FTIR data, PXRD data, and the FPD indicated that, with some error, each of the data sets showed similar trends to the concentration variation intrinsic in the 21 calibration samples. The use of FPCS was the key to improving the PLS prediction. A comparison of the PLS regression analysis of the FTIR data, the PXRD data, the FPD and the FPCS demonstrated that the FPCS produced the best prediction accuracies of unknown amounts of active pharmaceutical ingredients. The improvement in prediction accuracy of the FPCS method over the FPD method may be attributed to the noise removal and the data reduction associated with extracting the principal component scores prior to building the PLS regression model. The implications for developing distinctive signatures for unknown mixtures as in, e.g., natural products, are an exciting direction under investigation in the UHH laboratories.

References

[1] Guidance for Industry: Q8 Pharmaceutical Development, US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, Rockville, 2006.
[2] A. Gupta, G.E. Peck, R.W. Miller, K.R. Morris, Real-time near-infrared monitoring of content uniformity, moisture content, compact density, tensile strength, and Young's modulus of roller compacted powder blends, J. Pharm. Sci. 94 (2005) 1589–1597.
[3] S. Romero-Torres, J.D. Perez-Ramos, K.R. Morris, E.R. Grant, Raman spectroscopy for tablet coating thickness quantification and coating characterization in the presence of strong fluorescent interference, J. Pharm. Biomed. Anal. 41 (2006) 811–819.
[4] W. Cao, M.P. Mullarney, B.C. Hancock, S. Bates, K.R. Morris, Modeling of transmitted X-ray intensity variation with sample thickness and solid fraction in glycine compacts, J. Pharm. Sci. 92 (2003) 2345–2353.
[5] P.M. Ramos, I. Ruisanchez, K.S. Andrikopoulos, Micro-Raman and X-ray fluorescence spectroscopy data fusion for the classification of ochre pigments, Talanta 75 (2008) 926–936.
[6] V. Steinmetz, F. Sevila, V. Bellon-Maurel, A methodology for sensor fusion design: application to fruit quality assessment, J. Agric. Eng. Res. 74 (1999) 21–31.
[7] P. Boilot, E.L. Hines, M.A. Gongora, R.S. Folland, Electronic noses intercomparison, data fusion and sensor selection in discrimination of standard fruit solutions, Sens. Actuators B 88 (2003) 80–88.
[8] S. Roussel, W. Bellon-Maurel, J.M. Roger, Authenticating white grape must variety with classification models based on aroma sensors, FT-IR and UV spectrometry, J. Food Eng. 60 (2003) 407–419.
[9] J. Esteban, A. Starr, R. Willetts, P. Hannah, P. Bryanston-Cross, A review of data fusion models and architectures: towards engineering guidelines, Neural Comput. Appl. 14 (2005) 273–281.
[10] J. Llinas, D.L. Hall, An introduction to multi-sensor data fusion, Proc. IEEE Int. Symp. Circuit Syst. 6 (1998) 537–540.
[11] K. Esbensen, S. Schoenkopf, T. Midtgaard, D. Guyot, Multivariate Data Analysis – In Practice, Camo ASA, Trondheim, Norway, 1994.
[12] H. Martens, M. Martens, Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR), Food Qual. Pref. 11 (2000) 5–16.