

Chemical Sensing of Common Microorganisms Found in Biopharmaceutical Industries Using Mid-Infrared Laser Spectroscopy and Multivariate Analysis

Samuel P Hernandez-Rivera¹, Wilmer Carrión-Roca¹, Annette M. Colón-Mercado¹, Francheska M. Colón-González¹, John R. Castro-Suarez², Edwin R. Caballero-Agosto¹, José A. Centeno¹, and Carlos Ríos-Velázquez³

¹University of Puerto Rico

²Universidad del Sinu - Seccional Cartagena

³Universidad de Puerto Rico Recinto Universitario de Mayaguez Departamento de Biología

September 22, 2023

Abstract

We report on the spectroscopic investigation of common bacteria encountered in biopharmaceutical industries with spectroscopic definition and specificity using mid-infrared laser spectroscopy. This study describes the detection of three different bacteria species using quantum cascade laser spectroscopy coupled to a grazing angle probe (QCL-GAP). Stainless steel substrates were used as support for the bacterial samples. QCL-GAP spectroscopy was assisted by multivariate analysis (MVA) to assemble a powerful spectroscopic technique with classification, identification, and quantification resources. The bacterial species analyzed, *Staphylococcus aureus*, *Staphylococcus epidermidis*, and *Micrococcus luteus*, were used to challenge the technique's capability to discriminate microorganisms from the same family. Principal component analysis and partial least squares-discriminant analysis differentiated between the bacterial species using QCL-GAP. Spectral differences in the bacterial membrane were used to determine if these microorganisms were present in the samples analyzed. Results herein provided effective discrimination for the bacteria under study with high sensitivity and specificity values.

Title:

Chemical Sensing of Common Microorganisms Found in Biopharmaceutical Industries Using Mid-Infrared Laser Spectroscopy and Multivariate Analysis

Authors:

Wilmer Carrión-Roca^{1*}, Annette M. Colón-Mercado¹, Francheska M. Colón-González¹, John R. Castro-Suarez, Edwin R. Caballero-Agosto¹, José A. Centeno¹, Carlos Ríos-Velázquez², and Samuel P. Hernández-Rivera^{1*}

¹Center for Chemical Sensors (CCS)/Chemical Imaging and Surface Analysis (CISAC), University of Puerto Rico, Department of Chemistry, PO Box 9000, Mayaguez, PR, USA, 00681; wilmer.carrion@upr.edu (W.C.R.); annette.colon1@upr.edu (A.M.C.M.); edwin.caballero@upr.edu (E.R.C.A.); samuel.hernandez3@upr.edu (S.P.H.R.); jose.centeno10@upr.edu (J.A.C.); jacenteno@comcast.net (J.A.C.)

²University of Puerto Rico, Department of Biology, Mayaguez, PR, USA, 00681; carlos.rios5@upr.edu

³Universidad del Sinú, Unisinú, Cartagena 130015, Colombia; johncastro Suarez@gmail.com (J.R.C.S.)

*Correspondence: samuel.hernandez3@upr.edu (S.P.H.-R.); wilmer.carrion@upr.edu (W.C.R.) Tel.: (787)-832-4040; Ext 5404. (S.P.H.R.)

Abstract:

We report on the spectroscopic investigation of common bacteria encountered in biopharmaceutical industries with spectroscopic definition and specificity using mid-infrared laser spectroscopy. This study describes the detection of three different bacteria species using quantum cascade laser spectroscopy coupled to a grazing angle probe (QCL-GAP). Stainless steel material, like surfaces commonly used in biopharmaceutical industries, was used as support media substrates for the bacterial samples. QCL-GAP spectroscopy was assisted by multivariate analysis (MVA) to assemble a powerful spectroscopic technique with classification, identification, and quantification resources. The bacterial species analyzed, *Staphylococcus aureus*, *Staphylococcus epidermidis*, and *Micrococcus luteus*, were used to challenge the technique's capability to discriminate microorganisms from the same family. Principal component analysis and partial least squares discriminant analysis differentiated between the bacterial species, using (QCL-GAP) as the reference. Spectral differences in the bacterial membrane were used to determine if these microorganisms were present in the samples analyzed. Results herein provided effective discrimination for the bacteria under study with high sensitivity and specificity values.

Keywords: quantum cascade laser spectroscopy (QCLS), infrared spectroscopy (IRS), bacteria, stainless steel substrates (SS), principal component analysis (PCA)

1 Introduction

Microbial contamination happens because of the existence and proliferation of microorganisms in the environment [1]. The ability of microorganisms to grow in food, pharmaceutical and cosmetic products, and medical devices has been the subject of several studies [2]. From the infectious point of view, pathogenic microbes in pharmaceutical and food products make them hazardous and objectionable [2]. Microbial infections can alter active ingredients' physicochemical and biological properties or even become toxic materials [3]. Some diseases, such as diarrhea, acute gastroenteritis, and abdominal pain, result from microbial toxins. Nevertheless, depending on the individual sensitivity to the toxin, symptoms are different and range from mild distress to the individual's death [4].

Biological contamination in the drug manufacturing industry is a medical problem that can lead to drug degradation and sub-potency. Patients can be exposed to pathogens or opportunistic microorganisms that can cause serious metabolic harm or lead to death, especially if the patient is immunocompromised. It also represents a hazardous threat in the pharmaceutical and biotechnology clean-room production environment. Regulatory agencies require continuous, routine, or regular environmental monitoring of microorganisms to ensure that the environment during production activities is precisely controlled. Maintaining a controlled environment is vital to avoid contamination and protect the patient from contact with possible extraneous matter or product contamination. Environmental monitoring requirements may include but are not limited to the manufacturing room, surfaces, walls, ceiling roof HEPA filters, equipment, and personnel. Identifying and characterizing bacterial isolates starts by inspecting the colony morphology once the cells have been cultured in solid media, followed by microscopic analysis of Gram-stained preparations [5].

Today, pharmaceutical and biotechnology industries assess their activities, such as daily environmental monitoring for their facilities, bioburden, endotoxins, and sterility for product microorganism contamination. Industries, specifically the parenteral product industries, monitor the water used to manufacture the products. Biotechnology industries often use Polymerase Chain Reaction (PCR) and Micro sequencers. These techniques take long periods to implement because the operator needs to collect the sample in the area and inoculate it in a culture media to allow the microorganism's growth. Results obtained from these studies will commonly take about 12 to 24 hours, depending on the culture media used. Once the incubation period is completed, the analyst studies the morphology of the recovered microorganism under the microscope and performs the identification using the abovementioned techniques. The industry may select to ship the recovered microorganism to an external laboratory to complete the identification process.

Therefore, in these fields, fast identification of pathogenic and non-pathogenic microorganisms is necessary because the time required for identifying pathogens is essential in determining the contamination source promptly and reducing manufacturing costs. Rapid identification techniques play a critical role in reducing costs associated with contamination to mitigate the event faster and avoid further product processing that may increase product disposition expenses.

Bacteria can be divided and identified into groups based on their morphology (macro and micro), physiological, biochemical, serological, and genetic characteristics. Generally, the tests are combined in a series of solid or liquid media inoculated with bacteria and identified after a certain incubation period. Many different tests are often needed for the definitive identification of the bacteria. These tests require turnaround times from 24 hours to up to 5 days between receipt of material and identification results by the clinician [6].

Thus, empirical treatment with broad-spectrum antibiotics is often started while awaiting further identification of the bacteria. It has been reported that 10-30% of patients suffering from bloodstream infections in intensive care units (ICUs) do not receive the correct antibacterial therapy initially. Mortality rates in this group have been reported to be 30-60% higher than in the group that promptly receives appropriate therapy [7].

The effort has been invested in developing new techniques for identifying microorganisms, including molecular methods, such as mass spectrometry, electrospray ionization, matrix-assisted laser desorption ionization, Fourier Transform Infrared (FT-IR), and Raman spectroscopy. Among these methods, vibrational spectroscopy Quantum Cascade Laser Grazing Angle Probe (QCL-GAP) is a reagent less/solventless, in which there is no need to add chemical dyes or labels for identification. QCLs differ from traditional semiconductor laser diodes, which use p-n junctions for light emission, consisting of diode arrays with an active region where electrons and holes recombine to produce light emission. Instead, QCLs have multiple active regions composed of a multilayered semiconductor material structure specially designed to have the appropriate electronic bands [8].

QCL sources consist of semiconductor lasers based on sub-band transitions in a manifold of quantum-well structure [9].

QCL sources operate at wavelengths in the MIR starting from 3300 to approximately 750 cm^{-1} , which matches well with the fundamental vibrational absorption bands of many biological species, in contrast with typical diode sources where the laser emission generally matches the weaker overtone. The emission wavelengths of these lasers depend on the thickness of the quantum well and the barrier layers of the active region rather than the band gap of diode lasers. They operate near room temperature, produce milliwatts of radiation, and offer the possibility of tailoring the emission wavelengths within a broad range of frequencies [9,12].

Vibrational spectra of bacterial cells consist of signal contributions of all components in the cells and, therefore, reflect their overall molecular composition. FT-IR spectroscopy has proven that a wide range of microorganisms can be identified using the spectral response [13-15]. Gram-positive and Gram-negative sugar-based coating structures are also relevant properties that can contribute to the spectral differences. Other factors that have been used as discrimination in FTIR are the biochemical fingerprints observed on spectra consistently correlated with sugar-based coating structures that, besides reflecting strain variation, are also highly relevant for the specificity in pathogen-host interactions [16].

Thus, the present research aims to detect, identify, and discriminate between three bacteria from the same group: *Staphylococcus aureus*, *Staphylococcus epidermidis*, and *Micrococcus luteus* deposited over a stainless steel (SS) substrate-like material used in the biopharmaceutical industry clean rooms using QCL-GAP. Therefore, the bacteria concentration was not considered for this study since the investigation is associated with detection and discrimination. Studies regarding concentration are in the plan and will be included in a subsequent publication.

2 Materials and Methods

2.1 Materials and Reagents

Three environmental bacteria isolates (*Staphylococcus aureus* (*Sa*), *Staphylococcus epidermidis* (*Se*), and *Micrococcus luteus* (*Ml*)) from the manufacturing areas were inoculated from 30 - 35°C. The environmental isolates used in this study were sent to Microbiologics® Laboratory (St. Cloud, MN, 56303 USA) to certify and identify the bacteria using MALDI-TOF with supplemental tests, for example, ID determination with MALDI-TOF confirmation. The method used for concentration verification is the spread plate method, using the materials and incubation conditions required for the bacteria under characterization. After that process, the sample is lyophilized with a 50 EU/0.1mL concentration. The bacteria were reconstituted with a reconstitution buffer provided by Microbiologics company and inoculated in Tryptic Soy Broth (TSB). The bacteria were incubated at 30 - 35°C for 12 hours. A solution of 10 mL in 90 mL of TSB media was prepared. Each solution containing the microorganism was left to grow for 24 hours. The SS substrate (2-in x 2-in) used for the sample deposition of the bacteria was provided by Stainless Supply AJW Metal Product Company (Moroe, NC, 28110 USA). The SS substrate used in this study is a grade material similar to that used in clean rooms in biopharmaceutical industries.

2.2 Instrumentation

A QCL-GAP pre-dispersive spectrometer based on QCL technology (LaserTune, Block Engineering, LLC, Southborough, MA, 01772 USA) set up for reflection-absorption measurements was used for data acquisition in the MIR: 5.4 – 12.8 μm of the bacterial suspension samples (BSS). The microorganism's vibrational response was collected using a 4 cm^{-1} spectral resolution. The spectrometer was equipped with an internal thermoelectrically mercury-cadmium-telluride (MCT) detector and a 2 x 4 mm^2 MIR laser beam. The setup has two adjustable mirrors that allow the laser incidence at the grazing angle ($\sim 82^\circ$). The light is emitted towards the first gold (Au) mirror, which goes to the sample. A second gold mirror reflects the light from the sample, allowing a double-pass reflection-absorption measurement. The QCL-GAP array is mainly useful for detection at low concentration ranges with added sensitivity [18]

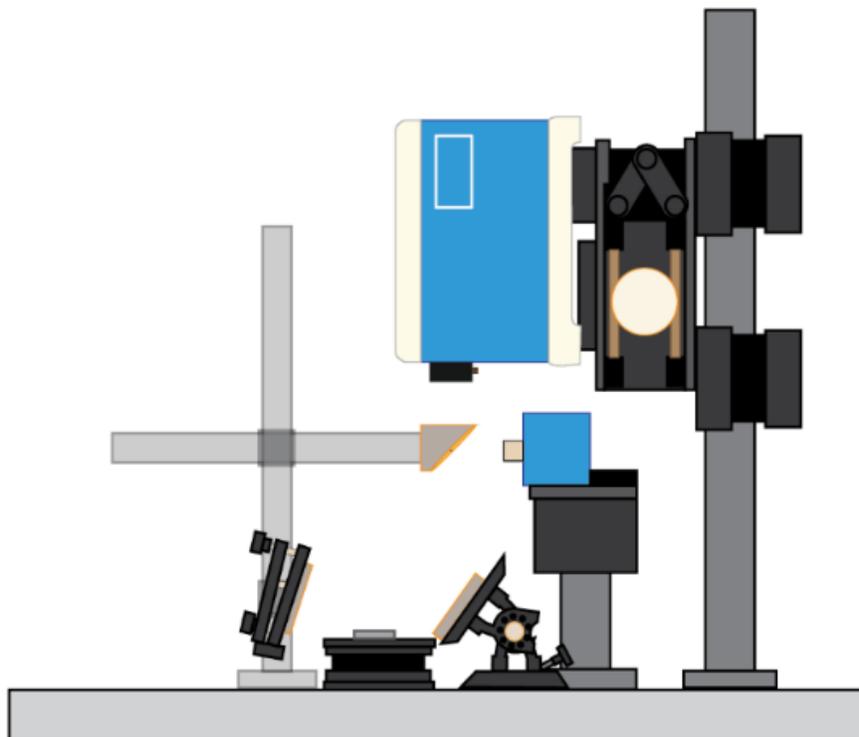


Figure 1 . QCL-GAP optical setup used for microorganism sample acquisition.

2.3 Sample Preparation

Aliquots of 20 μL of neat and mixtures of bacteria samples were deposited on the 2 X 2-in²(25.8-cm²) stainless-steel (SS) plates. Substrates were cleaned with isopropyl alcohol and left to dry in a chemical hood to avoid any sample contamination. Bacterial samples *Sa*, *Se*, and *Ml* were used to create the following mixtures: *Sa/Se*, *Sa/Ml*, and *Se/Ml*. A sample smearing technique using a micropipette tip (~1 mm diameter) was used to deposit the sample onto the substrate [20]. The smearing procedure accounts for minimal sample loss during handling. The solutions deposited on the substrate were left air-dry for 20 minutes until the solvent evaporated. Each bacterial sample acquired twenty (20) replicas by moving the optical stage in the QCL-GAP setup.

2.4 Multivariate Analysis

2.4.1 Model Development for Principal Component Analysis

MVA routines, including Principal Component Analysis (PCA), were applied to the vibrational spectra acquired for each sample. The models were created using The Unscrambler (®) X 2018 CAMO Software (0349 Oslo, Norway) version 10.5.1. Using OriginPro 8.6.0 (Northampton, MA, USA), Spectral analysis for unprocessed raw data was made. Several pre-processing methods were applied to the raw data. Standard normal variate (SNV), Multiplicative scatter correction (MSC), Savitzky-Golay first derivative (SG1), and Savitzky-Golay second derivative (SG2). These processing methods were applied to each data set, including combinations to separate the classes.

2.4.2 Model Development for Partial Least Squares Discriminant Analysis

The UnscramblerX software, similar to the PCA model development, was used to develop the PLS-DA models. Three PLS-DA models were developed to discriminate between the three microorganisms. These models were designed for binary classification with two classes: samples containing specific microorganisms ($y = 1$) or those without microorganisms ($y = -1$). The models were developed using 7 latent variables, cross-validated results with 20 segments of random samples, and the scores and loadings were calculated using the NIPALS algorithm. The threshold for distinguishing between the two classes was zero (0). Several figures of merit were used to evaluate the PLS-DA model performance: the percent of correctly classified samples (%T), percent of incorrectly classified samples (%F), sensitivity (SEN), specificity (ξ), and Matthews correlation coefficient (MCC). A value of negative one (-1) for MCC indicates the perfect classification of the opposite variable, a value of one (1) is the correct classification of the class, and zero (0) indicates random classification.

$$\%P = \frac{TP+TN}{TP+TN+FP+FN} \times 100\% \quad (1)$$

$$\%F = \frac{(FP+FN)}{TP+TN+FP+FN} \times 100\% \quad (2)$$

$$SEN = \frac{TP}{(TP+FN)} \quad (3)$$

$$\xi = \frac{TN}{(TP+TP)} \quad (4)$$

$$MCC = \frac{(TP \bullet TN) - (FP \bullet FN)}{\sqrt{(TP+FP) \bullet (TP+FN) \bullet (TN+FP) \bullet (TN+FN)}} \quad (5)$$

3 Results

3.1 QCL-GAP spectra of microorganisms

The vibrational response obtained from the microorganisms using the QCL-GAP was used to determine the functional groups used for characterization. Because of the spectral similarities among bacterial species, it is difficult to discriminate using spectrum profile differences associated with sample homogeneity at the substrate surface. Sample distribution in the deposition technique used and the uniformity of the bacteria on the surface after the sample drying process also affect the discrimination of bacterial species. These factors can impact the band characterization and the discrimination processes on bacteria mixtures. The optical

arrangement in the QCL-GAP creates an ellipse that covers a bigger surface area. This setup makes the beam size significantly larger than a typical IR system. Analysis of the spectral profile of the microorganisms under study shows characteristic signatures that allow identification and discrimination between the species.

Variation in the sample homogeneity during the deposition in the substrate is expected during spectral acquisition. However, the QCL-GAP ellipse effect can overcome those variations by covering more areas to enhance the signals during the spectrum acquisition. This effect played a role during the spectrum acquisition, providing more information that led to acceptable discrimination of the bacteria on pure and mixture of the bacteria.

The spectral measurements for the substrates (references or blanks) were measured at multiple regions of the deposited sample in the substrate. Before measuring the spectra of the neat bacteria and mixtures of bacteria, the spectrum of the substrate without the analyte is acquired as a reference. Then, the substrate containing the bacteria was collected. The instrument allows background subtraction without the need for additional processing. The performance was also verified before each run by collecting a blank spectrum using a gold (Au) substrate.

Band assignments in the fingerprint region for the bacteria sample's present spectra variations due to the substrate's reflectivity. In this experiment, the SS substrates used had a high reflectivity that allowed acquiring as many signals as possible. It is common knowledge that higher porosity of the substrate will reduce the signal to noise because it will absorb the light, resulting in fewer signals. That effect can lead to losing important information, impacting the resolution, and complicating the discrimination process.

Representative spectra for the *Se* and *Ml* mixture and neat *Se* and *Ml* spectra are shown in Figure 2. Although differences are observed in the mixture profile compared with the neat bacteria spectra, some bands have similarities. The same effect was observed for the other mixtures, such as *Sa* and *Ml*, shown in Figure 3, and *Sa* and *Se*, shown in Figure 4.

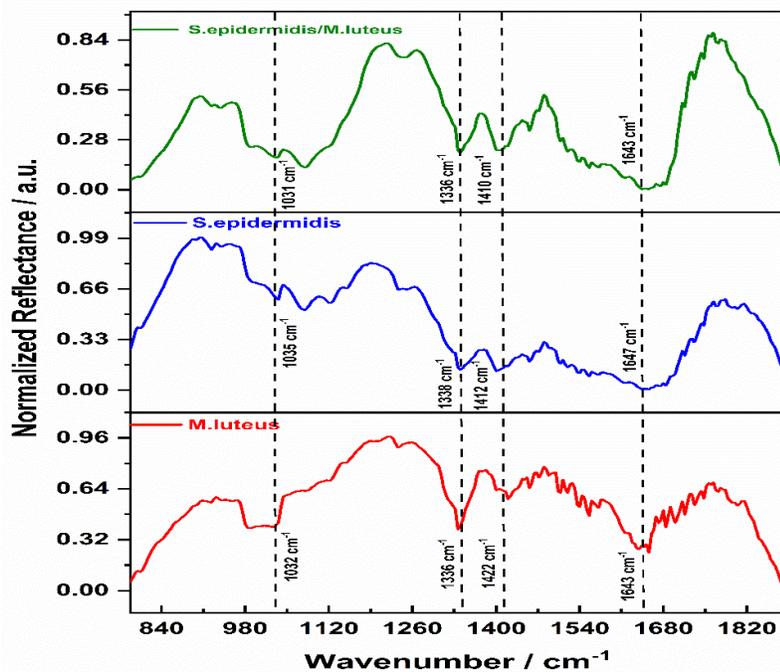


Figure 2. Normalized reflectance stacked spectra and proposed band assignment of *Se* and *Ml* mixture (top) and *Se*(center) and *Ml* (bottom) neat bacteria spectrum deposited in stainless steel substrate.

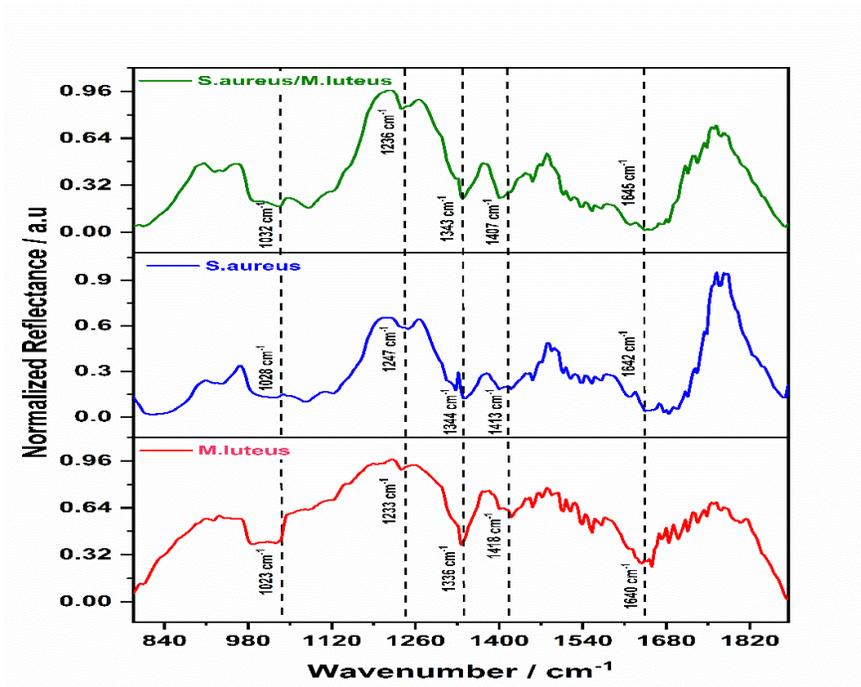


Figure 3. Normalized reflectance spectra for neat *Ml* (bottom), *Sa* (center), and a mix of *Sa:Ml* (top) deposited on an SS substrate. Tentative band assignments are included.

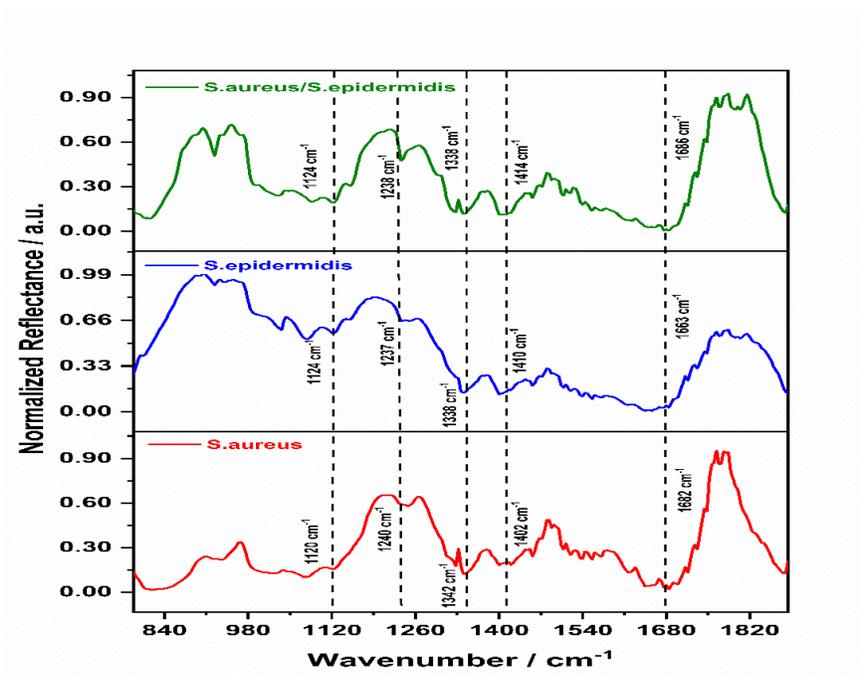


Figure 4. Normalized reflectance spectra for neat *Sa* (bottom), *Se* (center), and a mix of *Sa:Se* (top) deposited on an SS substrate. Tentative band assignments are included.

Some preliminary bands assignments summarized in Table 1, are identified in all bacteria, such as at 1642 - 1647 cm^{-1} Amide I, 1412 - 1422 cm^{-1} -C-C-O str (sym) of -COO, 1336 - 1344 cm^{-1} , Amide III, 1247 cm^{-1} -PO-PO⁻² asymmetric (phosphate I) and 1028 - 1032 cm^{-1} Glycogen. Once the bands were assessed and identified, the data were analyzed using principal component analysis (PCA). Results obtained from the analysis are summarized in section 3.22 of this paper.

Table 1. QCL-GAP spectra proposed band assignments

Bacterial Specie	Frequency (cm^{-1})	Band Assignment*
<i>Sa</i>	1642	amide I
	1413	C-O str (sym) of COO
	1344	amide III
	1247	PO-2 asymmetric (phosphate I)
	1028	glycogen
<i>Se</i>	1647	amide I
	1412	C-O str (sym) of COO
	1338	amide III
	1035	glycogen
<i>Ml</i>	1643	amide I
	1422	C-O str (sym) of COO
	1336	amide III
	1032	glycogen

*Tentative band assignment references.^[18,20]

3.2 Principal Component Analysis

MVA statistical routines were applied to the raw spectroscopic data of the microorganisms under study. The use of chemometrics in spectroscopy has been well documented in the MIR and Raman spectroscopies [20,22-23]. The MVA routine allows determining the variables responsible for separating and classifying the classes. For this work, the classification of the microorganisms was evaluated by performing pre-processing algorithms to reduce common scattering and background effects that are inevitable during spectral acquisition. It is common to see spectral variation effects due to sample heterogeneity, scattering attributed to particle size, solvent vapor pressure, and analyte solubility. These effects are corrected with the pre-processing algorithms, which allow correction in the data for accurate classification. A comparison between the raw unprocessed spectral data and pre-processed is shown in Appendix A, Figure 1A.

Principal Component Analysis (PCA) models were generated for the microorganisms considered for this discussion as neat and mixtures of these analytes representing a complex environment. The models were built on vibrational information using the 788 - 1884 cm^{-1} spectral region. First, the PCA model for the neat *Ec*, *Bt*, and *Se* bacteria is shown in Figure 5. This model was generated using the Standard Normal Variate (SNV) processing algorithm. It is represented by the scores of principal components 1 (PC-1) vs. PC-2 for the bacteria deposited on SS. PC-1 and PC-2 scores represented the largest percentage of variation for the acquired spectra. The PC-1 accounts for 52% of the total spectral variance and 19% for the PC-2.

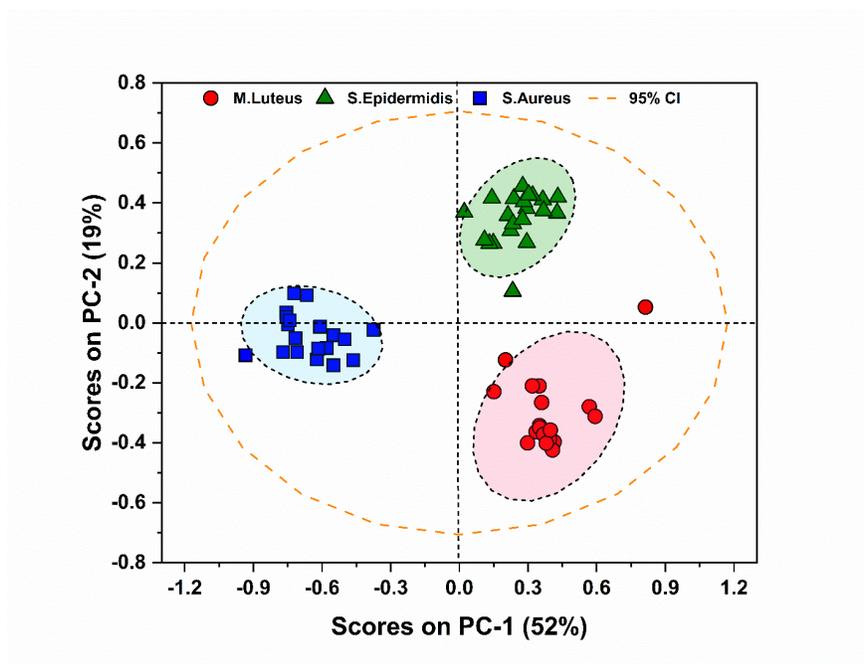
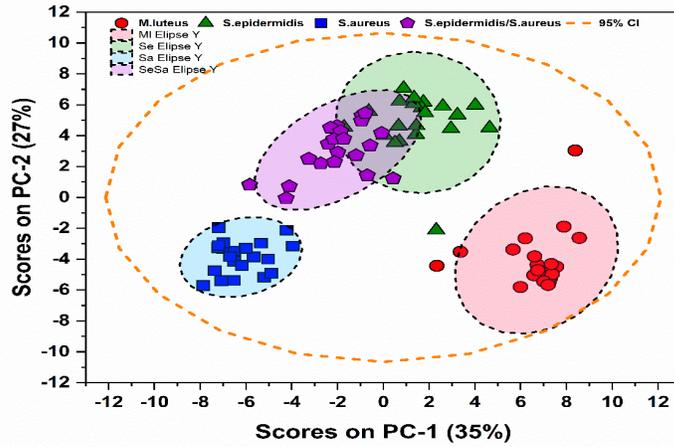
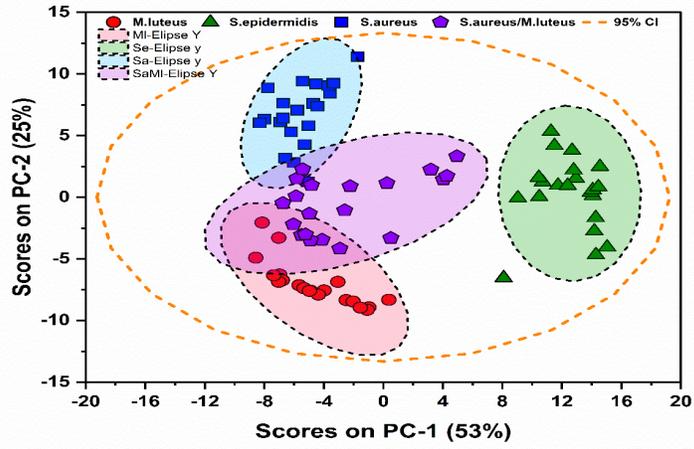


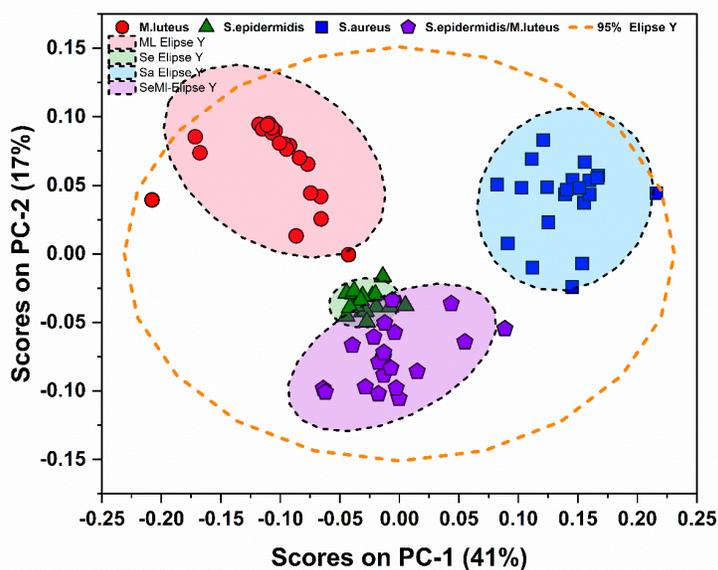
Figure 5. Scores plot of PC-1 vs. PC-2 for *Ml* , *Se* , and *Sa* as neat species.

After understanding the behavior of the *Ml* , *Sa* , and *Se* bacteria as neat species, the spectra for the mixtures were evaluated. Models shown in Figure 6 were generated to provide a clear view of a complex environment where a combination of these species might appear on the surface. Figure 6A shows the PCA model for the *Sa/Ml* mixture with a PC-1 score of 53% and a PC-2 score of 25%. An SNV pre-processing algorithm was applied to achieve class separation. The vibrational signatures of the mixture show a class that lies between the neat substances of the mixture. A difference is shown for the *Se* class, which appears distanced from the mixture and components. The same analysis was carried out for the *Se/Sa* mixture, Figure 6B, where the mixture classification showed a tendency for the *Se* class. The mixture is separated from the *Ml* component. Finally, a third PCA model was generated for the *Se/Ml* , shown in Figure 6C, by applying SNV and SG2 pre-processing algorithms. A second pre-processing step was expected since *Se* showed a higher turbidity than *Ml* . The effect may be attributed to a lower bacterial *Ml* concentration. The model also showed a tendency for the mixture towards the *Se* neat component. A further evaluation was performed by analyzing the models' contributing signals. This analysis required comparing the vibrational signature of the neat species with the loadings spectral values used for the PCA model.

(A)



(B)



(C)

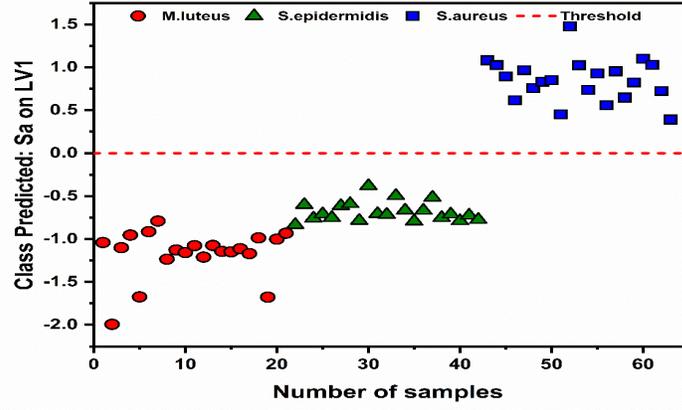
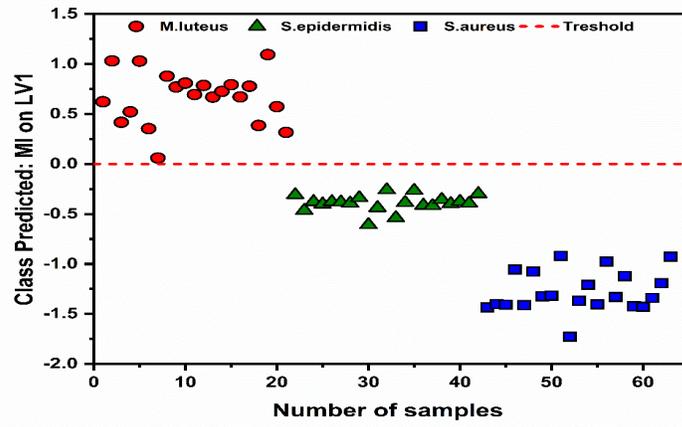
Figure 6. Scores plot of PC-2 vs. PC-1 (principal components for *Ml* , *Se* , *Sa*, and **A**) *Sa/Ml* mixture with SNV pre-processing algorithm; **B**) *Se/Sa* mixture with SG1-SNV and **C**) *Se/Ml* mixture with SNV-SG2.

3.3 PLS-DA Analysis

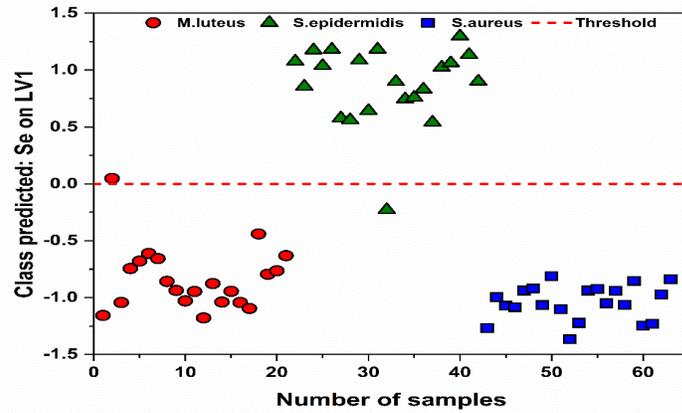
The PLS-DA models were developed for the three microorganisms using the two optimal pre-processing algorithms determined with the PCA models. Therefore, PLS-DA models were made using SG1+SNV and SNV+SG2, respectively. The best factor for each PLS-DA model was found by comparing the classification parameters. Next, the results from both pre-processing methods are compared.

The optimal results for *Ml* were obtained with SNV+SG2, *Se* with SG1+SNV, and *Sa* with SNV+SG2 from this analysis. For *Ml* , the PLS-DA model discriminates all samples perfectly with a %T of 100%, %F of 100%, SEN of 1.0, ξ of 1.0, and MCC of 1.0 for both pre-processing methods. Furthermore, the *Ml* PLS-DA model classifies *Sa* without difficulty, while the *Ml* samples are close to the threshold value of zero. *Se* appears close to the threshold, indicating slight difficulty; however, the samples are farther from the threshold than *Ml* . For *Se*, each group has a larger distance to the threshold than the *Ml* model; however, the model classifies two samples incorrectly from two different classes. Table 2 indicates that SNV+SG2 has better results than SG1+SNV for *Se* with a %T of 98.4%, %F of 1.6%, SEN of 1.0, ξ of 1.0, and MCC of 1.0. The PLS-DA models for *Sa* discriminate all samples perfectly with the same results as *Ml* , as shown in Table 2. Visualizing the predicted values vs. the number of samples shows that *Sa* has better classification than *Ml* and *Se* due to the values being far from the threshold and the clusters having a more compact structure. On the other hand, if separating all three microorganisms is desired, the *Ml* PLS-DA model with SNV+SG2 shows the best results since all the clusters do not overlap.

(A)



(B)



(C)

Figure 7. PLS-DA model for **A)** *Ml* with SNV+SG2;**B)** *Sa* with SG1+SNV; and **C)** *Se* with SNV+SG2 pre-processing.

Table 2. PLS-DA model evaluation results for *Ml*, *Se*, and *Sa* obtained with the optimal pre-processing methods: SNV+SG2 and SG1+SNV

	M. luteus	M. luteus	S. epidermidis	S. epidermidis	S. aureus	S. aureus
DP	SNV+SG2	SG1+SNV	SNV+SG2	SG1+SNV	SNV+SG2	SG1+SNV
%T	100.0%	100.0%	98.4%	96.8%	100.0%	100.0%
%F	0.0%	0.0%	1.6%	3.2%	0.0%	0.0%
SEN	1.0	1.0	1.0	1.0	1.0	1.0
ξ	1.0	1.0	1.0	1.0	1.0	1.0
MCC	1.0	1.0	1.0	0.9	1.0	1.0

4 Conclusions

Based on the study, it can be concluded that the results obtained demonstrate that the QCL-GAP operating in the 788 – 1884 cm^{-1} range produced high-quality spectral information from the bacterial species studied: *Sa*, *Se*, and *Ml*. Although the bacteria under investigation belong to the same family, the information provided by the QCL-GAP had sufficient data to discriminate the bacteria mixtures from the neat bacteria. This was demonstrated by the fact that in all PCA analyses with various combinations, the tendency was to observe the ellipse corresponding to the mixture positioned between the neat bacteria combination ellipses. For example, the ellipse for the mixture containing *Se* and *Sa* was positioned between the neat *Se* and *Sa* ellipses. The same effect was observed for the other combinations, such as *Se* / *Ml* and *Sa* / *Ml*, where the ellipses were positioned between *Se* and *Ml* and *Sa* and *Ml* neat bacteria, respectively. Developing PLS-DA models to discriminate one microorganism at a time, results show that an average of 99.2% of microorganisms were classified correctly,

Therefore, this study demonstrated the capability of QCL-GAP in combination with PCA to discriminate between bacteria from the same family. The development of this new methodology for analyzing bacteria using QCL-GAP provides fast and accurate analysis for detecting microorganisms. It is accompanied by a great potential to discriminate between similar types of microorganisms, as was demonstrated in this study.

Acknowledgments

I want to acknowledge and warmly thank my advisor, Professor Samuel P. Hernández-Rivera, Ph.D., for his trust in my capabilities, potential, and day-to-day support during my weakest moments during this journey. You made this possible. His guidance and advice carried me through all the stages of writing my project. I would also like to thank my committee members, Prof. Carmen A. Vega-Olivencia, Dr. Nairmen Mina-Camilde, Dr. Carlos Ríos-Velázquez, and Dr. Enrique Meléndez-Martínez for their support in the revision of my publications and dissertation, for making this challenging process an enjoyable moment and for your brilliant comments and suggestions, thanks to you. To my study partners and AbbVie, especially James Erker, Cormac Dalton, and Dave Thompson, for allowing me to continue through these challenges and support my goals.

I would also like to give special thanks to my family for their continuous support and understanding when undertaking my research and writing my project. Your prayer for me was what sustained me this far.

Finally, I would like to thank God for letting me through all the difficulties, giving me strength, and blessing all my steps during this process. I have experienced your guidance day by day. You are the one who let me finish my degree. You will continue building my future, career, and life goals and guide my family.

This material is based upon work supported by the US Department of Homeland Security under Grant Awards 22STESE00001-02-00 and 21-CWDARI 00042-01-00 Science and Technology Directorate, Office of University Programs. The views and conclusions contained in this document are those of the authors. They should not be interpreted as necessarily representing the official policies, either expressed or implied, of the US Department of Homeland Security. Additional support from the US Department of Agriculture, National Institute of Food and Agriculture award # 2022-77040-37623 is also acknowledged.

Conflicts of Interest:

The authors declare no conflict of interest.

Data availability

The data supporting this study's findings are available from the corresponding author upon reasonable request.

References

1. V. Mugoyela, K.D. Mwambete, Microbial contamination of nonsterile pharmaceuticals in public hospital settings. *Ther Clin Risk Manag* 2010;6:443- 8. doi: 10.2147/TCRM.S12253
2. H. Kobayashi, B. E. Rittmann, Microbial removal of hazardous organic compounds. *Environ Sci Technol* 1982;16(3):170A-83A. doi: 10.1021/es00097a002
3. T. R. Garrett, M. Bhakoo, Z. Zhang, Bacterial adhesion and biofilms on surfaces. *Prog Nat Sci*. 2008;18(9):1049-56. doi: 10.1016/j.pnsc.2008.04.001
4. M. Ratajczak, M. M. Kubicka, D. Kaminska, P. Sawicka, J. Dlugaszewska, Microbiological quality of nonsterile pharmaceutical products. *Saudi Pharm J* 2015;23(3):303-7. doi: 10.1016/j.jsps.2014.11.015
5. M. Harz, P. Rosch, K.-D. Peschke, O. Ronneberger, H. Burkhardt, J. Popp, Micro-Raman spectroscopic identification of bacterial cells of the genus *Staphylococcus* and dependence on their cultivation conditions August 2005 DOI: 10.1039/b507715j
6. K. Maquelin, C. Krischner, L. Choo-Smith, N. Ngo-Thi, M. Vreeswijk, G. J. Puppels, Clin. Identification of medically relevant microorganisms by vibrational spectroscopy. *Journal of Microbiological Methods* 51 (2002) 255 – 271
7. E. H. Ibrahim, G. Sherman, S. Ward, V. J. Fraser, H. Marin, The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting *Chest* 2000;118;146-155 DOI: 10.1378/chest.118.1.146
8. The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting *Chest* 2000;118;146-155 DOI: 10.1378/chest.118.1.146
9. D. Hofstetter, J. Faist, "High-performance quantum cascade lasers and their applications". *Top Appl Phys*. 2003. 89, 61-96.
10. J. Faist, F. Capasso, D. L. Sivco, C. Sirtori, A. L. Hutchinson, A. Y. Cho, Quantum cascade laser, *Science*, Vol. 264, Issue 5158, 1994, pp. 553-556.
11. F. K. Tittel, D. Richter, A. Fried, [Mid-Infrared Laser Applications in Spectroscopy], I. T. Sorokina, K. L. Vodopyanov, (Eds.): *Solid-State Mid-Infrared Laser Sources*, *Topics Appl. Phys.* 89, 445–516, **2003** .
12. A. A. Kosterev, F. K. Tittel, Chemical Sensors Based on Quantum Cascade Lasers, *IEEE J. Quantum Electron.* 38 (6), 582-591, **2002** .
13. L. Hvozدارa, N. Pennington, M. Kraft, M. Karlowatz, B. Miziakoff, Quantum cascade lasers for mid-infrared spectroscopy, *Vib. Spectrosc.* 30, 53-58, **2002** .
14. Naumann, D. Infrared Spectroscopy in Microbiology: In *Encyclopedia of Analytical Chemistry*, R.A. Meyers, 102-131, 2000.
15. A. C. Padilla-Jiménez, W. Ortiz-Rivera, C. Rios-Velazquez, I. Vazquez-Ayala, S. P. Hernández-Rivera, Detection and discrimination of microorganisms on various substrates with quantum cascade laser spectroscopy," *Opt. Eng.*, 53(6), 061611, **2014** . DOI:10.1117/10.1117/1.OE.53.6.061611.
16. H. Félix-Rivera, R. González, G. Rodríguez, O. M. Primera-Pedrozo, C. Ríos-Velázquez, S. P.

- Hernández-Rivera, Improving SERS Detection of *Bacillus thuringiensis* Using Silver Nanoparticles Reduced with Hydroxylamine and with Citrate Capped Borohydride, *International Journal of Spectroscopy*, vol. 2011, Article ID 989504, 9 pages, **2011** . <https://doi.org/10.1155/2011/989504>
17. A. Novais, A. R. Freitas, C. Rodrigues, L. Peixe, Fourier transform infrared spectroscopy: unlocking fundamentals and prospects for bacterial strain typing, *Eur. J. Clin. Microbiol. Infect. Dis.* , 38:427–448, **2019** .
 18. L. Xiaonan, M. Hamzah, A.-Q. Mengshi Lin, B. A. Rasco, Application of Mid-infrared and Raman Spectroscopy to the Study of Bacteria, *Food Bioproc. Tech.* **2011** . DOI 10.1007/s11947-011-0516-8
 19. J. Komal-Kumar, A. G. Devi-Prasad, Identification and comparison of Biomolecules in medicinal plants of *Tephrosia Tinctoria* and *Antylosia Albicans* by using FTIR, *Romanian J. Biophys.* , Vol. 21, No. 1, P. 63–71, **2011** .
 20. L. C. Pacheco-Londoño, N. J. Galán-Freyte, A. M. Figueroa-Navedo, R. Infante-Castillo, J. L. Ruíz-Caballero, S. P. Hernández-Rivera, Quantum cascade laser back-reflection spectroscopy at grazing-angle incidence using fast Fourier transform pre-processing, *J. Chemo* , 33(9), **2019** . [doi.org/10.1002/cem.3167]
 21. "Grazing Angle Probe Mount for Quantum Cascade Lasers." Inventors: Samuel P. Hernandez-Rivera and Leonardo C. Pacheco-Londoño. Ser.62/587,557; Filing Date: 11/17/2017.
 22. P. M. Fierro-Mercado, S. P. Hernández-Rivera, Highly Sensitive Filter Paper Substrate for SERS Trace Explosives Detection, *Int. J. Spectrosc.* , 2012, 1–7, **2012** .
 23. N. J. Galán-Freyte, A. M. Figueroa-Navedo, Y. C. Pacheco-Londoño, W. Ortiz-Rivera, L. C. Pacheco-Londoño, S. P. Hernández-Rivera, Chemometrics-enhanced fiber optic Raman detection, discrimination, and quantification of chemical agents simulants concealed in commercial bottles, *Anal Chem. Res.* , 2, 15–22, **2014** .
 24. P. M. Fierro-Mercado, S. P. Hernández-Rivera, Highly Sensitive Filter Paper Substrate for SERS Trace Explosives Detection. *Int. J. Spectrosc.* , 2012:1–7, **2012** .