We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

5,000
Open access books available

125,000
International authors and editors

140M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter

Mid-Infrared Laser Spectroscopy Applications in Process Analytical Technology: Cleaning Validation, Microorganisms, and Active Pharmaceutical Ingredients in Formulations


Abstract

Mid-infrared (MIR) lasers are very high-brightness energy sources that are replacing conventional thermal sources (globars) in many infrared spectroscopy (IRS) techniques. Although not all laser properties have been exploited in depth, properties such as collimation, polarization, high brightness, and very high resolution have contributed to recast IRS tools. Applications of MIR laser spectroscopy to process analytical technology (PAT) are numerous and important. As an example, a compact grazing angle probe mount has allowed coupling to a MIR quantum cascade laser (QCL), enabling reflectance-absorbance infrared spectroscopy (RAIRS) measurements. This methodology, coupled to powerful multivariable analysis (MVA) routines of chemometrics and fast Fourier transform (FFT) preprocessing of the data resulted in very low limits of detection of active pharmaceutical ingredients (APIs) and high explosives (HEs) reaching trace levels. This methodology can be used to measure concentrations of surface contaminants for validation of cleanliness of pharmaceutical and biotechnology processing batch reactors and other manufacturing vessels. Another application discussed concerns the enhanced detection of microorganisms that can be encountered in pharmaceutical and biotechnology plants as contaminants and that could also be used as weapons of mass destruction in biological warfare. In the last application discussed, the concentration of APIs in formulations was determined by MIR laser spectroscopy and was cross validated with high-performance liquid chromatography.
Keywords: mid-infrared (MIR) laser spectroscopy, grazing angle probe (GAP), reflectance-absorption infrared spectroscopy (RAIRS), active pharmaceutical ingredients (API), pharmaceutical cleaning validation, microorganisms

1. QCL-grazing angle reflectance-absorbance IR spectroscopy

A variety of optical sensing methods can be used for the detection of chemical contaminant residues on surfaces. These methods include QCL spectroscopy, Raman, FTIR, remote infrared spectroscopy (RIRS), and laser-induced thermal excitation (LITE) of infrared emission, among others [1–17]. Fast trace detection of chemical and biological threat agents on contaminated surfaces with high selectivity and specificity is fundamental in the prevention of terrorist attacks and rapid performance and execution of security protocols. Ideally, analyte sensing on surfaces would be a rapid, in situ, low-cost, portable, highly sensitive, and able to discriminate between components. A new setup with multi-reflection passes by a grazing angle probe (GAP) using a QCL source was employed to improve in situ detection of organic contaminants on a surface. This new prototype reduced the time analysis and improved the spectral S/N.

MIR spectroscopy operating at the grazing angle of incidence (~80–82° from the surface normal) is the most sensitive optical absorption technique available for measuring low chemical concentrations on surfaces such as metals [9, 18]. Under these conditions, reflection-absorption infrared spectroscopy (RAIRS) can be performed for optically thin samples. The technique can measure low concentrations of chemical compounds deposited on substrate surfaces, such as metals, glasses, and plastics [19–23]. LODs from 10 to 50 ng/cm$^2$ of a single analyte have been obtained [24]. Because of the above, this technique allows for the analysis of monolayers on surfaces. However, the low absorbance shown by monolayers requires longer analysis times (from 5 to 120 min of integration) to obtain spectra with good signal to noise ratios (S/N). The integration time can be reduced by multiple reflection pass systems or by increasing the power of the laser source. Both alternatives improve the absorbance of the sample. However, multiple reflections involve the loss of light by phenomena such as scattering due to the substrates where the target analytes reside, sample physical properties, and mirror imperfections. The divergence of the MIR beam and absorption by the substrate, samples, and mirrors are also factors that lead to signal losses.

Therefore, the main objective was to evaluate the QCL-GAP in back-reflection mode, which is suggested as a viable method to validate detection of explosives on metallic surfaces. A QCL-GAP was designed to obtain measurements in the lab. Reflectance spectra of RDX samples deposited on aluminum (Al) plates were obtained for a remote sensing modality.

Also, GAP-IRS can be used outside the confinement of the sample compartment, making it available for fieldwork. FTIR fiber optic-coupled grazing angle probe-reflection-absorption infrared spectroscopy (FOC-GAP-RAIRS) using a thermal excitation source (globar) has been investigated before to develop techniques for detection of HE residues on substrates [23, 25, 26]. The methodology can be used in situ to detect nanograms of the target compounds. Samples with surface concentrations ($C_S$) ranging from micrograms/cm$^2$ to nanograms/cm$^2$ of explosives (DNT, TNT, PETN, nitroglycerine (NG) and triacetone triperoxide (TATP)) have been studied on SS plates with excellent results yielding 10–100× LODs for HEs than for active pharmaceutical ingredients (APIs), for which the setup was originally developed [25, 26]. The main objective of this study was to design, develop and test a grazing angle probe (GAP) mount for coupling to MIR QCL spectrometer.
(QCL-GAP) as a viable tool to develop methodologies for detection of chemicals and microorganisms on surfaces at trace level quantities within the framework of homeland security applications. The QCL-GAP was designed to obtain measurements in the laboratory and the field. Back-reflection spectra of RDX samples deposited on SS plates outside the sample compartment were remote.

1.1 Optical system

A QCL source and a compact mount with mirrors fixed near the grazing angle (~82° from surface normal) were carefully coupled to improve detection, increase the S/N and reduce the time of analysis without saturating the MCT detector. A general view of the complete optical configuration of this novel system is shown in Figure 1. The beam was focused and expanded by a lens (ZnSe, 3 in. diameter) in the vertical direction. Next, the light was reflected by a mirror at 49° of the surface, deflecting the light at an angle of 8° with respect to the surface or 82° with respect to the surface normal, forming an elliptical beam spot on the surface. The axial size of the ellipse was \( d \times \left[ \frac{d}{\sin (8°)} \right] \). For example, for \( d = 4 \) mm, the axis of the ellipse was \( 4 \times \left[ \frac{4}{\sin (8°)} \right] = 4 \times 29 \) mm². The light was returned by a plane mirror (~82°) to the same surface producing a slightly larger image at the same position.

The QCL spectrometer has inherent limitations related to the instrument design in which the MIR detector is located within the spectrometer so that the system operates only collecting the back-reflected light. The results on the detection of explosives residues on SS plates with the setup illustrated in Figure 2 are highly promising: 73 pg./cm² for RDX. This value is ~102–103 times lower than currently reported LOD values for these explosives. In the next-generation design, the MIR...
detector will be placed at the plane of the second gold-coated mirror and placed at ~82° from the surface normal. The extrapolated value of the LOD can be obtained by plotting the S/N versus the reciprocal of the surface concentration (CS) to the value of ≈ 3 [26].

1.2 Removal of QCL-GAP interference fringes of RAIRS spectra

The proposition of using a QCL-GAP setup was evaluated for monolayer analysis of HE residues. PCA and PLS multivariate routines of chemometrics were employed to verify the effect of preprocessing options for the MIR RAIRS spectra obtained. When the optical system was well aligned, interference patterns were obtained that initially were considered problems that masked the spectroscopic information. Moreover, the observed patterns were modified when the surface had an analyte deposited on it. A fast Fourier transform (FFT) analysis was applied to determine if the interferences could be handled by FFT preprocessing for discrimination and quantification analysis using multivariate analysis (MVA). FFT is an algorithm that transforms a function from the time domain to the frequency domain and conversely. The resulting transformation is a complex function. In complex notation, the domain contains one signal made up of N complex points. Each of these complex points is composed of two parts, the real part and the imaginary part. In this case, FFT was carried out to find the frequencies of the interference patterns. This transformation was used for discrimination using PLS-DA and for quantification using PLS. Both PLS and PLS-DA are supervised methods in which the data are reduced to linear combinations containing the information to generate discrimination for PLS-DA or quantification for PLS.

Figure 3a shows the spectra of substrate clean (SS); the SS with the analyte loadings, ~16 ng/cm² RDX on SS and ~20 ng/cm² irbesartan (IRBS) on SS; and an active pharmaceutical ingredient (API) angiotensin II receptor and active component of AVAPRO® used as a potential chemical interferant. The figure also includes the reference spectra for RDX and IRBS acquired with a conventional diffuse-reflection system for bulk samples (90° with respect to the surface normal). Figure 3b shows a schematic diagram on how the interference patterns can be formed on the spectra of the substrate clean and with the analyte loading. These patterns are due to the interference by multiple reflections in the system. The
variation of interference patterns depending on whether the substrate is clean or loaded with the analyte can also be noticed. Figure 3c illustrates how the FFT is applied to the RAIRS spectra. In the RDX spectra, some signals were observed with difficulties such as NO$_2$ symmetric stretch at 1275 cm$^{-1}$ and N-N symmetric stretch at 1352 cm$^{-1}$ [27, 28]. Multiple reflections are generated by semitransparent analytes when low concentrations are deposited. This interaction of the light with the surface causes a modification of the interference patterns depending on the analyte and the concentration deposited. Transformation of the data with FFT preprocessing produces a complex function consisting of an imaginary part (Im), a real part (Re), and the magnitude of the function expressed as the absolute value of z(n). These parameters were used for MVA to build robust models.

Figure 4a–c show the spectra transformation for substrate SS clean (none) and SS contaminant with RDX IRBS. The FFT shows each frequency or modes that generated the interferences in the spectra. A detailed analysis of Figure 4a–c demonstrates that modes for RDX and IRBS are similar to clean SS and differ in small modifications of these modes. These small modifications are due to the nature of the layer, shown as absorbance, homogeneity, refractive index, particle size, and layer thickness. The Re function has a mode with a higher intensity than the other modes. This mode should be the principal interference which is generated between the lens and the mirror in back reflection. Each transformation and other preprocessing were used for a principal component analysis (PCA) to verify the differences between the analytes and the clean surface. A complete separation of spectra without analytes, with RDX and IRBS, was archived with the real part Re of FFT. The loadings for the analysis are shown in Figure 4d. Two components were necessary for a complete separation with Re FFT preprocessing.
In comparison to other parameters of FFT and other preprocessing algorithms, the separation was not complete. The preprocessing used were SNV, first derivative (FD), second derivative (SD), extended scatter correction (EMSC), multiplicative signal correction (MSC), Im, and \(|z(n)|\). Figure 5a–d show the correlation between the scores of PC1 and PC2 for the SNV, \(|z(n)|\), Im, and Re. A visualization of the separation between the classes using Re is clear and complete; the maximum separation was achieved with PC2.

PLS-DA was employed using Re(FFT) as a preprocessing routine. The complete analysis was done to measure the discriminant capacity. The number of points in FFT was changed to select the better resolution for the analysis. The sensitivity and specificity for leave-one-out cross validation (LOOCV) were also calculated for a different number of points for FFT. The PLS-DA model performance was evaluated through parameters of the confusion matrix such as sensitivity and specificity of the validation. The validation was initially evaluated regarding LOOCV. The sensitivity can be defined as the samples predicted as belonging to a class divided by the total samples in that class, and the specificity is the samples predicted as not belonging to the class divided by the total samples not belonging to that class. The sensitivity and specificity were calculated according to Eqs. 1 and 2:

\[
\text{Sensitivity} = \frac{TP}{TP + FN} \quad (1)
\]

\[
\text{Specificity} = \frac{TN}{TN + FP} \quad (2)
\]

Here TP, FN, TN, and FP represent the number of true positives, false negatives, true negatives, and false positives, respectively. The best models were generated using 75 and 100 number of points for the FFT preprocessing steps. These models
correspond to very high sensitivity and specificity values. Parameters for CV for all models are shown in Table 1 in which only two LVs were used. A model for quantification for RDX was generated [29–31].

### Table 1.

<table>
<thead>
<tr>
<th># points</th>
<th>SS</th>
<th>RDX/SS</th>
<th>IRBS/SS</th>
<th>SS</th>
<th>RDX/SS</th>
<th>IRBS/SS</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0.900</td>
<td>0.893</td>
<td>0.708</td>
<td>0.885</td>
<td>0.765</td>
<td>0.737</td>
</tr>
<tr>
<td>20</td>
<td>1.000</td>
<td>1.000</td>
<td>0.958</td>
<td>1.000</td>
<td>1.000</td>
<td>0.974</td>
</tr>
<tr>
<td>50</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.947</td>
</tr>
<tr>
<td>75</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>100</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>200</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.962</td>
<td>1.000</td>
</tr>
<tr>
<td>500</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
<td>0.904</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>750</td>
<td>0.900</td>
<td>1.000</td>
<td>0.833</td>
<td>0.885</td>
<td>0.912</td>
<td>0.912</td>
</tr>
<tr>
<td>1000</td>
<td>1.000</td>
<td>0.893</td>
<td>0.792</td>
<td>0.885</td>
<td>1.000</td>
<td>0.737</td>
</tr>
</tbody>
</table>

Leave one out cross-validation (LOOCV) sensitivity/specificity for two latent variables. The best models were generated using 75 and 100 number of points for the FFT preprocessing.

### 1.3 Conclusions

A compact GAP designed to be interfaced to a QCL-based spectrometer has been described. The unit enables RAIRS measurements in the MIR under conditions of a polarized, coherent, collimated, and high-brightness laser source. The new hyphenated technique has been used in analysis surface contaminants in two broad
area applications: pharma/biotech reactor cleaning validation and HE detection for defense and security applications. Interference back-reflection patterns were observed that initially hindered the successful application of the technique. A preprocessing algorithm based on FFT was implemented in MATLAB and successfully tested. Three derived functions were used: the absolute value of the complex function of the FFT ($|z(n)|$), the imaginary part of the FFT complex function ($\text{Im}(n)$), and the real part of the complex function of the FFT ($\text{Re}(n)$).

Optimization of preprocessing was obtained upon evaluation of preprocessing models for quantitative and qualitative analysis. PLS quantification models and PCA qualitative models improved by using $\text{Re}(n)$, allowing complete separation of three classes: clean substrates (SS), HE/substrates (RDX/SS), and active API/substrates (IRBS/SS). The values for the sensitivity and specificity were 1.000 for both RDX and IRBS. These results were attained using 75/100 pts. FFT preprocessing. The QCL-GAP back-reflection setup described herein can provide the basis for developing methodologies for high specificity and sensitivity results for monolayer analysis using RAIRS. These results will have a far-reaching impact on cleaning validation, defense/security, and other applications involving monolayer analysis.

2. MIR laser detection and discrimination of microorganisms

Driven by an imperative need to develop quick and precise methods for detection of biological warfare agents (BWAs), MIR laser spectroscopy study of selected microorganisms was undertaken. *Escherichia coli* (*Ec*) can be detected using electrochemical immunosensors, immobilized probes and solid-phase microextraction followed by GC–MS, which can also be used to detect other microorganisms. All these methods and others currently available are laborious and costly and comprise many preparation steps or selective pre-enrichment [32–39]. Though identification/discrimination of bacterial spores with FTIR has been reported, this contribution proposes the application of MIR laser spectroscopy for identification and discrimination of bacteria residing on reflective and matte surfaces [34, 35]. *Bacillus thuringiensis* (*Bt*) is a Gram-positive bacterium that forms spores that are highly chemoresistant and also tolerate high temperatures in their dormant state. *Bt* was selected as a simulant for BWAs based on its resemblances with *Bacillus anthracis* (anthrax), a well-known BWA. *Staphylococcus epidermidis* (*Se*) is a Gram-positive coccus that can be found usually on human skin. *Ec* is a Gram-negative bacteria affiliated to the family *Enterobacteriaceae*. *Ec* is a coliform that can be found in intestines of warm-blooded animals. Thus, the presence of *Ec* is associated with fecal contamination.

Several materials, including SS, CB, TB, glass, and W, were used as substrates for depositing the samples. Since the spectroscopic information of bacteria mainly consists of signature contributions from all the cell components, the reflectance spectra show the molecular compositions of the cells in general. Other IRS studies have focused the problem of detecting, identifying, and discriminating bacteria from the substrates they reside on using chemometric methods [36–40]. The methodology used in this work involves obtaining the MIR laser-enhanced reflectance spectra under high-brightness conditions. Certified bacterial strains of *Bt* (ATCC #35646), *Ec* (ATCC #8789), and *Se* (ATCC #2228) were acquired from the Microbial Biotechnology and Bioprospecting Lab at the Department of Biology at the University of Puerto Rico-Mayagüez campus. The microorganisms were selected based on their resemblance to real-world BWA simulants.

QCL spectra of microorganisms were used to identify the molecular vibrational markers in the biosamples. These vibrational signatures contain data on the
biochemical composition of microorganisms and of the molecules of which they composed [38]. Some of the cell wall components are different for Gram-positive and Gram-negative bacteria. On the one hand, Gram-positive bacteria have a denser and stiffer peptidoglycan coating that amounts from 40 to 80% more of the cell wall (by weight) than in Gram-negative bacteria. Also, Gram-positive bacteria contain teichoic acids that are covalently attached to peptidoglycan. Gram-negative cells do not contain teichoic acids. In contrast, they contain lipoproteins that are covalently attached to the peptidoglycan in the cell walls. Gram-negative bacteria have an external membrane outside the peptidoglycan layer that contains phospholipids in the interior and lipopolysaccharides in the exterior [23]. Each bacterial species has a unique MIR fingerprint spectrum due to the stretching and bending vibrations of its molecular bonds or protein functional groups (including nucleic acids, lipids, sugars, and lipopolysaccharides), as illustrated by the reference spectra presented in Figure 6 [40, 41]. MIR FTIR reference spectra of Bt, Ec, and Se are illustrated in Figure 6. Reference spectra were obtained in IR absorption using the Bruker Optics bench microspectrometer IFS66/v/S.

Representative QCL spectra of the microorganisms deposited on SS are shown in Figure 7. A total of 245 experiments are reported out of the 836 carried out. An experiment consisted of 15 replicate spectroscopic acquisitions for each bacterium/substrate arrangement. The spectral signatures were observed on the SS coupons, particularly in the fingerprint region, because of the highly MIR reflectivity of these surfaces. Tentative band assignments were based on a comparison of reported values. However, it was difficult to distinguish the different classes of species/surface arrangements studied based on the raw MIR spectral data due to the high degree of band overlapping. Thus, MVA routines were useful in handling the large dataset generated and facilitating the spectroscopic analysis. Vector normalization (VN) was used for data preprocessing before statistical analyses. Similar effects were found for the matte (nonreflective) substrates, although the classification and discrimination required more robust MVA routines and pretreatments. Table 2 shows the classification obtained between groups of bacteria on different surfaces using PCA. Bold values represent percentages of the discrimination predicted within a correctly classified group. Selecting 10–15 QCL reflectance spectra of samples from each bacterium on each of the substrate (225 spectra) led to generate other PCA models.

Figure 6.
Reference MIR reflectance spectra of Bt, Ec, and Se in the spectral region studied.
Spectra were pretreated by applying first dvt and MC algorithms to all bacterium/substrate combination spectra. SNV pretreatment had to be applied to the data involving the use of W substrates. A PCA model for MIR laser spectra of Bt, Ec, and Se deposited on the surfaces studied was generated, and the variance captured by the PCs was analyzed for each surface type. Score plot (PC-3 vs. PC-1) for TB is shown in Figure 8a. A relatively poor separation between the datasets of the bacteria was observed. PC-1 (53% variance) versus PC-2 (14% variance) was correlated with the differences among the microorganisms. Figure 8b illustrates the grouping of spectra according to PC-3 (10% variance) versus PC-2 (14% variance) was correlated with the differences among the microorganisms. Figure 8b illustrates the grouping of spectra according to PC-3 (10% variance) versus PC-2 (14% variance) was correlated with the differences among the microorganisms. Figure 8b illustrates the grouping of spectra according to PC-3 (10% variance) versus PC-2 (14% variance) was correlated with the differences among the microorganisms. Figure 8b illustrates the grouping of spectra according to PC-3 (10% variance) versus PC-2 (14% variance) was correlated with the differences among the microorganisms. Figure 8b illustrates the grouping of spectra according to PC-3 (10% variance) versus PC-2 (14% variance) was correlated with the differences among the microorganisms. Figure 8b illustrates the grouping of spectra according to PC-3 (10% variance) versus PC-2 (14% variance) was correlated with the differences among the microorganisms. Figure 8b illustrates the grouping of spectra according to PC-3 (10% variance) versus PC-2 (14% variance) was correlated with the differences among the microorganisms. Figure 8b illustrates the grouping of spectra according to PC-3 (10% variance) versus PC-2 (14% variance) was correlated with the differences among the microorganisms.

### Table 2.
Classification between groups of bacteria (Bt), (Ec), and (Se) on various substrates.

<table>
<thead>
<tr>
<th>Discrimination</th>
<th>Group size</th>
<th>Predicted discrimination</th>
<th>Substrates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bt</td>
<td>75</td>
<td>45 (60%)</td>
<td>18 (24%)</td>
</tr>
<tr>
<td>Ec</td>
<td>77</td>
<td>3 (4%)</td>
<td>64 (83%)</td>
</tr>
<tr>
<td>Se</td>
<td>75</td>
<td>14 (19%)</td>
<td>3 (4%)</td>
</tr>
</tbody>
</table>

Substrates 18 3 (17%) 4 (22%) 0 (0%) 11 (61%)

Percentage of cases correctly classified: 69.4%

PLS-DA was used as a classification methodology for differentiating between the bacterial species on the five matte substrates studied. In PLS-DA, the estimated experimental percentage of the correctly classified samples determines the sensitivity of the model. Moreover, the estimated experimental percentage of the samples that are rejected by the other classes in the model gives information on the specificity of the model. Therefore, in a perfect class model, the sensitivity and specificity have values of 1 or 100%. A total of 225 spectra corresponding to 15 spectra for of
each bacterium/surface arrangement were analyzed. Spectroscopic data were organized into two sets. About 75% of the reflectance spectra were randomly selected as the training set for the calibration and cross validation. The remaining 25% of the data comprised an external test set. The spectral windows used for the chemometric runs were 848–1012, 1022–1170, and 1173–1400 cm\(^{-1}\). Then data were preprocessed by smoothing before applying the first dvt.

A cross validation procedure using Venetian blinds with 10 splits was carried out. A classification model using this procedure was applied to 90% of the data. Then, the other 10% of the data was separated for the validation dataset to determine the accuracy of the models. The QCL spectra were pretreated by smoothing and taking the first derivative to improve the visualization of the spectra. The discrimination models for the bacteria/substrates are shown in Figure 9 for bacterial species deposited on TB. These illustrate the predicted (PRED) cross validation (CV) of classes for each sample (PLS-DA plot). The results obtained demonstrate that the use of QCL spectroscopy (840–1440 cm\(^{-1}\)) coupled to MVA-PCA and PLS-DA are suitable for discriminating between microorganisms (Bt, Ec, and Se) on several surfaces, including on reflective, matte substrates.

MIR laser spectroscopy was very effective for detecting microorganisms on various surfaces. When coupled to MVA, the combined methodology provided a quick response and efficient discrimination from the matte substrates. The methodology could be used to identify biofilms deposited on substrates, providing quick and precise analyses for national defense and security applications and for quality control purposes in industrial scenarios, when nondestructive analytical methods are preferred. Identification and discrimination of microorganisms from the acquired MIR laser reflectance spectra were attained with PCA and PLS-DA. In general, PLS-DA performed significantly better than PCA in the analyses of the bacteria studied [40, 41].

Figure 8.
PC-2 vs. PC-1 for the PCA models of Bt, Ec and Se spectra on: (a) CB; (b) glass for (a); (c) SS; (d) W.
3. MIR laser spectroscopy analysis of pharma formulations

The Beer–Lambert–Bouguer law is a relationship between the attenuation of radiant flux, the path length of the light traversing the media (l; m), and the molar concentration of the absorbing species (c; mol $\cdot$ dm$^{-3}$). At a fixed frequency or in small frequency interval, this relationship is characterized by the absorptivity or molar attenuation coefficient ($\varepsilon$; m$^2$ · mol$^{-1}$) of the material or components in a mixture. In the near-infrared (NIR) region, intrinsically small values of $\varepsilon$ make spectrometric measurements ideal for monitoring online industrial procedures because light travels a long way in the media or can be reflected off the sample [42]. Corresponding $\varepsilon$ values in the MIR region are much higher [43]. Thus, the main difficulties of using the MIR region for online industrial applications are the thickness of the samples, the concentrations of the samples, or both. For non-thin or dilute samples, radiation in the MIR is almost totally attenuated by the samples, and the light back-reflected by the detector is very little. Thus, diffuse reflectance measurements in the MIR lead to low intensities and accordingly to low S/N values. This makes the MIR inappropriate for online applications of monitoring industrial processes.

The main objective of this work was to investigate if MIR laser spectroscopy could be used to develop a remotely sensed method for quantitative analysis of APIs in pharmaceutical blends. A QCL spectrometer was used to collect the reflectance from powder samples. Provided that the API has a high $\varepsilon$ value when a MIR laser is used to excite the spectra, much more light would be reflected resulting in high S/N values. This study concluded that accuracies and precisions obtained using MIR laser spectroscopy are comparable to NIR spectroscopy [44, 45], Raman spectroscopy [46, 47], and attenuated total reflection (ATR) infrared spectroscopy [48]. Other process analytical technology (PAT) applications of MIR laser spectroscopy
are cleaning validation of pharmaceutical and biotechnology industrial batch reactors and other processing equipment [19], determination of API levels in tablets and formulations [49, 50], and vibrational circular dichroism [51].

3.1 Experimental details

Ibuprofen (IBP) and excipients, lactose monohydrate (54.25–74.25%), microcrystalline cellulose (25%), colloidal silicon dioxide (0.25%), and magnesium stearate (0.5%), were mixed concentrations from 0 to 21% (w/w); 21 compositions were prepared together with various samples of the control (0% API). Sample mixing was conducted by using a shaker/mixer. After initial mixing, the samples were ground in a mortar and pestle and remixed. QCL reflectance spectra of 21 powder mixtures and control were acquired on various locations on the sample surfaces. The parameters of the MIR laser spectrometer were 1.5 s scan time, the average power was 0.5–10 mW, and the spectral range was 600 cm$^{-1}$. Reference spectra for all chemicals were acquired in a model IFS 66v/S bench interferometer (FTIR, Bruker Optics, Billerica, MA, USA). This system had a cryo-cooled MCT detector and a KBr beam splitter. Reference spectra were acquired in transmission mode at a resolution of 4 cm$^{-1}$ using 32 scans at 10 kHz scan velocity.

Cross validation experiments were carried out using the industry standard method based on high-performance liquid chromatography (HPLC) on a model 1100 Agilent Technologies system (Santa Clara, CA, USA). The system was equipped with a diode array UV–VIS detector. A C18 HPLC column was used for the chromatographic experiments. Ultrapure water adjusted to pH of 2.5 with HPO$_3$ and acetonitrile (40/60, v/v) was used as mobile phase at a flow rate of 1.0 mL/min. Analyte detection was carried out in the UV at an excitation wavelength of 214 nm [52].

3.2 Results

The measured reflectance spectra were converted to the Kubelka-Munk (K-M) function. The main criterion to use the K-M transformation was that the intensity values obtained for the samples were very low [53]. For the chemometric analyses, a preprocessing step consisting of VN was used in the full spectral range to remove baseline shifts produced by scattering of the MIR light. These shifts are typically caused by variations in particle sizes of the crystalline components of the mixtures. VN involves calculating the average intensity first. Then, the average value of the intensities was subtracted from each spectrum. Subsequently, the sum of the intensities squared was calculated, and each spectrum was divided by the square root of this value. VN operates on each spectrum. This preprocessing step worked better than other pretreatment steps applied: MC, constant offset elimination (OFFSET), straight-line subtraction (SLS), minimum-maximum normalization (MIN-MAX), multiplicative scattering correction (MSC), and first and second dvt.

The control spectra were similar to the spectrum of the formulation containing 20% API. The differences observed were based on the fact that the excipients used to prepare the formulations contained many vibrational bands in the spectral region studied. A spectrum consisting of the difference between the 20% IBP formulation and the control (DIFF) was obtained and used to identify the vibrational markers of IBP in the formulations (Figure 10a). The IBP reflectance spectrum acquired by MIR laser (IBP-QCL) was compared to the corresponding FTIR spectrum (IBP-FTIR) to establish the exactness of the method. Vibrational signals corresponding to IBP were identified, establishing a good agreement between reflectance spectra obtained by IBP-QCL and IBP-FTIR. A slight shift of +4 cm$^{-1}$ for IBP-QCL relative to DIFF was observed for the band at 1303 cm$^{-1}$. 
Low-intensity signals were noticed in the DIFF spectrum at 1142, 1255, 1293, and 1540 cm\(^{-1}\). Thus, though API-matrix interactions were weak and did not affect the shape or position of the spectral markers of the API for the bands observed, the technique is sensitive enough to detect these weak interactions. Because the controls did not contain API and the excipients amounts were kept at the same levels as the samples with API, these were about 25% higher in concentration in the controls. The spectral profile of the various constituents in the control mixture at this substantial concentration change cannot be entirely compensated when applying a non-weighted spectral subtraction.

Values for the statistical parameters of the PLS model are included in Table 3. The root-mean-square error of estimation (RMSEE) and relative standard error (RSE) were used to estimate the exactness of the PLS-based models [54]. Spectra were acquired at various locations on the surface of the sample, 20 spectra/sample in total. The calibration and internal validation set consisted of 16 of these spectra. The external validation set, used for testing the model, consisted of the remaining four spectra. The resulting forecasted error was labeled RMSEP1. This procedure was repeated for each composition. The robustness of the model was verified by preparing five formulations with compositions different from those of the calibration set and predicting their concentration values. Another prediction error was calculated by the average of the differences between the nominal (gravimetric) value and the value predicted by the PLS model. This error was designated as RMSEP2.

The optimum model was generated by narrowing the spectral range. This was done by first using the complete spectroscopic window, then dividing the full range into equal spectral subregions, and so forth. The optimum arrangement of subregions was determined by starting with 10 subregions and sequentially excluding 1 subregion, each time determining the value of the RMSECV. The process continued until the values for RMSECV did not reduce any further. The region from 990 to 1295 cm\(^{-1}\) was selected as the best for the PLS modeling. The values achieved for RMSEE and RMSECV were comparable and small compared to the composition range of the experiments (0–21%) as per the values of RSE and RSCV of 2.3 and 3.1%, respectively, indicating the percentage of error in the model and...
the prediction capability, respectively. Challenging the model with new samples resulted in an RSE value increased to 6.5%, including the sample preparation errors. The bias, which is the average of the predicted values by cross validation, provides information on systematic errors. Since this value was small (Table 3), no significant deviations were attributed to the preparation of samples.

The very good linearity of the calibration model as can be evidenced by the value of the correlation coefficient squared ($R^2$). This means that the percentage of variance of the reference gravimetric values that are reproduced in the PLS regression is high (Table 3). This is also evident in Figure 10b: plot of prediction values (%) versus reference values (%) for the samples in calibrations, cross validations, and tests. Black dots represent fitted dataset for the calibration. Blue triangles symbolize cross validation dataset. Red squares represent the first test set; the second test set is shown in green circles. The ideal model ($y = x$), in which all the predicted values are equal to the reference values through the whole data interval, is represented by a dotted line. Figure 10b shows low dispersion of the data about the ideal model ($y = x$) for predicted values of calibrations, cross validations, and tests.

Finally, the most debatable figure of merit in the PLS analysis is the LOD [54–56]. The RSD was calculated for each mix from the predicted concentrations by cross validation and testing. A plot of the precision in terms of the RSD values versus nominal reference values was obtained. According to IUPAC recommendations [57–62], a power fit was applied, and the values for the LOD and the limit of quantification (LOQ) were obtained by the interpolation of the concentration for 33.3% of RSD and 10% of RSD, respectively [63]. LOD found was 1% of API, and the RSD was maintained below 5% for concentrations >5% (see inset of Figure 10b). This result suggests a good value for the LOD and a low uncertainty for the analytical quantification. These values are comparable to methodologies based on NIR and Raman spectroscopy measurements.

An evaluation between the proposed MIR laser spectroscopy methodology and the industry standard HPLC method was done. For the comparison study, a sample set of three formulations in the range of 0–20% was prepared, and the sample compositions were verified by HPLC and contrasted with the values obtained by the MIR laser spectroscopy methodology proposed (Table 4). A relative standard error of prediction for HPLC (RSP-HPLC) was obtained using the gravimetric values. The value obtained for RSP-HPLC was lower than other RSP values, but as can be

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td>0.998</td>
<td>LOD%</td>
<td>1</td>
</tr>
<tr>
<td>$R^2$-CV</td>
<td>0.996</td>
<td>LOQ%</td>
<td>3</td>
</tr>
<tr>
<td>RMSEE (%)</td>
<td>0.3</td>
<td>RPD-Fit</td>
<td>21.3</td>
</tr>
<tr>
<td>RMSECV (%)</td>
<td>0.41</td>
<td>RPD-CV</td>
<td>15.5</td>
</tr>
<tr>
<td>RMSEP (%)</td>
<td>0.7</td>
<td>RSD-R (%)</td>
<td>0.5</td>
</tr>
<tr>
<td>RMSEP1 (%)</td>
<td>0.54</td>
<td>RSD-H (%)</td>
<td>2.7</td>
</tr>
<tr>
<td>RMSEP2 (%)</td>
<td>0.82</td>
<td>RSD-REP (%)</td>
<td>5.4</td>
</tr>
<tr>
<td>RSE</td>
<td>2.3%</td>
<td>SEN</td>
<td>0.0063</td>
</tr>
<tr>
<td>RSP2</td>
<td>7.0%</td>
<td>LV</td>
<td>4</td>
</tr>
<tr>
<td>RSP-HPLC</td>
<td>1.3%</td>
<td>Bias</td>
<td>−0.002</td>
</tr>
</tbody>
</table>

*Error calculation used data of RMSEP1 and RMSEP2

Table 3. Values of relevant parameters for calculating the figures of merit of the PLS models.
observed in Table 4, the predicted concentration values (% composition, API) were very close to the gravimetric values and the HPLC predicted values. Thus, this cross validation experiment points to the robustness of the MIR laser methodology proposed.

3.3 Conclusions

QCL spectroscopy was used to acquire spectra of API in lab-prepared pharmaceutical formulations. MIR laser spectra of formulations were acquired outside the sample compartment, at 15 cm, in diffuse reflectance in the range of 1000–1600 cm$^{-1}$. Because of the convoluted MIR spectra of the formulations, the quantification had to be handled using PLS. The MVA method was shown to be capable of achieving good exactness and precision at forecasting the API compositions in the lab-made formulations. Significant effectiveness obtained for the model is indicative of a high analytical sensitivity equivalence (0.05% API), good repeatability (2.7%), and good reproducibility (5.4%). All these qualities allowed achieving a LOD of 1%. Furthermore, the proposed procedure is characterized by high specificity, high sensitivity, fast response, and sufficiently high accuracy and precision.

The proposed protocol is a demonstration that MIR laser spectroscopy can be used for off-line monitoring of APIs in pharmaceutical formulations. Further work in this area could lead to the next phase, where MIR laser spectroscopy combined with MVA routines of chemometrics possibly will be used to challenge results obtained by online supervision of manufacturing pharmaceutical and biotechnology processes. This would provide real-time data to control systems in continuous manufacturing practices (CMP) in agreement with modern PAT tendencies.

4. Summary

MIR laser spectroscopy has been demonstrated as a highly adaptable spectroscopic method for recasting traditional MIR spectroscopic techniques, such as absorption, reflectance, transmission, emission, and RAIRS, under the high-brightness conditions that a collimated, polarized, coherent laser source provides. When MIR laser-excited reflectance spectra are coupled to chemometric algorithms for classification, discrimination, and quantification, much lower LOD and LOQ values were obtained for the target chemical/biological agent simulants. A selection of several applications of MIR laser spectroscopy has been presented. These applications cover from the coupling a MIR laser to a compact grazing angle probe for trace detection of chemical and biological threat agents, experiments for detection of microorganisms, and PAT applications such as cleaning validation of batch reactors in pharmaceutical and biotechnology manufacturing plants and quantification of APIs in pharmaceutical formulations.

<table>
<thead>
<tr>
<th>% weight</th>
<th>% HPLC</th>
<th>% PLS-QCL</th>
</tr>
</thead>
<tbody>
<tr>
<td>14.4 ± 0.1</td>
<td>14.63 ± 0.03</td>
<td>14.4 ± 0.4</td>
</tr>
<tr>
<td>9.9 ± 0.1</td>
<td>10.07 ± 0.02</td>
<td>10.1 ± 0.3</td>
</tr>
<tr>
<td>4.48 ± 0.09</td>
<td>4.58 ± 0.03</td>
<td>4.5 ± 0.4</td>
</tr>
</tbody>
</table>

Table 4. Prediction of IBP Composition (w/w%) by HPLC and QCL.
Acknowledgements

This material is based upon work supported by the US Department of Homeland Security, Science and Technology Directorate, Office of University Programs, under Grant Award 2013-ST-061-ED0001. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the US Department of Homeland Security. The parts of the work were supported by the US Department of Defense, Agreement Number: W911NF-11-1-0152.

Author details


1 Department of Chemistry, University of Puerto Rico-Mayagüez, ALERT DHS Center of Excellence, Mayagüez, Puerto Rico, USA

2 School of Basic and Biomedical Sciences, University of Simón Bolívar, Barranquilla, Colombia

3 Department of Chemistry, University of Córdoba, Water Resources, Pesticides, and Heavy Metals Research Group, “GIAMP”, Montería, Colombia

4 Antonio de Arévalo Technological Foundation, TECNAR, Cartagena, Colombia

5 Department of Physics-Chemistry, University of Puerto Rico, Arecibo, Puerto Rico, USA

6 Department of Biology, University of Puerto Rico, Mayagüez, Puerto Rico, USA

*Address all correspondence to: samuel.hernandez3@upr.edu
References


[29] Barker M, Rayens W. Partial least squares for discrimination. Journal of
Infrared Spectroscopy - Principles, Advances, and Applications

Chemometrics. 2003;17(3):166-173. DOI: 10.1002/cem.785


[34] Siripatrawan U, Harte BR. Solid phase microextraction/gas chromatography/mass spectrometry integrated with chemometrics for detection of Salmonella typhimurium contamination in a packaged fresh vegetable. Analytica Chimica Acta. 2007;581:63-70. DOI: 10.1016/j.aca.2006.08.007


[43] Févotte G. In situ Raman spectroscopy for in-line control of
pharmaceutical crystallization and solids elaboration processes: A review. Chemical Engineering Research and Design. 2007;85(7):906-920


[48] Chan KLA, Hammond SV, Kazarian SG. Applications of attenuated total reflection infrared spectroscopic imaging to pharmaceutical formulations. Analytical Chemistry. 2003;75(9):2140-2146


[58] Boqué R, Faber NM, Rius FX. Detection limits in classical
multivariate calibration models.
Analytica Chimica Acta.
2000;423(1):41-49

[59] Boqué R, Rius FX. Multivariate
detection limits estimators.
Chemometrics and Intelligent
Laboratory Systems. 1996;32(1):11-23

[60] Faber K, Kowalski BR. Improved
estimation of the limit of detection
in multivariate calibration. Fresenius
Journal of Analytical Chemistry.
1997;357(7):789-795

[61] Bauer G, Wegscheider W,
Ortner H. Limits of detection in
multivariate calibration. Fresenius
Journal of Analytical Chemistry.
1991;340(3):135-139

[62] Thomsen V, Schatzlein D,
Mercuro D. Limits of detection
in spectroscopy. Spectroscopy.
2003;18(12):112-114

[63] Galán-Freyle NJ, Pacheco-
Londoño LC, Román-Ospino AD,
Hernández-Rivera SP. Applications of
quantum cascade laser spectroscopy
in the analysis of pharmaceutical
2017;70(9):1511-1519