Anal. Chem. **2007,** *79,* 1231−1236

Grazing-Angle Fiber-Optic Fourier Transform Infrared Reflection−Absorption Spectroscopy for the in Situ Detection and Quantification of Two Active Pharmaceutical Ingredients on Glass

Benjamin B. Perston,† Michelle L. Hamilton,† Bryce E. Williamson,*,† Peter W. Harland,† Mary A. Thomson,‡ and Peter J. Melling‡

Department of Chemistry, University of Canterbury, Christchurch, New Zealand, and Remspec Corporation, Sturbridge, Massachusetts 01566

Fourier transform infrared reflection−absorption spectroscopy has been used with a fiber-optic grazing-angle reflectance probe as a rapid, in situ method for trace surface analysis of acetaminophen and aspirin at loadings of ∼0−2 μg cm^{-2} on glass. Partial least-squares multivariate regression permits the loadings to be quantified, simultaneously, with root-mean-squared errors of prediction of RMSEP ≈ 0.1 μg cm^{-2} for both compounds. The detection limits are estimated to be L_D ≈ 0.2 μg cm^{-2}.

Clean equipment is important in many manufacturing situations: in particular, the pharmaceutical industry is subject to significant legislative requirements regarding cleaning and cleaning validation.[1] Acceptable residual limits (with dimensions mass/area) for active pharmaceutical ingredients (APIs) are calculated on the basis of the relative potencies of the residual API and the next API to be produced, as well as the ratio between the surface area of the equipment and the mass of API produced in a batch.[2] An upper limit is usually set (regardless of the calculated acceptable limit) at the point where the residue is visible to the naked eye.[3] This limit clearly depends on the contaminant and viewing conditions as well as the surface material and finish, but is typically[4] ∼1−4 μg cm^{-2}.

Currently, cleaning is validated predominantly by analysis of rinsewater[5] or of swabs taken from one or several regions of the equipment.[6] The analytical method often involves chromatographic separation (usually HPLC) for selectivity.[6] While they allow straightforward, sensitive, linear calibrations,[7,8] such methods are time-consuming and involve a significant period between sample collection and analysis, during which equipment usage is delayed by uncertainty of its cleanliness.

In situ infrared reflection−absorption spectroscopy (IRRAS; also called infrared external reflection spectroscopy) has the potential to provide much more rapid analyses than the commonly used wet-chemical methods. Modern FT-IR instruments are extremely stable and have good sensitivity, and infrared fiber-optics based on chalcogenide glasses can allow sampling up to several meters from the spectrometer.[9] Multivariate chemometric tools, such as partial least-squares (PLS) regression,[10] allow the implicit modeling of unquantified interfering species (such as atmospheric gases, surfactants and excipients) and instrumental artifacts (such as baseline changes). Mid-infrared absorption bands for organic compounds are sharp, so good selectivity is possible without the need for a separation step.

IRRAS has been used to study thin films on metallic,[11] semiconductor, and dielectric[12] substrates. Most of these studies have involved carefully prepared films produced, for example, by self-assembly or the Langmuir−Blodgett technique. The theory behind the measurements is well understood,[13,14] and detailed structural information about the films can be obtained, provided that the films are homogeneous and the substrate is flat and smooth. For organic films with thickness much less than the radiation wavelength (∼2−10 μm in the mid-infrared) on a glass substrate, IRRAS intensity is proportional to film thickness.[15]

In the setting of pharmaceutical cleaning validation, optically thin sample layers (that is, those that are difficult to detect visually) will most likely have been deposited by evaporation of contaminated solvent from the equipment surface. The physical presenta-

* Corresponding author. E-mail: bryce.williamson@canterbury.ac.nz.
† University of Canterbury.
‡ Remspec Corp.

(1) *Guide to Inspections of Validation of Cleaning Processes*; FDA: Rockville, MD, 1993.
(2) LeBlanc, D. A. *Pharm. Technol.* **2000**, *24*, 160−168.
(3) Fourman, G.; Mullen, M. *Pharm. Technol.* **1993**, *17*, 54−60.
(4) Jenkins, K. M.; Vanderwielen, A. J. *Pharm. Technol.* **1994**, *18*, 60−73.
(5) LeBlanc, D. A. *Pharm. Technol.* **1998**, *22*, 66−74.
(6) Shifflet, M. J.; Shapiro, M. *Am. Pharm. Rev.* **2002**, *5*, 35−41.
(7) Klinkenberg, R.; Streel, B.; Ceccato, A. *J. Pharm. Biomed. Anal.* **2003**, *32*, 345−352.
(8) Mirza, T.; Lunn, M. J.; Keeley, F. J.; George, R. C.; Bodenmiller, J. R. *J. Pharm. Biomed. Anal.* **1999**, *19*, 747−756.
(9) Melling, P. J.; Thomson, M. In *The Handbook of Vibrational Spectroscopy*; Chalmers, J. M., Griffiths, P. R., Eds.; Wiley: Chichester, 2002; Vol. 2, pp 1551−1559.
(10) Martens, H.; Næs, T. *Multivariate Calibration*; Wiley: Chichester, 1989.
(11) Umemura, J. In *The Handbook of Vibrational Spectroscopy*; Chalmers, J. M., Griffiths, P. R., Eds.; Wiley: Chichester, 2002; Vol. 2, pp 982−998.
(12) Jattner, J.; Hoffmann, H. In *The Handbook of Vibrational Spectroscopy*; Chalmers, J. M., Griffiths, P. R., Eds.; Wiley: Chichester, 2002; Vol. 2, pp 1009−1027.
(13) Heavens, O. S. *Optical Properties of Thin Solid Films*; Dover: New York, 1991.
(14) Parikh, A. N.; Allara, D. L. *J. Chem. Phys.* **1992**, *96*, 927−945.
(15) Tolstoy, V. P.; Chernyshova, I. V.; Skryshevsky, V. A. *Handbook of Infrared Spectroscopy of Ultrathin Films*; Wiley: Hoboken, NJ, 2003.

10.1021/ac061660a CCC: $37.00
Published on Web 01/05/2007

tion of the contamination will depend on the solvent, the chemical nature and quantity of the residual material, and the substrate surface. Previous work has shown that IRRAS calibrations can be established for individual surfactants and APIs on metallic and glass surfaces,[16,17] for an API in the presence of a surfactant on a glass surface,[17] and for an API in the presence of excipients on a stainless steel surface.[18] But the ability to quantify several APIs with a single rapid in situ measurement would also be valuable.

The work presented here concerns the simultaneous quantification of two chemically similar APIs on glass. Aspirin and acetaminophen were selected as model compounds since they are inexpensive and safe to handle. However, since almost all organic compounds exhibit strong infrared absorptivity, the technique will be of general applicability.

EXPERIMENTAL SECTION

Materials. Aspirin (*o*-acetylsalicylic acid) and acetaminophen (4-acetamidophenol) were obtained from Sigma Aldrich and used without further purification. The solvents used were Milli-Q water, ethanol (solvent grade), and acetone (solvent grade) or mixtures of these components. Glass coupons (15×15 cm^2) were cut from 3-mm-thick window (soda lime float) glass and were roughened by bead-blasting on one side to prevent reflection from the back face.

Sample Preparation. The spray method of sample preparation[17] was used for all results reported here. This gives more uniform coatings (in terms of the macroscopic distribution of the analyte) than the smear technique[17] but requires an additional calibration step to determine the loadings. Solvents of various volatility were used (acetone, ethanol, water, and mixtures thereof), with the spraying distance adjusted, accordingly, between ~0.5 and 1 m. A less volatile solvent allows the airbrush to be held further from the coupon, facilitating the preparation of more homogeneous samples. Both APIs were dissolved (at various, known concentration ratios) in the same spraying solution.

IRRAS Instrumentation and Data Collection. The FT-IR instrumentation has been described previously.[17,19] Briefly, the system comprises a prototype Remspec SpotView grazing-angle head coupled to an interferometer by a 1.5-m, 19-fiber chalcogenide−glass optical cable. Within the head, off-axis parabolic mirrors collimate the beam onto the sample (at an incidence angle of 80° to the surface normal) and focus the reflected light to an integrated MCT detector. The IR footprint produced by the grazing-angle probe is elliptical (~10 cm × 3 cm) with an intensity profile that decays from the center toward the edges. Transmission measurements were made using the internal sample compartment of the same spectrometer with a DTGS detector.

The wavenumber range was 4000−1000 cm^{-1}, and the resolution was 4 cm^{-1}. Single-beam background spectra (I_0) were obtained from clean glass coupons by averaging 100 interferometer scans (~30 s). Five to ten sample spectra (I) were collected from different regions of each loaded coupon by averaging 50 interferometer scans (~15 s) per spectrum. The IRRAS was calculated as $\log(I_0/I)$. Since there is a considerable (~20 cm) beam path through laboratory air outside the spectrometer, absorbance bands arising from small changes in CO_2 and H_2O vapor concentrations often appear in the spectra. To test whether these features pose any problem for the chemometric methods, a single background spectrum was used for each batch of four to eight samples, rather than for each sample. The result of this is that, in most spectra, the atmospheric absorption bands are stronger than the bands due to the analyte.

Primary Calibration. The loading (mass per unit area) of each sample was determined after the measurement of IRRAS. The sample coupons were rinsed with 50 or 100 mL of ethanol, and the analysis was conducted by UV colorimetry using a GBC 920 UV/visible spectrometer. The spectra of eight standards were measured over the range 210−340 nm at ~2-nm intervals with a resolution of 2.0 nm and a scan speed of 130 nm/min. The full-spectrum classical least-squares method[20] was used for calibration and subsequent prediction. The uncertainty in the reference method was estimated by analyzing eight fresh standards and calculating the root-mean-squared error of prediction (RMSEP), which was found to be ~0.013 μg cm^{-2} for both compounds.

Two tests verified that the entire analyte loadings were recovered from the coupons. The first involved depositing known loadings by pipet on a clean coupon and verifying that the analysis method gave the correct result; the second involved collecting an additional volume of rinsate for the samples and verifying that the analysis result was insignificant.

RESULTS AND DISCUSSION

IRRA Spectra. Typical IRRAS of acetaminophen and aspirin on glass are shown in Figure 1 (2000−1000 cm^{-1}) where they are compared with absorption spectra obtained in transmission from pressed KBr pellets. The effective loadings for the pellets were determined by dividing the mass of API (~1 mg) by the pellet area (1.33 cm^2). The transmission spectra in Figure 1 have been scaled by the ratio of the IRRAS sample loading to the KBr pellet sample loading.

In contrast with transmission spectra, the IRRAS can exhibit both positive and negative features, depending on the refractive indices of the substrate and film materials, the polarization of the light, and the incidence angle.[15] In addition, band shape distortions due to optical effects[21] and spectral features due to the substrate can appear.[22] The strong positive feature at ~1280 cm^{-1} in the IRRAS is due to the Si−O stretching mode of the substrate. As a consequence of this strong resonance, the real part of the refractive index crosses that of the incident medium (air; $n \approx 1$) while the imaginary part remains very small. This results in a deep minimum in the reflectance of the bare substrate and a correspondingly greater relative change in the reflectance due to the layer. To the blue (higher frequency) of the substrate feature the API bands are negative, while to the red they are positive.

Model Optimization by Cross-Validation. The chemometric modeling was achieved by using the PLS-1 variant of the PLS regression method,[23] which permits different spectroscopic re-

(16) Mehta, N. K.; Goenaga-Polo, J.; Hernandez-Rivera, S. P.; Hernandez, D.; Thomson, M. A.; Melling, P. J. *Biopharmacology* **2002**, *15*, 36−71.
(17) Hamilton, M. L.; Perston, B. B.; Harland, P. W.; Williamson, B. E.; Thomson, M. A.; Melling, P. J. *Org. Process Res. Dev.* **2005**, *9*, 337−343.
(18) Teelucksingh, N.; Reddy, K. B. *Spectroscopy* **2005**, *20*, 16−20.
(19) Melling, P. J.; Shelley, P. H. U.S. Patent 6,310,348, 2001.
(20) Beebe, K. R.; Pell, R. J.; Seasholtz, M. B. *Chemometrics: A Practical Guide*; Wiley: New York, 1998.
(21) Yamamoto, K.; Ishida, H. *Vib. Spectrosc.* **1994**, *8*, 1−36.
(22) Mielczarski, J. A.; Mielczarski, E. *J. Phys. Chem. B* **1999**, *103*, 5852−5859.
(23) Haaland, D. M.; Thomas, E. V. *Anal. Chem.* **1988**, *60*, 1193−1202.

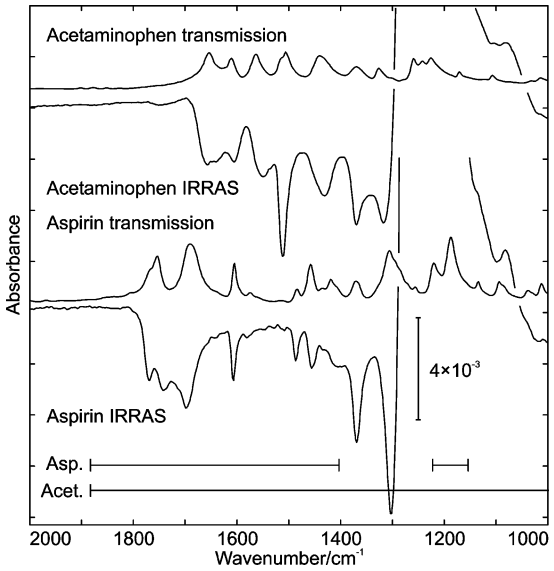

Figure 1. IRRA (glass substrate) and absorption (in transmission; KBr pellet) spectra of acetaminophen and aspirin. The transmission spectra have been scaled to the same effective loading as the IRRAS (acetaminophen 2.08 μg cm^{-2}; aspirin 2.95 μg cm^{-2}). The spectra have been offset along the ordinate for clarity; the scale is indicated. The horizontal bars labeled Asp and Acet indicate the wavenumber ranges used to build PLS-1 chemometric models for aspirin and acetaminophen, respectively.

gions and ranks (number of PLS factors) to be used for each analyte. The procedure was implemented in MATLAB[24] using algorithms from the literature,[23,25] with cross-validations and test set validations.[20,23,26] It entails two parts: first, generation of an optimized calibration model and, second, evaluation of the quality of the model. To ensure that the optimization process did not bias the quality evaluation, the 45 samples were split into separate calibration (31 samples) and test (14 samples) sets, chosen to cover simultaneously the loading range ~0−2 μg cm^{-2} for both compounds. (The loadings are given in Figure S-1 of the Supporting Information.) Using a larger proportion of the standards for the test set would have improved the precision of the quality evaluation, but at the expense of the performance of the model. For optimization, model quality was judged in terms of the rms error of cross-validation (RMSECV). Once the model parameters had been selected, the rms error of prediction for the test set (RMSEP) was calculated by applying the optimized model to the test set.

The wavenumber ranges for the chemometric models were selected with the aim of maximizing the incorporation of analyte bands (Figure 1) while minimizing the influence of features that are uncorrelated with the analyte loading. Of several ranges that were used in trials, those indicated in Figure 1 and Table 1 gave the lowest RMSECV values and were used for the results presented here. The range chosen for acetaminophen essentially encompasses the "fingerprint" region. For aspirin, a narrower range was used and the prominent Si−O feature was excluded. Inclusion of the API O−H and N−H bands did not improve the

Table 1. Model Parameters and Statistics for Cross-Validations and Test Set Validations for Acetaminophen and Aspirin Mixtures on Glass[a]

	acetaminophen	aspirin	
loading range/μg cm^{-2}	0−1.7	0−2.2	
wavenumber range/cm^{-1}	1880−1000	1880−1398 and 1225−1157	
rank	8	8	11
cross-validations:			
RMSECV/μg cm^{-2}	0.12	0.12	0.12
RMSECV$_{<1}$/μg cm^{-2}	0.09	0.10	
R^2	0.96	0.97	0.97
p_1	0.10	0.14	0.67
p_2	0.27	0.21	0.63
test set validations:			
RMSEP/μg cm^{-2}	0.09	0.15	0.13
RMSEP$_{<1}$/μg cm^{-2}	0.06	0.08	0.06
R^2	0.97	0.96	0.97
p_1	0.24	0.01	0.07
p_2	0.51	0.002	0.17

[a] RMSECV$_{<1}$ and RMSEP$_{<1}$ are calculated from the residuals of standards with analyte loadings of <1 μg cm^{-2}.

RMSECV. Optimal ranks (those judged to give the highest quality models) for each analyte were determined by two methods[23,27] that gave the same results.

Several spectroscopic preprocessing procedures were investigated with the aim of building better-optimized models. Mean centering, which is commonly used with PLS but is unsuitable for some kinds of data,[28] did not reduce the optimal rank or significantly reduce the RMSECV. The first-derivative, (Savitzky−Golay method[29,30] with 15 smoothing points) reduced the optimal rank but did not significantly change the RMSECV. The results that follow were obtained in the absence of preprocessing.

In the cross-validations, the spectra for each sample were treated collectively, being either included or left out together. The resultant RMSECV values, plotted against rank in Figure 2, exhibit broad minima between ranks of ~8 and 12. The optimal ranks were 8 for each compound (Figure 2 and Table 1), corresponding to the left edge of the broad minimum.

The cross-validation plots for acetaminophen (Figure 3) and aspirin (Figure S-2 of the Supporting Information) at a rank of 8 are very similar and indicative of a satisfactory model. As the loadings increase, the spread of the predictions for individual spectra for each sample about the mean for that sample increases, but the deviations of the means from the diagonal ideal-fit (zero intercept and unit slope) line do not change significantly. This apparent heteroscedasticity is at least partly due to sample heterogeneity, as discussed below.

Test Set Validations and Bias Tests. The quality evaluations were performed by applying the calibration models to the test sets. The results for acetaminophen (Figure S-3 of the Supporting Information) essentially mirror those of the cross-validation. However, as shown in Figure 4a, the predictions for the test set of aspirin appear to show a systematic bias (referred to here as type-1 bias) to low values, with only two of the mean predictions

(24) MATLAB 6.5 Release 13; The Mathworks, Inc.: Natick, MA, 2002.
(25) Dayal, B. S.; MacGregor, J. F. *J. Chemom.* **1997**, *11*, 73−85.
(26) Kramer, R. R. *Chemometric Techniques for Quantitative Analysis*; New York, 1998.
(27) Martens, H. A.; Dardenne, P. *Chemom. Intell. Lab. Syst.* **1998**, *44*, 99−121.
(28) Seasholtz, M. B.; Kowalski, B. R. *J. Chemom.* **1992**, *6*, 103−111.
(29) Savitzky, A.; Golay, M. J. E. *Anal. Chem.* **1964**, *36*, 1627−1639.
(30) Madden, H. H. *Anal. Chem.* **1978**, *50*, 1383−1386.

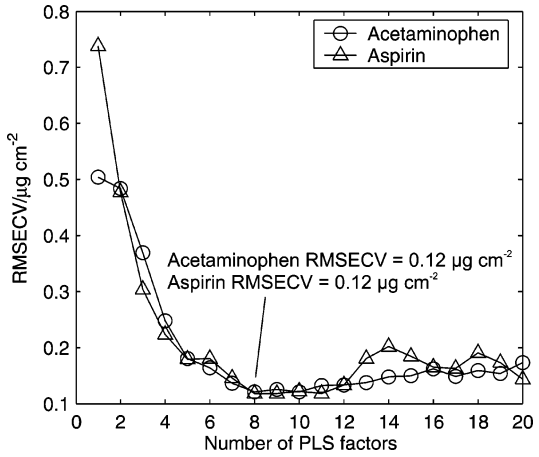

Figure 2. Root-mean-square errors of cross-validation plotted against rank for acetaminophen (circles) and aspirin (triangles) on glass. The optimal model ranks (determined by the methods of refs 23 and 27) are 8 for both APIs; the corresponding RMSECV values are indicated.

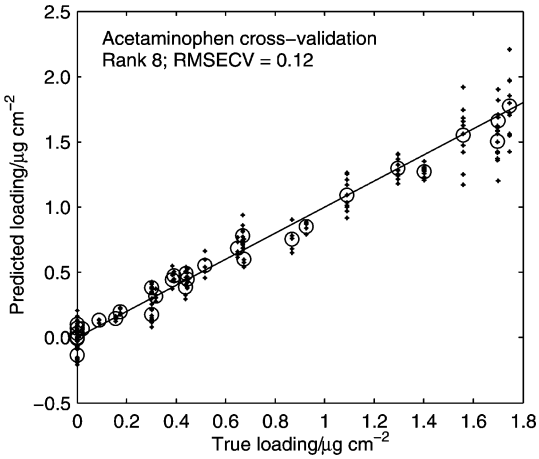

Figure 3. Predicted versus true loadings for the acetaminophen cross-validation with a rank of 8. Dots are predictions from individual spectra while open circles are mean predictions per sample. The solid diagonal (zero intercept and unit slope) is the line of perfect agreement.

falling above the ideal-fit line. A statistical test, based on the joint confidence region for the slope and intercept of a least-squares line fitted through the predicted and true loadings,[31] confirms that this bias is significant ($p_1 = 0.01 < \alpha$) at the $\alpha = 0.05$ level. By the same measure, type-1 bias is absent from the other rank-8 cross and test set validations (Table 1).

A known cause of bias in PLS regression is the use of too few factors to account for all of the relevant variation in the spectra.[32] Consistent with this, when the rank for aspirin is increased, the type-1 bias is reduced. At a rank of 11, the right edge of the broad minimum in Figure 2, it is insignificant ($p_1 = 0.07$) and the RMSEP is slightly decreased (Figure 4b and Table 1).

A second kind of bias (referred to here as type-2 bias) occurs when the loading of one compound has a systematic effect on the predicted loading of the other. This would not necessarily show up in the test for type-1 bias since the loadings of the two compounds are uncorrelated (Figure S-1). However, it would be revealed by an equivalent test that investigates correlation between the residuals (differences between predicted and true loadings) for one compound and the loadings of the other, where the ideal-fit line now has zero slope and intercept. For acetaminophen (Figure S-4), this test indicates no significant type-2 bias in either the cross-validation or test set validation at a rank of 8 (Table 1). For aspirin (Figure 5a), type-2 bias is insignificant in the rank-8 cross-validation, but it is significant ($p_2 = 0.002$) in the test set validation at that rank. As was the case for type-1 bias, the type-2 bias weakens as more factors are included, and for the 11-factor model (Table 1 and Figure 5b), $p_2 = 0.17$.

Sample Heterogeneity. An important limitation of the data presented here is that IRRA spectra are measured from several different regions of each sample, whereas the loading determined by the reference method is the average for the entire coupon. Although the latter can be determined with good precision, sample heterogeneity means that it may not accurately represent the true, local loading pertaining to any particular IRRAS measurement. The extent of the loading heterogeneity (variations in the local analyte loading in the region actually sampled by the probe) was estimated by spraying coupons constructed from four separate strips and then determining the loading of each strip by the UV colorimetric method. The relative standard deviation so determined (after correcting for the difference in area between the

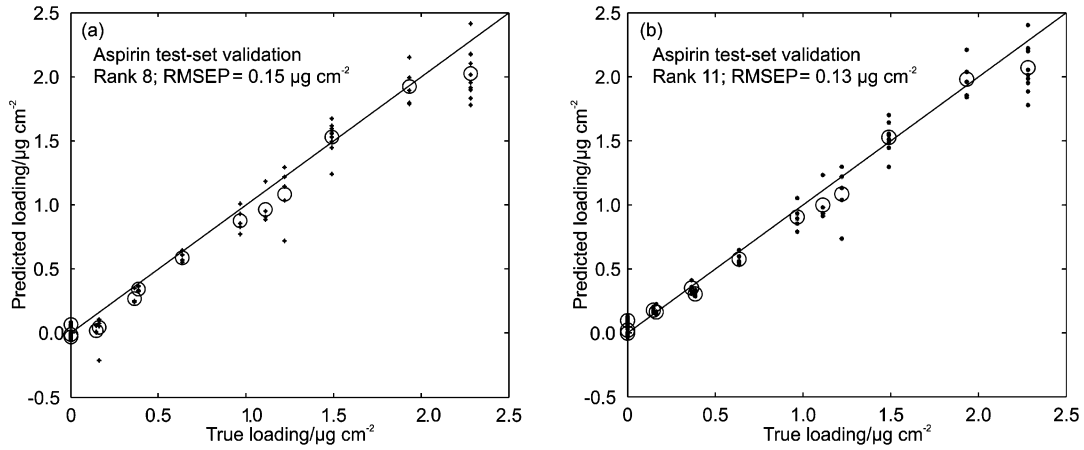

Figure 4. Predicted versus true loadings for the aspirin test set validations with ranks of 8 (a) and 11 (b). Dots indicate predictions from individual spectra, while open circles show the mean predictions per sample. The solid diagonals (zero intercept and unit slope) are the lines of perfect agreement.

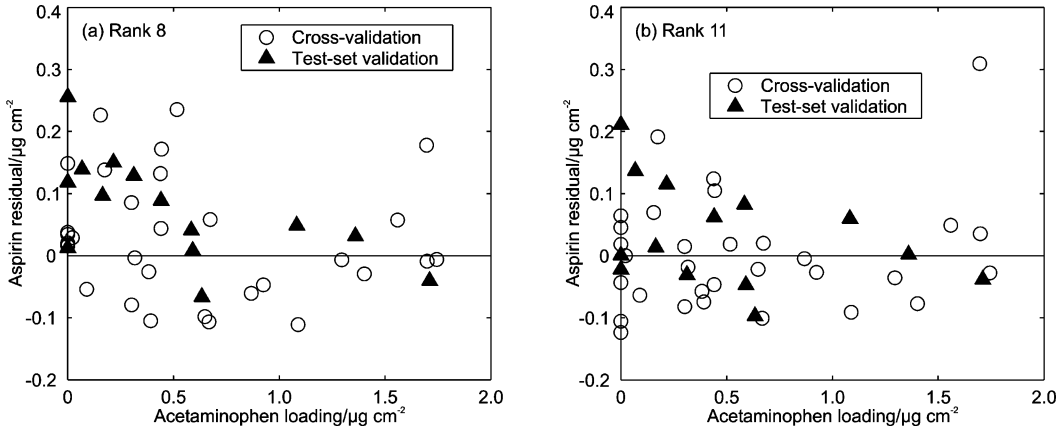

Figure 5. Loading residuals for the cross-validations (open circles) and test set (solid triangles) validations of aspirin plotted against acetaminophen loading for ranks of 8 (a) and 11 (b). The only significant type-2 bias revealed (at the $\alpha = 0.05$ level) in these data is for the rank-8 test-set validation for which $p_2 = 0.002$.

strips and the IR footprint) was ~8%. This error is equivalent tothe presence of significant and heteroscedastic errors in the reference loading values.

There are two consequences of these errors. First, as a result of the heteroscedasticity of the errors, unweighted regression methods will not give the best estimate of the regression vector, but due to the absence of appropriate algorithms for weighted multivariate regressions in most commercial software (and since the number of measurements is sufficient that improvements to the model would probably be minor), we have not attempted to address this issue. Second, the rms prediction errors in Table 1 represent convolutions of contributions intrinsic to the spectroscopic (IRRAS) method with contributions arising from sample heterogeneity and are therefore pessimistic estimates of the true accuracy of the spectroscopic method.[33] This inflation of the prediction error occurs whenever the reference errors are sizable, and other workers have attempted to correct it by subtracting the estimated error in the reference values from the RMSEP.[32,33] Heteroscedastic reference method errors do not imply heteroscedastic errors in the IRRAS method, and the subtraction method may still be valid: however, it is not immediately obvious how to generalize it to accommodate the heteroscedastic reference errors, and this approach has not been pursued here.

Samples will also exhibit heterogeneity relating to variations of thickness and roughness. This issue is not addressed here, but is the subject of parallel investigations described elsewhere.[34,35] The success of the calibrations reported in this and related, earlier work[17,34] indicates that the effects of this kind of heterogeneity either are insignificant or are averaged out during the FT-IR measurements.

Detection Limits. A method for cleaning validation is more usefully characterized by a detection limit than by an RMSEP. Prediction uncertainties and figures of merit for multivariate calibration are topics of active research in chemometrics.[36] Formally, in multivariate inverse regression, the detection limit does not have a fixed value but varies from sample to sample;[37] however, as shown below, provided that the calibration set is sufficiently large, the variability of the detection limit for most samples is small enough that a single value can usefully characterize the performance of the analytical method.

If measurement errors in the spectroscopic and reference methods are neglected, an approximate expression for the error variance of the predicted loading, y_u, for a new sample is[37]

$$V(y_u) \approx (1 + h_u) \times \mathrm{MSEP} \qquad (1)$$

where MSEP = $(\mathrm{RMSEP})^2$ and h_u is the leverage of the new sample, a weighted measure of its distance from the origin of the model space. The leverage depends on the contributions to the spectrum of the analyte and all interfering species. Samples with greater leverage have greater uncertainty in their predicted loadings. However, if there are many degrees of freedom in the calibration set (many calibration standards and/or few factors) and the new sample is not unusual, the leverage is likely to be negligible ($h_u \ll 1$). In this work, the mean test set leverages are small (0.03 for the 8-factor acetaminophen model and 0.05 for the 11-factor aspirin model), although there are several spectra with leverages of 0.1 or greater.

For all samples with negligible leverage, the detection limit can be obtained as the product of the RMSEP and an appropriate t statistic, following the IUPAC recommendation[36] of considering both false positive and false negative detection decisions. The critical level, L_C, is the loading that the prediction for a blank sample could erroneously exceed with probability α. In the limit of negligible leverage, it is given by

$$L_C = t_{1-\alpha,n} V(y_u)^{1/2} \approx t_{1-\alpha,n} \times \mathrm{RMSEP} \qquad (2)$$

where n is the number of degrees of freedom associated with the

(31) Mandel, J.; Linnig, F. J. *Anal. Chem.* **1957**, *29*, 743−749.
(32) Faber, N. M.; Kowalski, B. R. *J. Chemom.* **1997**, *11*, 181−238.
(33) DiFoggio, R. *Appl. Spectrosc.* **1995**, *49*, 67−75.
(34) Hamilton, M. L.; Perston, B. B.; Harland, P. W.; Williamson, B. E.; Thomson, M. A.; Melling, P. J. *Appl. Spectrosc.* **2006**.
(35) Perston, B. B. Ph.D. thesis, University of Canterbury, Christchurch, New Zealand, 2006.
(36) Olivieri, A. C.; Faber, N. M.; Ferré, J.; Boqué, R.; Kalivas, J. H.; Mark, H. *Pure Appl. Chem.* **2006**, *78*, 633−661.
(37) Boqué, R.; Larrechi, M. S.; Rius, F. X. *Chemom. Intell. Lab. Syst.* **1999**, *45*, 397−408.

RMSEP (i.e., the number of test set standards) and the t statistics are single-tailed. The detection limit, L_D, is then the loading for which there is a probability β that the model will falsely return a predicted value lower than L_C (leading to an erroneous "not detected" decision); it is given by

$$L_D = L_C + t_{1-\beta,n} V(y_u)^{1/2} \approx (t_{1-\alpha,n} + t_{1-\beta,n}) \times \text{RMSEP} \quad (3)$$

With $\alpha = \beta = 0.05$, eq 3 yields detection limits of ~0.3 and ~0.4 μg cm^{-2} for acetaminophen and aspirin, respectively. If the test set is large, the RMSEP estimate has many degrees of freedom and the sum of the t statistics is ~3.3, so eq 3 is equivalent to the "three times the standard deviation of the noise" rule of thumb.

It must be stressed that, in multivariate calibration, there is no guarantee that a low-loading sample will have low leverage, so while the detection limits calculated in this manner are indicative of the general performance of a model, the sample-specific detection limit, taking into account the leverage, should be calculated for all new samples.

Because of the significant contribution of error in the reference loadings, the RMSEP values and hence these detection limit estimates are biased high. Due to the heteroscedasticity in the reference loading errors, this effect is exacerbated by the inclusion of samples with higher loadings. RMSECV and RMSEP values calculated using the residuals of only the standards with loadings of <1 μg cm^{-2} in the analyte of interest are also given in Table 1. These values correspond to detection limits of 0.2 μg cm^{-2} for both APIs.

CONCLUSIONS

The results presented here demonstrate that grazing-angle fiber-optic IRRAS can be used to quantify simultaneously two chemically similar APIs on a glass surface at loadings well below those visible to the naked eye, without making frequent background measurements or taking any special precautions regarding absorption by atmospheric gases. This is made possible by the use of PLS regression, although other multivariate inverse calibration methods are also likely to be suitable.

The results presented here relate to simultaneous quantification of just two surface contaminants, but the method can be generalized to any number of components whose spectra are sufficiently distinct. Its main advantages over the traditional swab-HPLC technique are that it samples those contaminants that are actually on the surface, rather than those that can be removed by the swab, and is much more rapid. With regard to the latter of these points, the in situ spectroscopic measurements can be made in ~15 s or less using a conventional laboratory spectrometer. The use of the chemometric calibration to determine the loadings requires just a few more seconds, so that the entire process of quantification can be completed in significantly less than 1 min after the grazing-angle head has been placed on the sample surface.

An important issue underlined by this work is the risk of underfitting when employing the usual safeguards against overfitting. Methods such as those suggested by Martens and Dardenne[27] and by Haaland and Thomas,[23] are intended to strike a balance between excessive bias due to underfitting and excessive variance due to overfitting, but in this work, they seemed to err on the side of underfitting, as illustrated by the significant bias encountered in the aspirin test set validation at rank 8.

Later work will address the issue of separating the error intrinsic to the method from error due to sampling of a heterogeneous standard. In application to real samples, the sampling error can be mitigated by making measurements at several different physical locations on the substrate surface, so it is important to have an accurate estimate of the error pertaining to a homogeneous sample. In the first instance, additional effort will be devoted to improving the homogeneity of the standards and to more thorough treatments of prediction intervals and detection limits. Finally, it should be noted that other factors can contribute to heteroscedastic deviations. For example, as the loading increases, the risk of nonlinear spectroscopic response also increases: this would introduce another loading-dependent contribution to the error arising from lack of model fit. Provided that homogeneous standards are available, it may be possible to account for such effects by employing nonlinear multivariate methods.

ACKNOWLEDGMENT

B.B.P. acknowledges financial assistance from a University of Canterbury Doctoral Scholarship and a New Zealand Vice Chancellors' Committee William Georgetti Scholarship. M.L.H. thanks Remspec Corporation for provision of a graduate student scholarship.

SUPPORTING INFORMATION AVAILABLE

Plots of the calibration and test set loadings along with additional cross-validations and test set validations are available, as referred to in the text. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review September 4, 2006. Accepted November 14, 2006.

AC061660A