The main objective of the project was to strengthen the existing research program in "Counter Weapons of Mass Destruction". Commercial QCL based, near field spectroscopic detection systems have been fully evaluated and successful used in detection and discrimination of explosives on test surfaces. The focus of the third year was to locate at the “needle in the haystack”. That is, discriminating for nearly trace amounts of explosives and other threat agents against a highly interfering background: the substrates on which these are deposited. Highly interfering substrates (non-reflective, non-ideal substrates) were used to test powerful multivariate analysis routines utilized for...
ABSTRACT

The main objective of the project was to strengthen the existing research program in “Counter Weapons of Mass Destruction”. Commercial QCL based, near field spectroscopic detection systems have been fully evaluated and successful used in detection and discrimination of explosives on test surfaces. The focus of the third year was to locate at the “needle in the haystack”. That is, discriminating for nearly trace amounts of explosives and other threat agents against a highly interfering background: the substrates on which these are deposited. Highly interfering substrates (non-reflective, non-ideal substrates) were used to test powerful multivariate analysis routines utilized for discriminating for the target chemical/biological threats against the predominant signals from the background substrates. Among the substrates used were cardboard, wood, plastics and luggage materials. Research has also focused on measuring the infrared signals at off conventional back-reflection geometry. Both thermal and laser sources were used to collect signals from target chemicals at ranges of 1-4 m and angles from 0 to 80°. Ambient IR spectra will also be measured under various atmospheric conditions to establish baseline measurements. This work focused on enhancing established technology and introducing novel concepts in standoff detection.
Enter List of papers submitted or published that acknowledge ARO support from the start of the project to the date of this printing. List the papers, including journal references, in the following categories:

(a) Papers published in peer-reviewed journals (N/A for none)

<table>
<thead>
<tr>
<th>Received</th>
<th>Paper</th>
</tr>
</thead>
</table>


07/09/2014 40.00 Ricardo Infante-Castillo, Samuel P. Hernández-Rivera. Predicting heats of explosion of nitrate esters through their NBO charges and 15N NMR chemical shifts on the nitro groups, Computational and Theoretical Chemistry, (02 2011): 0. doi: 10.1016/j.comptc.2010.10.038


TOTAL: 15

Number of Papers published in peer-reviewed journals:

(b) Papers published in non-peer-reviewed journals (N/A for none)

Received Paper


TOTAL: 3
Number of Papers published in non peer-reviewed journals:

4.00

(c) Presentations


Number of Presentations: 4.00

Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

Received    Paper


TOTAL: 6
Number of Non Peer-Reviewed Conference Proceeding publications (other than abstracts):

<table>
<thead>
<tr>
<th>Received</th>
<th>Paper</th>
</tr>
</thead>
</table>

TOTAL:
Number of Peer-Reviewed Conference Proceeding publications (other than abstracts):

(d) Manuscripts

<table>
<thead>
<tr>
<th>Received</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>07/07/2014 15:00</td>
<td>William Ortiz-Rivera, Amira C. Padilla-Jiménez, Carlos Ríos-Velázquez, Iris Vázquez-Ayala,, Samuel P. Hernández-Rivera. Detection and discrimination of microorganisms on various substrates with quantum cascade laser spectroscopy, Optical Engineering (09 2013)</td>
</tr>
<tr>
<td>07/07/2014 41:00</td>
<td>Hilsamar Félix-Rivera, Michael L. Ramírez-Cedeño, Rhaisa A. Sánchez-Cuprill, Samuel P. Hernández-Rivera. Triacetone triperoxide thermogravimetric study of vapor pressure and enthalpy of sublimation in 303–338K temperature range, Thermochimica Acta (02 2011)</td>
</tr>
<tr>
<td>08/31/2012  3:00</td>
<td>John R. Castro-Suarez, Leonardo C. Pacheco-Londono1, Max Diem, Thomas J. Tague, Jr., and Samuel P. Hernandez-Rivera. ACTIVE MODE FT-IR STANDOFF DETECTION OF TRINITROTOLUENE ON ALUMINUM SUBSTRATES, Vibrational Spectroscopy (08 2012)</td>
</tr>
</tbody>
</table>

TOTAL: 8
Books

Received Book

TOTAL:

Received Book Chapter


TOTAL: 4

Patents Submitted

Patents Awarded
## Graduate Students

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
<th>Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td>John R. Castro-Suarez</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Jorge Castellanos</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Jose L. Ruiz-Caballero</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Amira C. Padilla-Jimenez</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Nataly Y. Galan-Freyle</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td>Carlos Ortega-Zuñiga</td>
<td>0.50</td>
<td></td>
</tr>
<tr>
<td><strong>FTE Equivalent:</strong></td>
<td><strong>3.00</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Total Number:</strong></td>
<td><strong>6</strong></td>
<td></td>
</tr>
</tbody>
</table>

## Names of Post Doctorates

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Leonardo C. Pacheco</td>
<td>0.25</td>
</tr>
<tr>
<td>William Ortiz-Rivera</td>
<td>0.25</td>
</tr>
<tr>
<td><strong>FTE Equivalent:</strong></td>
<td><strong>0.50</strong></td>
</tr>
<tr>
<td><strong>Total Number:</strong></td>
<td><strong>2</strong></td>
</tr>
</tbody>
</table>

## Names of Faculty Supported

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
<th>National Academy Member</th>
</tr>
</thead>
<tbody>
<tr>
<td>Samuel P. Hernandez-Rivera</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td><strong>FTE Equivalent:</strong></td>
<td><strong>0.25</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Total Number:</strong></td>
<td><strong>1</strong></td>
<td></td>
</tr>
</tbody>
</table>

## Names of Under Graduate students supported

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
<th>Discipline</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amanda M. Figueroa-Navedo</td>
<td>0.50</td>
<td>Chemical Engineering</td>
</tr>
<tr>
<td>Doris Laguer</td>
<td>0.50</td>
<td>Geosciences</td>
</tr>
<tr>
<td>Gabriel Nieves</td>
<td>0.50</td>
<td>Chemical Engineering</td>
</tr>
<tr>
<td>Roxannie Gonzalez</td>
<td>0.50</td>
<td>Biosciences</td>
</tr>
<tr>
<td>Migdalia Hidalgo</td>
<td>0.50</td>
<td>Chemical Engineering</td>
</tr>
<tr>
<td><strong>FTE Equivalent:</strong></td>
<td><strong>2.50</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Total Number:</strong></td>
<td><strong>5</strong></td>
<td></td>
</tr>
</tbody>
</table>
Student Metrics
This section only applies to graduating undergraduates supported by this agreement in this reporting period

The number of undergraduates funded by this agreement who graduated during this period: ...... 6.00
The number of undergraduates funded by this agreement who graduated during this period with a degree in science, mathematics, engineering, or technology fields: ...... 6.00
The number of undergraduates funded by your agreement who graduated during this period and will continue to pursue a graduate or Ph.D. degree in science, mathematics, engineering, or technology fields: ...... 4.00
Number of graduating undergraduates who achieved a 3.5 GPA to 4.0 (4.0 max scale): ...... 3.00
Number of graduating undergraduates funded by a DoD funded Center of Excellence grant for Education, Research and Engineering: ...... 0.00
The number of undergraduates funded by your agreement who graduated during this period and intend to work for the Department of Defense: ...... 3.00
The number of undergraduates funded by your agreement who graduated during this period and will receive scholarships or fellowships for further studies in science, mathematics, engineering or technology fields: ...... 4.00

Names of Personnel receiving masters degrees

<table>
<thead>
<tr>
<th>NAME</th>
<th>Total Number:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nataly Y. Galan-Freyle</td>
<td>2</td>
</tr>
<tr>
<td>Carlos Ortega-Zuñiga</td>
<td></td>
</tr>
</tbody>
</table>

Names of personnel receiving PHDs

<table>
<thead>
<tr>
<th>NAME</th>
<th>Total Number:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sandra N. Correa-Torres</td>
<td>8</td>
</tr>
<tr>
<td>Hilsamar Felix-Rivera</td>
<td></td>
</tr>
<tr>
<td>William Ortiz-Rivera</td>
<td></td>
</tr>
<tr>
<td>Pedro Fierro Mercado</td>
<td></td>
</tr>
<tr>
<td>Marcia del R. Balaguera-Gelves</td>
<td></td>
</tr>
<tr>
<td>G. Marcela Herrera-Sandoval</td>
<td></td>
</tr>
<tr>
<td>Eduardo A. Espinosa-Fuentes</td>
<td></td>
</tr>
<tr>
<td>Amira C. Padilla-Jimenez</td>
<td></td>
</tr>
</tbody>
</table>

Names of other research staff

<table>
<thead>
<tr>
<th>NAME</th>
<th>PERCENT_SUPPORTED</th>
<th>FTE Equivalent:</th>
<th>Total Number:</th>
</tr>
</thead>
</table>

Sub Contractors (DD882)

Inventions (DD882)

Scientific Progress
Technology Transfer
I. ABSTRACT

The need for rapid detection and identification instrumentation for purposes of deterring chemical and biological threats that could potentially be used as Weapons of Mass Destruction (WMD) in national defense and security applications has become a concern of immense importance in modern society. From security and anti-terrorist personnel, to first responders and law enforcing employees such as police officers, airport screeners, and border patrols personnel, to Navy, Army, Air Force, and National Guard workforces, the possibility of coming in contact with a chemical or biological threat is high. To achieve the required countermeasures in chemical/biological threats mitigation, vibrational spectroscopic techniques are frequently used. In the past few years, many of the published reports have focused on the detection of these important chemical/biological threats. However, the majority of them require some type of sampling. Obtaining samples in the field is the principal disadvantage of most threat detection devices because the
person doing the sampling is at risk. This project focused on evaluating mid-infrared (MIR) sources including eye safe laser sources such as quantum cascade lasers (QCLs) in remote detection experiments of chemical and biological threat agents (including explosives).

The main objective of the project was to strengthen the existing research program in “Counter Weapons of Mass Destruction”. Commercial QCL based, near field spectroscopic detection systems have been fully evaluated and successful used in detection and discrimination of explosives on test surfaces. The focus of the third year was to locate at the “needle in the haystack”. That is, discriminating for nearly trace amounts of explosives and other threat agents against a highly interfering background: the substrates on which these are deposited. Highly interfering substrates (non-reflective, non-ideal substrates) were used to test powerful multivariate analysis routines utilized for discriminating for the target chemical/biological threats against the predominant signals from the background substrates. Among the substrates used were cardboard, wood, plastics and luggage materials. Research has also focused on measuring the infrared signals at off conventional back-reflection geometry. Both thermal and laser sources were used to collect signals from target chemicals at ranges of 1-4 m and angles from 0 to 80°. Ambient IR spectra will also be measured under various atmospheric conditions to establish baseline measurements. This work focused on enhancing established technology and introducing novel concepts in standoff detection.

Keywords: Standoff Detection, Quantum Cascade Laser Spectroscopy, Chem/Bio Threats, Weapons of Mass Destruction
II. TABLE OF CONTENTS

III. PROJECT DESCRIPTION 4

IV. SUMMARY OF THE MOST IMPORTANT RESULTS 5

V. PROJECTS

1. DETECTION AND DISCRIMINATION OF MICROORGANISMS ON SUBSTRATES USING QUANTUM CASCADE LASER SPECTROSCOPY 6

2. DETECTION OF HIGHLY ENERGETIC MATERIALS ON NON-REFLECTIVE SUBSTRATES USING QCL SPECTROSCOPY 28

3. SPECTROSCOPIC DETECTION OF BACTERIA USING TRANSMISSION MODE QCL SPECTROSCOPY 51

VI. STATUS AND OVERALL PROJECT PROGRESS: FINAL DEVELOPMENTS 65

VII. LEVERAGING OF RESOURCES 66

VIII. PROJECT DOCUMENTATION AND DELIVERABLES 71

IX. BIBLIOGRAPHY 77
III. PROJECT DESCRIPTION

The objective of this work was to develop new technologies and extending the existing technologies based on Remote Infrared Spectroscopy (RIRS) as applied to detection of chemical and biological threats (CBTs). The pursued enhancements are based on lowering current detection limits in near-field detection (1-10 m), improving range detection (distances up to and longer than 100 meters), and detecting signatures of realistic CBT-related materials as they pertain to the detection of improvised explosive devices (IEDs), homemade explosives (HMEs) and weapons of mass destruction (WMDs). It is expected that applications of the technologies developed as part of this DoD HBCU/MI project will contribute significantly to solving problems related to remote detection of hazardous chemicals, including HMEs, chemical warfare agents (CWAs), toxic industrial compounds (TICs) and biological threat agents (BTAs). The technologies will enable the detection of small amounts of chemicals, their formulations and their degradation products via the methodologies to be developed. The techniques based on remote detection are non-invasive, rapid, real-time or nearly so (1-60 s), and eye safe, even when using laser beams. The work addressed advancing the instrumentation by designing optimized versions of mid-infrared (MIR) QCL based standoff detection systems. The issue of selectivity and interferences was addressed as well as other issues that can be used for the prediction of spectral signals (signatures and sensitivity) of environmentally exposed CBTs. The outcome of these studies will be a better understanding of the spectroscopic signatures of exposed CBTs and the ability to predict the performance of spectroscopic measurement techniques for measuring many different types of threats.

Specific aims for achieving objectives and project goals were the following:

- Advance the sensitivity and range of remote infrared detection of chemical and biological threats by developing remote detection methodologies based on a quantum cascade scanning laser system operating in the mid infrared region.
- Determine spectral signatures from samples of threat agents deposited on realistic sample matrices (non-ideal, non-reflective substrates).
• To build a library of samples containing different amounts of target compounds adsorbed onto real sample matrices and observe how these materials age and are chemically altered following environmental exposure and how these specific changes influence the spectral signatures and intensities.

• Perform extensive discrimination studies. Develop and test Chemometrics-based (partial least squares, PLS) routines for spectroscopic quantification coupled to linear discrimination analysis routines to increase the selectivity required to examine complex matrices containing HE, CWAS and BTAs.

IV. SUMMARY OF THE MOST IMPORTANT RESULTS

Several projects have been completed as part of the efforts in vibrational spectroscopy based chemical detection of chem/bio threats. These include MIR spectroscopy based detection of explosives using QCL sources. Work continued in remote detection of explosives on other substrates such as plastic, wood, leather, cloth, and other metallic surfaces in order to quantify and obtain chemometrics models that can be applied in real world environments. Detailed theoretical study about optical processes on these surfaces is required to understand the nature of the results of the experiments. Transition of standoff laser induced thermal emission (LITE) of explosives to real world substrates using complex matrices such as sand, plastic, wood, leather, cloth, and other metallic surfaces has also begun.

Work on QCL based detection of explosives was pursued in collaboration with industrial partners for development of hand-held MIR explosives detection equipment. Still other companies are being considered as partners for collaborative and transitioning work. The commercial applications have been envisioned in the standards market for cyclic organic peroxides (CAP) HME, including DADP, HMTD and TMDD.

The Chemical Measurement and Imaging Program supports research focusing on chemically-relevant measurement science and imaging, targeting both improved understanding of new and existing methods and development of innovative approaches and instruments. For which the primary focus is on development of new instrumentation enabling chemical measurements likely to be of wide interest and utility to the chemistry research community.
V. PROJECTS

PROJECT 1: DETECTION AND DISCRIMINATION OF MICROORGANISMS ON SUBSTRATES USING QUANTUM CASCADE LASER SPECTROSCOPY

QCLS/PLS-DA detection and discrimination of microorganisms.

INTRODUCTION

Defense and security agencies, as well as the private industries, are highly interested in finding new ways to detect and identify unknown biological threats. By developing new capabilities and by expanding current experiences, food industries, environment protection agencies, and pharmaceutical and biotechnology industries may benefit from the application of infrared sensing. Recent studies have focused on the development of rapid and accurate methods for recognizing agents that represent specific microorganisms. For example, disposable electrochemical immunosensor [1] and immobilized probes for the detection of *Escherichia coli* (*Ec*) [2] and solid phase microextraction/gas chromatography/mass spectrometry for distinguishing bacteria have been developed [3]. However, these techniques are time consuming, expensive, and involve many preparation steps and selective pre-enrichment. Although the identification
and discrimination of bacterial spores with mid-infrared (MIR) technologies have been reported [4, 5], this contribution describes the first application of a tunable MIR quantum cascade laser (QCL)-based spectrometer for identification and discrimination of bacteria. The QCLs have revolutionized many areas of research and development related to applications in defense and security [6, 7].

The QCL sources are unipolar semiconductor injection lasers that are based on interband transitions in a multiple quantum-well heterostructure [8]. These lasers operate in the MIR at wavelengths starting at approximately 5 μm and expanding (in narrow or wide coverage) to 12.5 μm, which matches very well with the fundamental vibrational absorption bands of many chemical and biological species relative to the conventional diode sources, where laser emission generally matches the weaker overtones, or to oxide based thermal sources (such as the globar source), which are much less intense. The emission wavelength of a QCL depends on the thickness of the quantum well and barrier layers in the active region rather than on the band gap of the diode lasers. The QCLs operate at near room temperature, producing from milliwatts to watts of IR radiation, both in continuous mode and pulsed laser operation. More recently, tunable QCL systems within a broad range of frequencies available for MIR experiments have become accessible [9]. In reflection mode, the same optical device is used to project the beam onto the analyte and to collect the reflected radiation.

*Bacillus thuringiensis (Bt)*, a gram-positive bacterium that can form endospores which are highly resistant to chemical and thermal extremes in their latent state, was one of the bacteria chosen for this study [10]. The life cycle of *Bacillus* species includes vegetative cell growth that can form endospores as a defense mechanism. Endospores are highly resistant to environmental stresses including high temperatures, irradiation, strong acids, and disinfectants. In addition, endospores can tolerate extreme environments. This tolerance makes endospores suitable for transport before or during biological attack [11]. *Bt* is not harmful to humans and was chosen as a model in this study due to its similarities with *Bacillus anthracis* (anthrax), which has a high potential for use in terrorist attacks. Similarly, *Staphylococcus epidermidis (Se)* is a gram-positive cocci bacterium that is commonly associated with humans (usually with human skin) [12]. *Ec* is a member of the Enterobacteriaceae family of gram-negative bacteria. *Ec* is a thermotolerant coliform that
occurs in the intestines of warm-blooded animals. Thus, Ec is an indicator of fecal contamination [13].

Various substrates, including cardboard (CB), glass, travel bags (TBs), wood, and stainless steel (SS), were used as support surfaces for the bacteria. Because the vibrational spectra of bacterial cells consist of signal contributions from all the cell components, these spectra reflect the overall molecular compositions of the cells. The MIR studies of Naumann et al., [14–16] were extended to identify and discriminate between different vegetative bacteria with chemometric methods. These data were reduced to demonstrate the capability of this spectroscopic technique for identifying and discriminating between the types of bacterial cells that are considered as biological simulants [17].

EXPERIMENTAL SETUP

Samples preparations

Bacterial strains of Bt (ATCC #35646), Ec (ATCC #8789), and Se (ATCC #2228), were provided by the Microbial Biotechnology and Bioprospecting Lab at the Biology Department of the University of Puerto Rico-Mayagüez. These biological agents were selected for testing based on the previous studies, which included their resemblance to real-world biothreats and microbiological differences between the individual bacteria. Pure cultures were grown in Miller-modified Luria–Bertani (LB) agar and broth (Fisher Scientific International, Thermo Fisher Scientific, Waltham, MA). Prior to analysis, these cultures were stored at −80°C in microvials that contained 20% glycerol (cryoprotectant). After satisfactory growth of Bt, Ec, and Se, colonies were attained in agar, isolated on LB plates, and inoculated into 5 mL of LB or tryptic soy broth. These colonies were allowed to grow overnight. Se and Bt were placed in an orbital shaker at 32°C (∼120 rpm) for 24 h before culturing for 72 h. Ec was placed in an orbital shaker at 37°C and allowed to grow for 5 h after inoculation. The subcultures were diluted (1:50) in their appropriate media and centrifuged at 5 K for 5 min at 4°C to obtain bacterial pellets.

These pellets were washed once with 20 mL of a 1% phosphate buffered saline (PBS) solution to remove the growth media. Finally, the pellets were resuspended in 4.0 mL of PBS (Ec and Se) or distilled water (Bt) before allowing them to grow for 72 h at 32°C and
at 250 rpm. The harvested spores were stored at 4°C in distilled-deionized H₂O until the experiment was performed. Serial dilution was performed by plating the spores on LB agar to count the bacteria that were suspended in the PBS. These plates were incubated at 37°C to obtain colonies. The colonies were counted to determine the bacterial concentration in the stock suspension based on the dilution series. The following concentrations were obtained: 5.9x10¹⁰/5.0x10³, 3.5x10⁹ and 9.5x10⁹ colony forming units per milliliter (CFU/mL) for Bt vegetative cell/endospores, Ec and Se respectively. Because the incubation time for the microorganisms must be determined from their growth curves (starting at an optical density of 0.025 at a wavelength of 600 nm), the OD₆₀₀ (Biophotometer, Eppendorf North America, Hauppauge, New York) was measured before and after centrifugation.

The scanning electron microscope (SEM) images were obtained with a JEOL-JSM 6500 and a XL series 30S Philips/FEI to verify the size and morphology of bacterial cells. Figure 1 shows rod-shaped Bt and Ec with a size of between 0.5 to 1.0 × 1.4 to 3.0 μm. In addition, Se was spherically shaped with a size of 0.5 × 1.5 μm.

![SEM micrographs of bacteria](image)

**Figure 1.** The SEM micrographics of bacteria used in this study, including Bt, Ec and Se. Operating voltage: 15 kV.

**Instrumentation**

A QCL-based dispersive IR spectrometer (LaserScan™, model 712, Block Engineering, LLC, Marlborough, MA) was used to acquire MIR reflectance data for the three bacterial samples on the wavenumber range of 830 to 1430 cm⁻¹. LaserScan™ uses infrared reflection spectroscopy to provide high-quality reflectance spectra from
materials deposited on surfaces and from bulk substrates. The MIR laser provides the LaserScan™ with a source capable of sensitivities that are orders of magnitude larger than IR spectroscopy systems equipped with thermal sources. Added to that are the properties related to laser sources: polarized and coherent. The spectral resolution achievable by QCL sources can be very high at the cost of sacrificing tunability. However, by coupling the spectral information provided with the QCL spectroscopic system with the powerful multivariate analysis of chemometrics, highly interfering contributions from the substrates can be accounted for and separated from spectral information of bacteria.

**Figure 2.** Experimental setup of the quantum cascade laser (QCL) mid-infrared (MIR) spectroscopic system that was used to detect *Bl, Ec* and *Se* on real-world substrates, including stainless steel (SS), glass, cardboard (CB), travel bag (TB) and wood.

Bacteria detection on the different substrates required a simple sample preparation procedure that consisted of depositing 10 μL of each bacterial sample suspension on the selected solid substrates (CB, TB, wood, glass, and SS) for in situ measurements [18, 19]. Biosamples were then allowed to dry to simulate real-world samples. Aluminum plates were used to support the substrates, which contained between 10 and 15 samples.
of each of the three analyzed bacteria that were deposited on the substrates over an area of $1 \times 1$ cm. Experimental setup of the QCL MIR spectroscopic system that was used to detect $Bt$, Ec, and Se on real-world substrates, including SS, glass, CB, TB, and wood, is shown in Fig. 2.

Substrates with and without bacterial samples were used to measure the reflectance spectra of bacteria deposited on the substrates with the LaserScan™ (Fig. 4.2). Clean substrates were used to measure the background spectra. The QCL spectrometer had a focal point distance of 6 in. The system was sensitive to the way the material was deposited, with a light-capturing efficiency of approximately 1% for the ideal substrate. Therefore, the angle of sample placement was critical. In addition, the light from the source could be lost due to absorption or multiple reflections from the material of interest, especially for thick surfaces. Visible pointers (HeNe laser beams) were aligned with the invisible MIR radiation to form a laser spot (approximately $2 \times 4$ mm) on the surface. In addition, built-in algorithms were used for nearly real-time sample detection. Trace and bulk concentrations of the other substances tested were detected in the field with a minimum of 500 ng/cm$^2$ [20].

OPUS 6.0™ software package for data acquisition and analysis (Bruker Optics, Billerica, Massachusetts) was used for data analysis. During analysis, the spectra were normalized to be compared. For multivariate data analysis, the matrix interlinks of the individual measurements were generated in rows, and the selected wavenumber values were generated in columns. The spectral intensity of each intensity-wavenumber combination formed the elements of the matrix. Principal component analysis (PCA) regression analysis algorithm was initially used to analyze the data. The PCA is frequently used to reduce the number of variables in an experiment. With PCA, datasets with many variables can be simplified (by data reduction) to easily interpret the results. Partial least squares-discriminant analysis (PLSDA) is a multivariate method that is used to classify samples and to reduce the number of variables. For this analysis, it is assumed that the differences between groups will dominate the total variability of the samples. The PCA and PLS-DA routines were applied using the MATLAB™ computational environment (The MathWorks, Inc., Natick, Massachusetts) coupled to the PLS-Toolbox™ v. 7.0.3 (Eigenvector Research, Inc., Wenatchee, Washington).
RESULTS AND DISCUSSION

MIR spectra of microorganisms were used to identify the molecular vibrational modes in the biosamples. These vibrational modes contain valuable information regarding the biochemical makeup of microorganisms and thus of the biomolecules of which they are composed [21]. Multivariate analyses are useful for handling large datasets and to make spectral analysis viable. Spectral analysis included tentative assignment of bands based on the reported MIR absorption frequencies characteristic of the agents that represented the tested microorganisms (Bt, Ec, and Se) and measurement of reflected intensities for applying multivariate analysis [22, 23].

Overall, 836 different MIR experiments were conducted. However, only 245 experiments are reported for simplicity. Each experiment consisted of 15 replicate spectral measurements for each bacterium for each of the substrates studied. Previous studies have used the single spectrum concept to represent replicate spectral dispersions. This method is acceptable because it yields meaningful and concrete results [24, 25]. MIR spectra of the three strains of bacteria (Bt, Ec, and Se) are illustrated in Fig. 3. It was difficult to differentiate the different classes of the various tested surfaces based on the raw MIR spectra data due to the high degree of band overlapping. To overcome this problem, several pretreatments were used. The first pretreatment was spectral normalization.

Each bacterial species has a unique IR fingerprint spectrum due to the stretching and bending vibrations of its molecular bonds or protein functional groups (including nucleic acids, lipids, sugars, and lipopolysaccharides) [26], as illustrated by the reference spectra presented in Fig. 4. Reference spectra were obtained in absorption mode using a bench microspectrometer model IFS66/v/S (Bruker Optics, Billerica, MA).

Cell surface characteristics vary between the gram-positive and gram-negative bacteria. Gram-positive bacteria have a thicker and more rigid peptidoglycan layer that makes up to 40% to 80% more of the cell wall (by weight) than in gram-negative bacteria. In addition, gram-positive bacteria contain teichoic acids that are covalently bound to peptidoglycan. In contrast, gram-negative cells do not contain teichoic acids, but contain lipoproteins that are covalently bound to the peptidoglycan in the cell walls. Gram-
negative bacteria have an outer membrane outside the peptidoglycan layer that contains phospholipids in the inner layer and lipopolysaccharides in the outer layer [27].

![Figure 3. Normalized QCL spectra of bacterial suspensions.](image)

Amino acid composition of peptide chains in the Ec gram-negative bacterium consists of D-alanine, D-glutamic acid, and meso-diaminopimelic acid. In contrast, the Se gram-positive bacterium consists of L-alanine, D-glutamine, L-lysine, and D-alanine [28]. MIR spectra of intact bacterial cells are generally complex with broad peaks due to the overlaid contributions of all the biomolecules in the bacterial cell (Fig. 4.4) [26].

Following the recommendations of Naumann [14] and Naumann et al., [15,28] the IR spectra should be analyzed for bending vibrations between 1200 and 1500 cm\(^{-1}\) to identify fatty acids, proteins, and phosphate-carrying compounds in bacteria. The region from 900 to 1200 cm\(^{-1}\) contains carbohydrate bands that result from microbial cell walls. In addition, the region between 700 and 900 cm\(^{-1}\) is a fingerprint region that contains weak but unique vibrations that are characteristic of specific bacteria [29]. Tentative functional group assignments in these regions and their associations with the major vibration bands in the MIR spectra of each bacterium are depicted in Table 1.
Figure 4. Reference MIR spectra for *Bacillus thuringiensis* (Bt), *Escherichia coli* (Ec), and *Staphylococcus epidermidis* (Se) across the studied spectral region.

The spectral information for 245 MIR spectra (and replicas) of bacteria and substrates was subjected to several preprocessing steps to reduce the number of variables, maintaining the variance between low bacterial classes.

Chemometrics models were constructed for which the datasets with many variables could be simplified by performing data reduction, making results more easily interpretable in terms of multivariate analyses models.

The PCA models were represented in terms of score plots and well-defined separations between the classes of different bacteria on various substrates were not obtained. However, the PCA models represented most of the data variability.
Table 1. Tentative band assignments used for bacterial identification.

<table>
<thead>
<tr>
<th>Bt</th>
<th>Se</th>
<th>Ec</th>
<th>Tentative band assignments [15, 27, 28–30]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1426</td>
<td>1427</td>
<td></td>
<td>Symmetric stretching vibrations of -COO- functional groups of amino acid side chains or free fatty acids.</td>
</tr>
<tr>
<td>1411</td>
<td></td>
<td></td>
<td>Glutamic acid, CO₂⁻ asymmetric and symmetric stretching</td>
</tr>
<tr>
<td>1403</td>
<td>1404</td>
<td>1402</td>
<td>Symmetric stretching vibrations of -COO- functional groups of amino acid side chains or free fatty acids.</td>
</tr>
<tr>
<td>1392</td>
<td>1390</td>
<td></td>
<td>C-H bending, -CH₃ stretch in fatty acids</td>
</tr>
<tr>
<td>1387</td>
<td></td>
<td></td>
<td>COO stretching, Amines C-N CH₂ and CH₃ bending modes of lipids and proteins, stretching amides II</td>
</tr>
<tr>
<td>1377</td>
<td>1372</td>
<td></td>
<td>CH₃ symmetric bending</td>
</tr>
<tr>
<td>1364</td>
<td></td>
<td></td>
<td>Dipicolinic acid</td>
</tr>
<tr>
<td>1317</td>
<td>1318</td>
<td></td>
<td>Amide III protein</td>
</tr>
<tr>
<td>1306</td>
<td>1310</td>
<td></td>
<td>CH₂ and CH₃ bending modes of lipids and proteins</td>
</tr>
<tr>
<td>1294</td>
<td>1295</td>
<td>1293</td>
<td>Amide III</td>
</tr>
<tr>
<td>1249</td>
<td></td>
<td></td>
<td>Pectin -cellulose</td>
</tr>
<tr>
<td>1240</td>
<td>1241</td>
<td></td>
<td>RNA PO₂⁻ asymmetric stretching</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bt</th>
<th>Se</th>
<th>Ec</th>
<th>Tentative band assignments [15, 27, 28–30]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1087</td>
<td>1088</td>
<td></td>
<td>DNA PO₂⁻ symmetric stretching</td>
</tr>
<tr>
<td>1082</td>
<td>1073</td>
<td></td>
<td>RNA PO₂⁻ symmetric stretching</td>
</tr>
<tr>
<td>1059</td>
<td>1031</td>
<td></td>
<td>Ribose C-O stretching, carbohydrates, RNA Ribose C-O stretching</td>
</tr>
<tr>
<td>1010</td>
<td></td>
<td></td>
<td>DNA ribose C-O stretching and RNA ribose stretching</td>
</tr>
<tr>
<td>992</td>
<td>995</td>
<td></td>
<td>Dipicolinic acid, phenylalanine</td>
</tr>
<tr>
<td>968</td>
<td>964</td>
<td>962</td>
<td>RNA Uracil Ring stretching; uracil ring bending</td>
</tr>
<tr>
<td>968</td>
<td>944</td>
<td></td>
<td>DNA Ribose phosphate skeletal motion</td>
</tr>
<tr>
<td>929</td>
<td>930</td>
<td>931</td>
<td>C-OH out of plane</td>
</tr>
<tr>
<td>915</td>
<td>917</td>
<td>915</td>
<td>DNA Ribose Phosphate skeletal motion</td>
</tr>
<tr>
<td>907</td>
<td>907</td>
<td></td>
<td>Pectin</td>
</tr>
<tr>
<td>891</td>
<td>887</td>
<td>894</td>
<td>Aromatic ring vibrations of phenylalanine, tyrosine, tryptophan and the various nucleotides.</td>
</tr>
</tbody>
</table>
Representative MIR-QCL spectra for each bacterium after deposition on SS substrate are shown in Fig. 5. Based on these spectra, the bacteria analyzed by QCL in reflectance mode resulted in unique spectral signals in the MIR region. These spectral signals were clearly observed on SS substrate, particularly in the fingerprint region. In addition, similar results were obtained for the others substrates (CB, TB, wood, and glass).

<table>
<thead>
<tr>
<th>1177</th>
<th>1171</th>
<th>1168</th>
<th>1167</th>
<th>1151</th>
<th>1104</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amines</td>
<td>C-N,</td>
<td>Polysaccharides</td>
<td>RNA Ribose stretching</td>
<td>C-O</td>
<td></td>
</tr>
<tr>
<td>1177</td>
<td>1171</td>
<td>1168</td>
<td>1167</td>
<td>1151</td>
<td>1104</td>
</tr>
<tr>
<td>Polysaccharides</td>
<td>RNA Ribose stretching</td>
<td>C-O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>878</td>
<td>875</td>
<td>873</td>
<td>871</td>
<td>867</td>
<td>866</td>
</tr>
<tr>
<td>862</td>
<td>862</td>
<td>862</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>854</td>
<td>849</td>
<td>851</td>
<td>NH$_2$ wagging, tyrosine</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 5. The MIR spectra of *Bacillus thuringiensis* (Bt), *Escherichia coli* (Ec), and *Staphylococcus epidermidis* (Se) at room temperature following deposition on a SS substrate.

The classification between groups of bacteria on different substrates is shown in Table 2, in which the bold values represent the percentages of predicted discrimination within a group that were correctly classified. Other PCA models were built by selecting 10 to 15 MIR spectra of samples from each bacterium on each substrate (CB, TB, wood, glass,
and SS, for 225 MIR spectra). Next, these spectra were preprocessed based on their first derivative (order: 2, window: 15 pt.) and mean centered for all substrates. In addition, to obtain optimal results, standard normal variate preprocessing was used for the data that were acquired for wood substrates. The data from Bt, Ec, and Se were run together in the PCA models, and the variance that was described by each PC was examined for each substrate.

**Table 2.** Classification between groups of bacteria (Bt) *Bacillus thuringiensis*, (Ec) *Escherichia coli*, and (Se) *Staphylococcus epidermidis* on different substrates.

<table>
<thead>
<tr>
<th>Actual discrimination</th>
<th>Group size</th>
<th>Predicted discrimination</th>
<th>Substrates</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bt</strong></td>
<td>75</td>
<td>45 (60.0%)</td>
<td>18 (24.0%)</td>
</tr>
<tr>
<td><strong>Ec</strong></td>
<td>77</td>
<td>3 (3.89%)</td>
<td>64 (83.1%)</td>
</tr>
<tr>
<td><strong>Se</strong></td>
<td>75</td>
<td>14 (18.7%)</td>
<td>3 (4.0%)</td>
</tr>
<tr>
<td><strong>Substrates</strong></td>
<td>18</td>
<td>3 (16.7%)</td>
<td>4 (22.2%)</td>
</tr>
</tbody>
</table>

Percentage of cases correctly classified: **69.4%**

Scores plot (PC-3 versus PC-1) for CB is shown in Fig. 6 (a). A poor separation between the different bacterium datasets is shown. PC-1 (52.84% variance) was correlated to the differences between the three bacteria. Overall, 60% of the Bt samples were classified as Ec, while 93% and 100% of the Ec and Se were correctly classified, respectively. Scores plot (PC-2 versus PC-1) on glass substrates indicated good separation between the Bt samples [see Fig. 6 (b)]. However, the Ec and Se samples were 30% closer together, because they resulted in similar spectra when placed on glass.

This trend was also investigated for scores plot of PCs on TB (Fig. 7). The scores plots did not indicate class separation between the three types of bacteria on TB substrates [PC-2 versus PC-1, Fig. 7 (a)]. However, these plots only represent portions of the data variance (14.2% and 52.7%, respectively). For example, PC-3 versus PC-2 [Fig. 4.7 (b)] scores plot accounted for little variance (10.2% and 14.2%, respectively).
Figure 6. PCA for the QCL spectra of *Bacillus thuringiensis* (Bt), *Escherichia coli* (Ec), and *Staphylococcus epidermidis* (Se) deposited on (a) CB and (b) glass.

Figure 7. PCA for *Bacillus thuringiensis* (Bt), *Escherichia coli* (Ec), and *Staphylococcus epidermidis* (Se) on TB for (a) PC-2 versus PC-1 and (b) PC-3 versus PC-2.

The SS substrates resulted in the largest discrimination among bacterial species studied. However, scores plot (PC-2 13.9% versus PC-1 44.7%) indicated no significant variance between the three types of bacteria. Se samples (40.0%) were classified near Bt and Ec samples (94%) on the substrates. These results suggest that the bacteria were differentiated by their cell wall characteristics [gram-positive and gram-negative, see Fig. 4.8 (a) and Table 1]. Scores plots for the bacteria deposited on wood substrates required
a larger number of PCs (PC-3 5.65%, PC-4 3.17%, and PC-5 2.93%) for classification. In addition, the Ec samples were well classified [100%, as displayed in Fig. 4.8 (b); PC-2 10.7% versus PC-1 63.4%], whereas the Se samples (53.3%) were incorrectly characterized as Bt. Next, Bt samples were confused as Ec species (74.0%). These results in turn indicate that PCA can adequately discriminate between these three types of bacteria. However, PCA, which is frequently used as an unsupervised classification method, is not statistically powerful enough to achieve class separation, because it is not effective when "within-group" variations are larger than "between-group" variations. In this case, the use of supervised classification methods, such as PLSDA, must be considered.

![Figure 8](image_url)

**Figure 8.** The PCA for QCL spectra of *Bacillus thuringiensis* (Bt), *Escherichia coli* (Ec), and *Staphylococcus epidermidis* (Se) deposited on (a) SS and (b) wood.

The PLS-DA was used as a chemometrics tool and as a classification method for differentiating between bacterial species (*Bt*, *Ec*, and *Se*) on five different substrates. In PLS-DA, sensitivity is defined as the estimated experimental percentage of correctly classified samples. In addition, specificity is defined as the estimated experimental percentage of the samples that are rejected by the other classes in the model. Thus, a perfect class model has sensitivity and specificity values of 100% [31]. Fifteen MIR sample spectra for each bacterium and surface were analyzed (225 MIR total spectra).
Figure 9. Partial least squares-discriminant analysis (PLS-DA) plots for discriminating bacteria on CB: (a) cross-validation (CV) and prediction (PRED) of *Bacillus thuringiensis* (Bt) on CB; (b) CV and PRED of *Staphylococcus epidermidis* (Se) on CB; (c) CV and PRED of *Escherichia coli* (Ec) on CB; (d) percent variation accounted for by each latent variable (LV) used in the model for each specimen on CB.
Figure 10. PLS-DA plots for discriminating bacteria on luggage: (a) CV and PRED of *Escherichia coli* (*Ec*) on TB; (b) CV and PRED of *Bacillus thuringiensis* (*Bt*) on TB; (c) CV and PRED of *Staphylococcus epidermidis* (*Se*) on TB; (d) percentage variation accounted for by each LV used in the model for each specimen on TB.
Figure 11. PLS-DA plots for discriminating bacteria on SS: (a) CV and PRED of Bacillus thuringiensis (Bt) on SS; (b) CV and PRED of Escherichia coli (Ec) on SS; (c) CV and PRED of Staphylococcus epidermidis (Se) on SS; (d) percentage variation accounted for by each LV used in the model for each specimen on SS.
Vibrational information was organized into two groups. Approximately 75% of the sample spectra were randomly selected as training sets for the calibration and cross-validation (CV) models. The other 25% of the spectra were used as an external test set. To build a robust model, the spectral data were limited to the spectral ranges: 848 to 1012 cm\(^{-1}\), 1022 to 1170 cm\(^{-1}\), and 1173 to 1400 cm\(^{-1}\). Next, these data were pretreated by smoothing (order: 0; window: 15 pts.) before using the first-derivative (order: 2, window: 15 pts.) preprocessing technique. A cross-validation procedure was performed by using venetian blinds with 10 splits. This procedure was used to build a classification model for 90% of the spectra. Then, the remaining 10% were assessed to determine the accuracy of the models. Smoothing and first-derivative preprocessing improved the visualization of the spectra of the bacteria. The discrimination models for each bacterium and its surface are shown in Figs. 9 to 11, which illustrate the predicted cross-validated classes for each sample (as shown in the PLS-DA plots).

The outlier data were determined by the evaluation of residuals versus Hotelling's T\(^2\) plot on PLS Toolbox™ software for MATLAB®. According to the plot, the data with both high-residual variance and high leverage are outliers and can distort the model. Five percent of the calibration data were detected as outliers in calibration and cross-validation process and were removed from the analysis.

The discriminated bacterium class was located at approximately 1, while the other bacteria classes were located at approximately 0. The dashed line denotes the threshold parameter. These PLS-DA models were evaluated with several statistical parameters, including sensitivity, specificity, classification error of calibration, cross-validation (CV), and prediction (PRED) of the lowest latent variable. These values are shown in Table 3.

Based on these results, all the bacteria classes were successfully classified with only four latent variables. Sensitivity and specificity values of 100% for Bt, Ec, and Se were obtained for CB, TB, SS, wood, and glass. The root mean square error of cross-validation (RMSECV) and root mean square error of prediction (RMSEP) values were below 0.30, 0.35, 0.25, 0.30, and 0.40 for all the bacteria on CB, SS, TB, wood, and glass, respectively. These statistical parameters confirm that the PLS-DA performed significantly better than PCA in discriminating between bacterial strains on the substrates investigated.
Table 3. The PLS-DA summary for discriminating bacteria on substrates.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Stainless steel</th>
<th>Travel baggage</th>
<th>Cardboard</th>
<th>Wood</th>
<th>Glass</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bt</td>
<td>Ec</td>
<td>Se</td>
<td>Bt</td>
<td>Ec</td>
</tr>
<tr>
<td>Sensitivity (CV)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Specificity (CV)</td>
<td>100</td>
<td>100</td>
<td>96</td>
<td>100</td>
<td>96</td>
</tr>
<tr>
<td>Specificity (PRED)</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Specificity (PRED)</td>
<td>100</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>88</td>
</tr>
<tr>
<td>Class. Err (PRED)</td>
<td>0.00</td>
<td>0.00</td>
<td>0.13</td>
<td>0.00</td>
<td>0.06</td>
</tr>
<tr>
<td>Num. LVs</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>RMSEC</td>
<td>0.32</td>
<td>0.19</td>
<td>0.25</td>
<td>0.22</td>
<td>0.25</td>
</tr>
<tr>
<td>RMSEP</td>
<td>0.28</td>
<td>0.20</td>
<td>0.24</td>
<td>0.14</td>
<td>0.19</td>
</tr>
</tbody>
</table>

*Bacillus thuringiensis* (Bt), *Escherichia coli* (Ec), and *Staphylococcus epidermidis* (Se).

CONCLUSIONS

This study was aimed to detect and discriminate between bacteria cells on different substrates using QCL spectroscopy. The MIR diffuse reflectance spectroscopy was used on the range of 840 to 1440 cm⁻¹. In addition, PCA and PLS-DA classification models were built. Although the spectra appear very similar, slight differences occurred and revealed spectral variations. The obtained results demonstrated that the use of MIR spectroscopy (840 to 1440 cm⁻¹) with multivariate PCA and PLS-DA was useful for discriminating between bacterial strains (Bt, Ec, and Se) on various substrates. In addition, QCL spectroscopy was highly suitable for detecting microorganisms on different substrates, providing a quick response and analysis. Furthermore, this new technology can potentially be used to identify biological contaminants on surfaces and provides fast and accurate analyses for security and quality control purposes when nondestructive analytical methods are preferred. Substrate porosity clearly plays an important role in diminishing the number of microorganisms available to reflect MIR light on the substrate surface. This aspect needs to be addressed in detail in the further studies.

In summary, QCL spectroscopy was successfully used to identify biological samples in complex matrices by using their characteristic fingerprint spectrum in the MIR.
Differentiation and discrimination of microorganisms from the acquired spectra were achieved by using chemometrics tools such as PCA and PLS-DA. Overall, PLS-DA performed significantly better than PCA in the analyses of bacteria in this study.

References


PROJECT 2: DETECTION OF HIGHLY ENERGETIC MATERIALS ON NON-REFLECTIVE SUBSTRATES USING QUANTUM CASCADE LASER SPECTROSCOPY

ABSTRACT
A quantum cascade laser spectrometer (QCLS) was used to obtain mid-infrared spectra of highly energetic materials (HEMs) deposited on non-ideal, low reflectivity substrates such as travel bags, cardboard and wood. Various deposition methods, including spin coating, sample smearing, partial immersion and spray deposition, were used to prepare the standards and samples used in this study. The HEMs used were the nitroaromatic explosive 2,4,6-trinitrotoluene (TNT), the aliphatic nitrate ester pentaerythritol tetranitrate (PETN), and the aliphatic nitramine 1,3,5-trinitroperhydro-1,3,5-triazine (RDX). Hit quality index (HQI) values were calculated from a spectral library and used as a first identification method through correlation coefficients. In addition, two chemometrics algorithms were applied to analyze the spectral characteristics recorded using QCL spectroscopy: partial least squares (PLS) regression was used to find the best correlation between the infrared signals and the surface concentrations of the samples, and PLS combined with discriminant analysis (PLS-DA) was used to discriminate, classify and identify similarities in the spectral data sets. Several multivariate analysis preprocessing steps were used to analyze the obtained infrared spectra of HEMs deposited as contaminants on the target substrates. The results clearly demonstrate that the infrared vibrational method used in this study can be useful for the detection of HEMs on real-world low reflectivity substrates.

Index Headings: quantum cascade laser spectroscopy; highly energetic materials; non-reflective substrates; chemometrics; partial least squares analysis; discriminant analysis.

INTRODUCTION
The detection and identification of hazardous chemical compounds, such as highly energetic materials (HEMs), is of great interest to defense and security agencies and is also of importance in forensic applications. When detonated, HEMs have the potential to destroy public and private buildings and may also jeopardize the lives of civilians, first
responders and law enforcement employees. Over the past few years, vibrational spectroscopy techniques, such as Raman and infrared spectroscopies, have frequently been used to deter possible terrorist threats by providing the basis for the required countermeasures to prevent explosive terrorist events.\textsuperscript{1-5}

Mid-infrared (MIR) electromagnetic radiation is located in the spectral region from approximately \(350\) to \(4000\) cm\(^{-1}\). In this region, molecules have characteristic vibrational energy states that can be excited upon interaction with photons from an appropriate excitation source, enabling the detection of trace amounts of compounds by measuring the intensities of absorbed, reflected or transmitted MIR.\textsuperscript{6} Fourier transform infrared (FT-IR) spectroscopy has been extensively used in both active and passive modalities in the MIR in defense and security applications.\textsuperscript{1,2,7-10} Active mode FT-IR spectroscopy has been used for the post-blast detection of energetic materials using infrared radiation produced from both globar and synchrotron sources. Reports have validated FT-IR spectroscopy as a useful tool for forensic science applications.\textsuperscript{7,8} Emission (passive mode) and absorption MIR spectroscopies have recently been used as vibrational techniques for the standoff detection of explosives and other chemical agents deposited on metallic surfaces.\textsuperscript{2,9,10}

The need to develop more powerful IR sources that enable detection at longer distances when a target is deposited on substrates at trace levels suggests the use of collimated, coherent, and polarized sources. These sources were first developed in 1994 at Bell Labs with the invention of quantum cascade lasers (QCLs).\textsuperscript{11} A QCL is a unipolar semiconductor injection laser based on sub-interband transitions in a multiple quantum-well heterostructure. As a semiconductor laser that has the ability to produce varying wavelengths and to operate at various temperatures, this type of laser has various advantages over other types of lasers.\textsuperscript{11,12} QCLs are capable of producing from a few tens to hundreds of milliwatts of continuous mode or pulsed power under ambient conditions, are commercially available and have enabled the development of a ruggedized system for the detection of hazardous chemical compounds. Recent developments in QCL technology include size reduction, which has enabled the transition from tabletop laboratory instruments to easy-to-handle, portable, and small instrument designs that can be used by first responders and military personnel outside the
confinement of a sample compartment. Moreover, the increase in output power has enabled the use of QCL-based spectrometers in long-distance (range) applications, making the detection of chemical and biological threat agents possible from tens of meters from the source.\textsuperscript{13} Furthermore, QCLs can be operated in field conditions, allowing for the sensitive detection of HEMs such as the homemade explosive triacetone triperoxide (TATP), the aliphatic nitrate ester pentaerythritol tetranitrate (PETN), the aliphatic nitramine 1,3,5-trinitroperhydro-1,3,5-triazine (RDX) and the nitroaromatic 2,4,6-trinitrotoluene (TNT) in the vapor phase using photoacoustic spectroscopy.\textsuperscript{14-16} The detection of TATP and TNT in the vapor phase has also been achieved using infrared absorption spectroscopy with satisfactory results.\textsuperscript{3,17} Moreover, the use of QCL sources has been useful for the standoff detection of HEMs deposited on surfaces using photoacoustic and traditional infrared absorption spectroscopies.\textsuperscript{18-22} Thundat’s group recently demonstrated that nanomechanical infrared spectroscopy provides high selectivity for the detection of TNT, RDX and PETN without the use of chemoselective interfaces by measuring the photothermal effect of the adsorbed molecules on a thermally sensitive microcantilever.\textsuperscript{23} However, the majority of previous investigations focused on the detection of HEMs deposited on nearly ideal, highly reflective substrates (such as metallic surfaces), and there are few published reports on the effects of non-ideal, low reflectivity substrates on the spectra of the analyzed targets.

In this study, non-contact detection experiments using QCL spectroscopy (QCLS) were performed. The experiments were conducted using an active mode QCL source to excite the MIR molecular vibrations from the investigated HEMs. Hit quality indices (HQIs) were calculated for the spectra of analytes recorded from complex, highly interfering surfaces. In this detection modality, two chemometrics routines were applied to analyze the spectral characteristics recorded using QCLS: partial least squares (PLS), which assisted in finding the best correlation between the IR signals and the analyte surface concentrations, and partial least squares coupled with discriminant analysis (PLS-DA), which was used to discriminate, classify and identity similarities between the spectral data. Several preprocessing steps were used in the multivariate analysis protocols. The results indicate that the investigated infrared vibrational technique is useful for the
detection of HEMs on the types of non-ideal, low reflectivity substrates studied in this work.

MATERIALS AND METHODS

Reagents and materials. The reagents and materials used in this study included HEMs, solvents and substrates. TNT was acquired as a crystalline solid (99% min.; 30% min. water content; standard grade) from Chem Service, Inc. (West Chester, PA, USA); PETN and RDX were synthesized and purified in our laboratory according to the method described by Ledgard. Methanol (99.9%, HPLC grade), dichloromethane (CH$_2$Cl$_2$, HPLC grade), and acetone (99.5%, GC grade) were purchased from Sigma-Aldrich (Milwaukee, WI, USA) and were used as solvents to deposit the HEM samples with various surface concentrations onto the test surfaces. The substrates used were travel baggage (TB; black polyester), cardboard (CB; corrugated single-wall) and wood (grade C-C, plugged soft 3-ply spruce plywood). Aluminum (Al) plates were used as high reflectivity reference substrates. The substrates were purchased from local suppliers and were cut into square pieces of approximately 31 mm x 31 mm. These substrates were gently cleaned using acetone or methanol and were stored in desiccators or Dry Keeper™ cabinets to stabilize their weights by controlling their moisture contents. Attempts to stabilize the weights of the substrates were only successful for the TB substrates. Because the humidity levels in the laboratory varied on a daily basis, the wood and CB substrates did not acquire stable weights, and deposition methods that were not dependent on the weight of the substrate had to be used. These deposition methods are explained in the following section.

Sample preparation. Sample preparation is a fundamental step in this type of experiment. To obtain calibration models that provided good predictions for the validation samples, a method that allowed for the proper transfer of the target HEMs to the tested substrates had to be used. This was required to achieve reliable and consistent analyte surface concentrations (mass/area), thereby allowing validation of the experimental methodology. Various deposition methods were tested, including sample smearing, spin coating, spray deposition, and partial immersion. A brief discussion of each of these
methods follows to provide information regarding the experimental protocols, focusing on the advantages and disadvantages of each deposition method.

Deposition by sample smearing required the use of a fixed volume (~ μL) of the HEM solution in an appropriate solvent (e.g., acetone) transferred to one side of a substrate using micropipettes. This solution was then smeared over the entire surface using a Teflon stub. This method resulted in a quantitative transfer of the analyte, but the homogeneity of the deposition obtained in this study was typically poor because the surfaces of the employed substrates were non-flat and porous. There was also a slight sample loss during the process, and a relatively high degree of technical skill was required to implement the sample smearing approach in a consistent and reproducible manner. In an attempt to obtain more homogeneous depositions with less sample loss, spin coating and spray deposition methods were also tested. During spin coating deposition, an excess amount of solution was placed on the substrate, which was then rotated at high speed to spread the fluid through centrifugal force. The spin coating deposition method resulted in high homogeneity for flat substrates (such as metals) but not for porous substrates such as TB, CB and wood. For the spray deposition method, a solution of analyte in a volatile solvent was sprayed onto the substrate using an airbrush gun, and intermediate homogeneities were achieved. Substrate preparation was required for both of the aforementioned methods, in which the substrates had a stable weight for determining the amount of explosive deposited. This was accomplished by measuring the difference in the substrate weight before and after deposition of the analyte. Because of this limitation, a different method was tested in which the amount of explosive deposited on the substrate was independent of the weight of the substrate. During partial immersion deposition in a borosilicate glass Petri dish, the desired amount of HEM was dissolved in acetone (2 mL), and then the substrates were partially immersed in this solution (only one side of the substrate was immersed). This technique required weighing the Petri dish that contained the HEM before adding the acetone-HEM solution, after the immersion, and after complete evaporation of the solvent. The amount of explosive deposited on the substrate was determined from the difference in the weight of the Petri dish. When using this method, the sample loss was negligible, a good homogeneity was attained, and no substrate preparation was required. A spectrophotometric (UV-Vis) technique was used
to validate this method. Finally, the obtained nominal surface concentrations were 1 – 15 μg/cm² on the tested substrates.

**Setup.** Fig. 1 shows a diagram that illustrates the experimental setup, including (a) sample preparation using the aerosol spray and partial immersion deposition methods, (b) non-contact detection of the target chemicals and (c) chemometrics multivariate analyses. The detection of PETN, RDX and TNT deposited on non-ideal, low reflectivity substrates was performed using a LaserScan™ dispersive spectroscopy system acquired from Block Engineering (Marlborough, MA, USA). The specifications for the spectroscopy system and for the laser source used were nearly continuous wavenumber coverage (1000-1600 cm⁻¹), laser pulse modulation of 200 kHz, beam divergence of < 5 mrad, beam size of 2 x 4 mm collimated light and a thermoelectrically (TE) cooled MCT internal detector. The QCLS measurements were obtained in the diffuse reflectance mode.

**Fig. 1.** Experimental setup. (a) Sample preparation: HEM samples deposited on substrates. (b) In situ QCL spectral measurements. (c) Multivariate statistical analyses.
MIR spectra were recorded at a distance of 6 in. from the substrates with 2 co-additions and a 4 cm$^{-1}$ resolution. Eleven surface concentrations for each HEM, ranging from 1 to 15 μg/cm$^2$, were prepared. Four spectra per sample were measured, resulting in 396 independent measurements. The spectra were stored in Thermo-Galactic™ SPC format and then analyzed using the PLS and PLS-DA chemometrics models with PLS Toolbox™ version 6.5 (Eigenvector Research Inc., Wenatchee, WA, USA) for MATLAB™ (The MathWorks, Inc. Natick, MA, USA). In addition, HQI values, which are a measure of the similarity between the measured spectrum and reference (library) spectrum, were also calculated.

RESULTS

Spectral analysis. MIR vibrational spectra of TNT, RDX and PETN deposited on TB, CB and wood substrates were obtained using a LaserScan™ spectrometer. The spectra were recorded over the spectral region of 1000-1600 cm$^{-1}$, where the symmetric and asymmetric nitro group vibrations of the HEMs occur. Fig. 2 presents the MIR reflectance spectra of TNT, PETN and RDX deposited on Al, TB, CB and wood. The spectra on Al are provided as a reference to assist in identifying the obtained MIR vibrational bands of the HEMs after being deposited on non-ideal substrates.

As shown in Fig. 2a, characteristic vibrational signals from the HEMs clearly evidence the highly reflective nature of the Al substrate. Some of the vibrational bands that were tentatively assigned to TNT were 1024 cm$^{-1}$ (CH$_3$ deformation), 1086 cm$^{-1}$ (C–H ring in-plane bending), 1350 cm$^{-1}$ (symmetric stretching of nitro groups) and 1551 cm$^{-1}$ (asymmetric NO$_2$ stretching).$^{25}$ For PETN, some of the important signatures appeared at 1003 cm$^{-1}$ (CO stretching), 1038 cm$^{-1}$ (NO$_2$ rocking), 1272 cm$^{-1}$ (ONO$_2$ rocking), 1285 cm$^{-1}$ (NO$_2$ stretching) and 1306 cm$^{-1}$ (NO$_2$ rocking).$^{26}$ Finally, important markers for RDX were detected at 997 cm$^{-1}$ (N-N and ring stretching), 1220 cm$^{-1}$ (C-N stretching), 1270 cm$^{-1}$ (NO$_2$ stretching), 1310 cm$^{-1}$ (N-N stretching), 1420 and 1445 cm$^{-1}$ (H-C-N asymmetric bending) and 1570 cm$^{-1}$ (N-O asymmetric stretching).$^{27,28}$
Typical spectra of the HEMs deposited on CB, wood and TB are shown in Figs. 2b-2d. In addition to the HEM signatures, additional vibrational bands from the substrates were also observed. However, providing details about the signals from the substrates that give rise to these infrared bands is beyond the scope of this contribution. The focus of this work is the analysis of the interferences produced by the substrate contributions and the effects of these interferences on the ability to detect and discriminate HEMs deposited on the investigated substrates. The spectral profiles of each HEM were similar when deposited on CB and wood (Figs. 2b and 2c). The salient MIR signatures of the HEMs on CB and wood were the NO$_2$ bands at approximately 1270 cm$^{-1}$ for the aliphatic explosives PETN and RDX and at 1350 cm$^{-1}$ for the aromatic explosive TNT. The infrared bands of the HEMs on TB are shown in Fig. 2d and in Fig. 3.

The vibrational frequencies observed during the detection of the HEMs are in agreement with the above discussion. The MIR spectral profiles of bands due to the NO$_2$
group from the HEMs deposited on Al differed from those of the HEMs deposited on non-ideal substrates. In all cases, when the HEMs were deposited on highly IR-absorbing substrates, the spectral signatures were inverted compared with those deposited on Al substrates. This behavior was also reported by Fuchs et al., who deposited TNT on an auto body piece covered with car paint.\textsuperscript{19} The reflectance of a material is a measure of its capacity to reflect incident energy, which is determined by two components: the fraction of reflected radiation on the surface of the material and the total energy of the incident radiation at any wavenumber. Fig. 2 presents the relative reflectance spectra of the HEMs when present as contamination residues on non-ideal substrates. The relative reflectance spectra were calculated by dividing the IR light reflected from the samples by the IR light reflected from the reference, a highly reflective Al substrate. When the HEMs were deposited on ideal surfaces, such as Al, the intensity of the reflected light was lower in the spectral region where the HEMs have IR absorption bands, which generated “downward-looking” peaks when the relative reflectance spectrum was calculated (as shown in Fig. 2a). When the HEMs were deposited on low reflectivity, non-ideal substrates, such as TB, CB and wood, changes in the reflective optical properties of the substrate surface occurred such that the intensity of the light reflected from the HEM-substrate surface was higher in the spectral ranges in which the HEMs have IR absorption bands, generating inverted “upward-looking” peaks.

\textbf{Fig. 3.} QCL spectra of HEMs deposited on TB. The spectra are shown in two spectral regions: (a) 1240-1300 cm\textsuperscript{-1} and (b) 1340-1370 cm\textsuperscript{-1}. Spectra of substrates are also included.
**Target Identification.** After the samples were prepared, spectra were recorded and multivariate analyses were performed. First, a spectral search was applied to each spectrum with the purpose identifying the detected HEM. The simplest spectral search is based on the calculation of the hit quality index (HQI) values. The HQI is a numerical quantity that indicates the correlation between two spectra and has been widely used in spectroscopy to indicate the degree of spectral matching in library searches. In the protocol, a spectrum of an unknown sample is compared with all spectra of known samples in a library, and the best match is determined based on the calculated HQI values. Routines for calculating the HQI are available in most commercial portable spectrometers to facilitate the identification of unknown compounds in the field. HQI values can be calculated using various algorithms, but the two most commonly used are the Euclidean distance and spectral correlation algorithms. In the Euclidean distance algorithm, the HQI values are calculated from the square root of the sum of the squares of the difference between the vectors for the unknown spectrum and each library spectrum. In contrast, the spectral correlation algorithm utilizes the Pearson product-moment correlation coefficient, $r_{xy}$, which is a measure of the strength and direction of the linear relationship between two variables. This parameter is defined as the covariance of the variables divided by the product of their standard deviations. In this case, the spectral correlation algorithm is applied between two spectra: a reference (or library of spectra) and an unknown spectrum to be identified. The $r_{xy}$ values were calculated using Eq. 1:

$$r_{xy} = \frac{\sum x_i y_i - \frac{1}{n}(\Sigma x_i)(\Sigma y_i)}{\sqrt{\left(\sum x_i^2 - \frac{1}{n}(\Sigma x_i)^2\right)\left(\sum y_i^2 - \frac{1}{n}(\Sigma y_i)^2\right)}}^{1/2} \tag{1}$$

where $x$ and $y$ represent the spectral responses of the reference spectrum and of an unknown spectrum, respectively, measured at the $i^{th}$ wavenumber for a set of $n$ corresponding wavenumber points. In the present case, the HQI assumes values between +1 and -1. A HQI value of +1 is obtained when the spectral similarities are maximum (the unknown spectrum is identical to a library spectrum), -1 when the spectral similarities are inverted (the unknown spectrum is identical to a library spectrum but with peaks inverted with respect to the library spectrum) and 0 when there is no spectral similarity. Often, the values are rescaled to values between 0 and 1, where values...
between -1 and 0 are equated to zero and a value of 1 indicates maximum spectral similarity.\textsuperscript{8} Another possibility is to rescale the square of $r$ ($r^2$)\textsuperscript{32-33} to yield positive values between 0 and 1. However, in this investigation, the rescaling step was not applied such that the effects of the contributions of complex substrates on the reflectance spectra (analyte + substrate) could be evaluated when HQI algorithms were used to identify unknown spectra from reference spectra.

Table I presents the spectral correlation coefficients for unknown spectra on the various types of substrates tested. Before calculating the $r$ values, a spectral normalization was performed on the entire spectral range of 1000-1600 cm\textsuperscript{-1} for both the unknown (measured) and library spectra (measured on Al substrates). The normalization of a spectrum for library searching is a two-step process, as recommended by ASTM E2310-04 (2009): “Standard Guide for Use of Spectral Searching by Curve Matching Algorithms with Data Recorded Using MIR Spectroscopy”.\textsuperscript{29} First, the minimum spectral response value in the selected spectral range is subtracted from the entire spectral response in the same range. The resulting values are then scaled by dividing by the maximum value in that range. The net result is a spectrum in which the minimum intensity value is zero (0) and the maximum value is one (1). The values of the spectral correlation coefficients for the HEMs shown in Table I are severely influenced by the type of substrate used, which can be confirmed when metal substrates such as Al were used, resulting in high values of $r$. When non-metallic substrates were used, such as TB, CB and wood, low values of $r$ were obtained. It is generally acceptable to consider a spectrum of an unknown compound to be similar to one of the library when the spectral correlation coefficient is greater than $\sim 0.85$. Given this restriction, spectra measured on only Al substrates were correctly identified when QCLS coupled with spectral correlation algorithms were used. All of the spectral correlation coefficients for unknown spectra of HEMs deposited on CB, wood and TB were well below the minimum accepted $r$ value (0.85). In some cases, the $r$ values were negative (i.e., some unknown spectra had peaks inverted with respect to the library spectra). These results highlight two factors. First, low $r$ values were obtained when the QCL spectra of HEMs deposited on low reflectivity, complex substrates were measured. Second, negative $r$ values were obtained due to the effect of such substrates, resulting in spectra with inverted peaks.
**TABLE I.** Values of HQI for spectra of HEM deposited on various substrates.

<table>
<thead>
<tr>
<th></th>
<th>Spectra Queried</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HEM spectra on Al</td>
<td>HEM spectra on CB</td>
<td>HEM spectra on wood</td>
<td>HEM spectra on TB</td>
</tr>
<tr>
<td>library spectra</td>
<td>PETN</td>
<td>RDX</td>
<td>TNT</td>
<td>PETN</td>
</tr>
<tr>
<td>PETN</td>
<td>0.95</td>
<td>0.39</td>
<td>0.07</td>
<td>0.06</td>
</tr>
<tr>
<td>RDX</td>
<td>0.39</td>
<td>0.94</td>
<td>0.09</td>
<td>0.15</td>
</tr>
<tr>
<td>TNT</td>
<td>0.10</td>
<td>0.14</td>
<td>0.96</td>
<td>-0.03</td>
</tr>
</tbody>
</table>

Because the spectral correlation coefficients were not efficient for the identification and classification of HEMs when they were deposited on non-ideal, low reflectivity substrates (such as CB, TB, and wood) for direct field detection applications, multivariate analysis methods, such as PLS and PLS-DA, were applied for spectral identification, classification and quantification. PLS Toolbox™ version 6.5 for MATLAB™ was used to analyze the data. Fig. 1c presents a summary of the statistical treatment performed on the data. PLS-DA is one of the most widely used chemometrics tools, particularly when the goal is to discriminate, classify and identify spectral similarities in a multivariate data set. PLS-DA is a supervised pattern recognition method. A detailed explanation of how PLS-DA works and its numerous applications is not included in this paper. Excellent in-depth mathematical support and various applications in natural sciences and engineering are available in the literature.\textsuperscript{35-37} PLS-DA is a linear classification method that combines the properties of PLS regression with the discrimination power of a classification technique. PLS-DA is based on the PLS regression algorithm, which searches for latent variables with a maximum covariance between a descriptor matrix $\mathbf{X}$ and a response matrix $\mathbf{Y}$ (containing the membership of samples to the $G$ classes expressed with a binary code: 1 or 0). The primary advantage of PLS-DA is that the relevant sources of data variability are modeled by the so-called latent variables (LVs), which are linear combinations of the original variables, and consequently, allowing graphical visualization and understanding of the different data patterns and relations by LV scores and loadings. The scores represent the coordinates of the samples in the LV projection hyperspace.\textsuperscript{36,37}
The data were divided into two groups for each model: a calibration set and a prediction (test) set. The calibration set contained approximately 70% of the data, and the prediction set contained the remainder of the data. PLS-DA was applied to the spectral data to discriminate, classify or group all spectra by HEM type and to discriminate between clean substrates and the substrates with HEMs. An overview of the performance of PLS-DA on the spectral data (matrix $X$) is presented in Fig. 4. The first step was intended to arrange the spectral data in a matrix with dimensions of “$n \times m$” ($n$ rows and $m$ columns), where $n$ represents the number of samples used in the calibration set and $m$ represents the number of spectral variables in the study. In this work, the matrix dimensions were 93x3056 for the calibration set of each type of substrate tested. The second step was to apply the required preprocessing handling of the data. The purpose of preprocessing the data was to remove or to decrease the effect of the interferences from the background/substrate, thereby enhancing the vibrational signatures of the HEMs. PLS-DA was able to decompose the original matrix into two new matrices: one that has most of the relevant information containing the largest variability in the data and another that contains information that is not as relevant to the data, generally termed noise. The matrix that contains the information of interest has dimensions of “$n \times p$”, where $p$ is a new column matrix called the latent variable (LV), which is the product of the transformation from the original variables. LVs are orthogonal vectors that are linear combinations of the original data. In general, two or three LVs are sufficient to capture the majority of the variability in the data. In the LVs, each sample that was initially represented by an IR spectrum is now represented by only one value, termed the “score” of the sample. These scores describe how the samples are related to each other. Score plots are obtained by graphing LV2 vs. LV1. These plots provide important information regarding how different samples are related to each other. Finally, the original matrix with dimensions of 93x3056 for the calibration set of each substrate tested was reduced to a matrix with dimensions of 93x3 using three LVs.

In PLS-DA, the optimal number of LVs that provided the best classification for the calibration set was determined using cross validation procedures. The Venetian blinds procedure with 10 group splits was used to divide the calibration set into cross validation groups. Several preprocessing steps were applied to the data with the objective of
generating multivariate models capable of discriminating and clustering the spectra based on chemical similarities, i.e., PETN, RDX, TNT and clean substrates (TB, CB and wood). The preprocessing steps used were auto scaling (AS), mean centering (MC), standard normal variate (SNV), 1<sup>st</sup> / 2<sup>nd</sup> derivatives and their combinations. For all the models that were generated, those that performed the best were those that used the full spectral range (1000-1600 cm<sup>-1</sup>).

**Fig. 4.** Representation of multivariate analysis (PLS-DA) on spectral data used in this research to identify and classify HEMs deposited on non-ideal, non-reflective substrates.

For PLS-DA, classification parameters, such as sensitivity and specificity, derived from the confusion matrix from the calibration, cross validation and prediction sets were used to evaluate the performances of the classification models. The sensitivity (SEN) describes the model’s ability to correctly recognize samples that belong to that class, and the specificity ($\xi$) describes the model’s ability to reject samples of all other classes and can be calculated as follows:

$$\text{SEN} = \frac{TP}{(TP+FN)} ; \quad \xi = \frac{TN}{(TN+FP)}$$

(2)

where TP = true positive, TN = true negative, FP = false positive, and FN = false negative. When dealing with two classes A and B, A: positive (a specific HEM) and B: negative (the other HEM or the clean substrates), TP is the number of members of class A that have been correctly classified, FP (false positive) is the number of members of class B that
have been incorrectly predicted as class A, TN is the number of members of class B that have been correctly classified and the number of members of class A that have been incorrectly predicted as class B. Both SEN and $\xi$ take values between 0 and 1, where 1 is the desired result.\textsuperscript{35} Table II presents a summary of the results obtained from applying PLS-DA to the QCLS spectra of the HEMs on the investigated surfaces for the calibration, cross validation and prediction sets.

**Table II:** Summary of results of PLS-DA of QCLS spectra of HEMs on surfaces obtained for calibration, cross validation (10 groups split in venetian blinds) and prediction set.

<table>
<thead>
<tr>
<th></th>
<th>HEM spectra on TB</th>
<th>HEM Spectra on CB</th>
<th>HEM spectra on Wood</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PETN</td>
<td>RDX</td>
<td>TNT</td>
</tr>
<tr>
<td>SEN (Cal)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\xi$ (Cal)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SEN (CV)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\xi$ (CV)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>SEN (Pred)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>$\xi$ (Pred)</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Class. Err. (Cal)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Class. Err. (CV)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Class. Err. (Pred)</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>RMSEC</td>
<td>0.136</td>
<td>0.137</td>
<td>0.123</td>
</tr>
<tr>
<td>RMSECV</td>
<td>0.149</td>
<td>0.142</td>
<td>0.132</td>
</tr>
<tr>
<td>RMSEP</td>
<td>0.109</td>
<td>0.104</td>
<td>0.100</td>
</tr>
<tr>
<td>LV</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>Variance Captured (%)</td>
<td>87.8</td>
<td>87.8</td>
<td>87.8</td>
</tr>
</tbody>
</table>

The score plots, which allow visualization of the clustering of the spectral data, resulting from the PLS-DA runs are shown in Figs. 5. These plots demonstrate that the best results were obtained for the various generated models after the preprocessing steps mentioned above had been applied. For the multivariate analysis of the MIR vibrations of the HEMs deposited on TB and CB, taking the first derivative (15 pt.) and applying MC were sufficient to obtain the best chemometrics models. Using derivatives in spectroscopy facilitates the identification of the wavenumber location of the maxima/minima of poorly resolved spectral features in complex spectra. Taking derivatives can also be used as a
background correction technique to reduce the effect of spectral background interferences in quantitative analytical methods. MC was used to decrease the dimensionality of the spectral data.

The PLS-DA models resulting from the discrimination of each HEM from the others and from the neat substrate for TB are shown in Fig. 5. The class predictions of PETN, RDX, and TNT on TB from the cross validation are shown in Figs. 5a-5c. Four LVs were required to obtain the best multivariate classification model with SEN and $\xi$ equal to 1 (see Table II) for the HEMs tested from the calibration, cross validation and prediction data sets. The variance captured from matrix $\mathbf{Y}$ was 87.8%, which is sufficient for a good classification of the predicted spectra set on TB. For the multivariate (clustering) analysis of the HEMs on TB, a total of six LVs were necessary to capture 80% of the total variance in the spectral data. As shown in the score plot in Fig. 5d, two latent variables with 60% of the spectral variance were sufficient to obtain an excellent classification according to the type of HEM deposited and to discriminate from the TB substrate. In this model, spectra from the prediction set (RDX Test, TNT Test, and PETN Test) were well grouped according to chemical characteristics with spectra from the calibration sets.

Class predictions of PETN, RDX, and TNT on CB from the cross validation required four LVs were to obtain the best classification model allowing SEN and $\xi$ to equal 1 (see Table II) for the HEMs tested for the calibration, cross validation and prediction data sets. The variance captured from matrix $\mathbf{Y}$ was 91.5%, which is sufficient for a good classification of the predicted spectra set on CB. In the PLS-DA clustering analysis for the QCL spectra of the HEMs on CB, five LVs were required to capture 80% of the total variance in the spectral data using the first derivative (15 pt.) and MC as prepossessing steps. Two LVs corresponding to 63% of the variance were sufficient to obtain excellent spectral classification (according to the type of explosive deposited and to discriminate from the substrate. Using this model, spectra from the validation set (RDX Test; TNT Test; PETN Test) were grouped reasonably well with spectra from the calibration set according to vibrational signatures.

Class predictions of HEMs on wood from the cross validation required five LVs to obtain the best multivariate classification model that allowed SEN and $\xi$ to equal 1 (see Table II) for the HEMs tested from the calibration, cross validation and prediction data.
sets. The variance captured from matrix \( Y \) was 92.3\%, which is sufficient for a good classification of the prediction spectra set on wood. For the multivariate analysis of the MIR vibrations of the HEMs deposited on wood, preprocessing of the data using the first derivative (15 pts.), SNV transformation and MC was necessary to obtain the best chemometrics results. The first derivative was used to eliminate spectral differences on the baseline and to smooth the spectra. SNV transformation and MC were used to normalize and center the data. In the PLS-DA clustering model of the HEMs on wood, six LVs were required to capture 80\% of the total variance in the spectral data. Two LVs accounting for 61\% of the spectral variance were sufficient to obtain a good spectral classification according to the type of HEM deposited and to discriminate from the wood substrate. Spectra from the prediction set were precisely grouped with the spectra from the calibration set according to chemical similarities.

Figure 5. PLS-DA model for discrimination of HEM on TB: (a) class prediction for PETN; (b) class prediction for RDX; (c) class prediction for TNT; (d) scores plot of LV2 vs. LV1 for detection of PETN, RDX and TNT on TB. Preprocessing steps applied were: 1st deriv. (15 pt.) and MC. Threshold for discrimination and 95\% confidence level for clustering are represented with red dotted lines.
A more general and robust model based on IR spectral data for the three HEMs on the three substrate types was generated. In the multivariate analysis, preprocessing of the data required applying the second derivative (17 pts.), as well as using SNV transformation and MC. In this analysis, twelve LVs were required to capture 80% of the total variance in the spectral data. The first three LVs explained 47% of the spectral variance (Fig. 8b) according to the type of HEM. The PLS-DA model generated using the three HEMs deposited on three non-ideal, low reflectivity substrates was capable of differentiating substrates without HEMs (region A, Fig. 6a) from surfaces with HEMs (region B, Fig. 6a). In the model representation, there are a few outliers (substrate spectra classified as spectra of substrates with HEMs, and vice versa), which is a consequence of the plot representation using only 2 LVs. However, in general, the ambitious task of grouping HEMs deposited on non-ideal substrates based on spectral similarities and discriminating the spectral features of the HEMs from those of the substrates was largely achieved.

**Quantification of HEMs Using PLS Regressions.** PLS regressions were used to analyze the spectral data and to find the best correlation between the multivariate spectral information and the HEM surface concentrations. The main concept of PLS is to obtain the most information possible concentrated in the first few loading vectors or latent variables (LV). PLS regressions were used to generate models for all the HEMs deposited on the investigated substrates. In addition, the root mean square error of cross validation (RMSECV), root mean square error of prediction (RMSEP), coefficient of determination from cross validation (R²-CV) and coefficient of determination from prediction (R²-Pred) were calculated and used as indications of the quality of the obtained spectral correlations. Chemometrics models based on PLS regressions were obtained using the complete spectral range measured (1000 to 1600 cm⁻¹). The sample concentrations were the same as those in the PLS-DA analyses. MC was applied to each spectrum from the spectral data set as a preprocessing step. Fig. 7 shows some of the PLS regression plots obtained for the predicted concentration vs. measured concentrations for each HEM deposited on each substrate. Fig. 7a shows a PLS model for RDX on TB, Fig. 7b displays a PLS model representing TNT on CB, and Fig. 7c presents a PLS model for PETN on wood.
The best results for RMSECV, RMSEP, R\textsuperscript{2}-CV and R\textsuperscript{2}-Pred, including the number of LVs required in the PLS model, are shown in Table III. Eight LVs were required to obtain the best PLS models, resulting in determination coefficients higher than 0.99, with the exception of TNT on wood, in which 9 LVs were required. Higher values of RMSECV and RMSEP were found for TNT in the three substrates tested. These results can be attributed to the higher vapor pressure of the nitroaromatic HEM when compared to PETN and RDX, leading to mass losses via sample sublimation during the course of the experiments. Based on the obtained results, it can be concluded that PLS-based models were highly successful in correlating surface concentrations of HEMs on the investigated substrates.

**Fig. 6.** PLS-DA model for QCL spectra of PETN, RDX and TNT deposited on TB, CB and wood substrates. Preprocessing steps applied were 2nd deriv. (17 pts.) + SNV + MC: (a) 2D-Score plot using LV1 and LV2; (b) 3D-Score plot using LV1, LV2 and LV3. 95% confidence level for clustering is represented with red dotted line.
**Fig. 7.** PLS regression plots: predicted vs. measured surface concentrations for HEM. (a) RDX on TB, (b) TNT on CB and (c) PETN on wood. Spectral range: 1000-1600 cm\(^{-1}\).

**Table III:** Statistical parameters of PLS calibration models for spectra of HEM deposited on non-ideal, low reflectivity substrates using QCLS. Spectral range: 1000-1600 cm\(^{-1}\).

<table>
<thead>
<tr>
<th>Substrate</th>
<th>Explosive</th>
<th>LV</th>
<th>R(^2) CV</th>
<th>R(^2) PRED</th>
<th>RMSECV (µg/cm(^2))</th>
<th>RMSEP (µg/cm(^2))</th>
</tr>
</thead>
<tbody>
<tr>
<td>TB</td>
<td>PETN</td>
<td>8</td>
<td>0.998</td>
<td>0.999</td>
<td>0.12</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>RDX</td>
<td>8</td>
<td>0.996</td>
<td>0.998</td>
<td>0.17</td>
<td>0.13</td>
</tr>
<tr>
<td></td>
<td>TNT</td>
<td>8</td>
<td>0.973</td>
<td>0.971</td>
<td>0.48</td>
<td>0.41</td>
</tr>
<tr>
<td>CB</td>
<td>PETN</td>
<td>7</td>
<td>0.980</td>
<td>0.985</td>
<td>0.56</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>RDX</td>
<td>8</td>
<td>0.980</td>
<td>0.987</td>
<td>0.61</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>TNT</td>
<td>8</td>
<td>0.945</td>
<td>0.993</td>
<td>1.06</td>
<td>0.47</td>
</tr>
<tr>
<td>Wood</td>
<td>PETN</td>
<td>5</td>
<td>0.989</td>
<td>0.996</td>
<td>0.43</td>
<td>0.29</td>
</tr>
<tr>
<td></td>
<td>RDX</td>
<td>8</td>
<td>0.984</td>
<td>0.985</td>
<td>0.46</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>TNT</td>
<td>9</td>
<td>0.918</td>
<td>0.982</td>
<td>1.03</td>
<td>0.45</td>
</tr>
</tbody>
</table>
4. CONCLUSIONS

A QCL-based spectroscopy system allowed for the detection of HEMs deposited at low surface concentrations (1-15 μg/cm²) on three types of non-ideal, low reflectivity substrates: travel bag fabric (TB), cardboard (CB) and wood. Spectral identification using spectral correlation algorithms was not sufficiently efficient for identifying the HEMs when present on non-ideal, low reflectivity, strong mid-infrared-absorbing substrates. However, multivariate analyses were sufficiently efficient to attain the goals of this investigation. The chemometrics-based multivariate analyses used to detect the target HEMs deposited on TB and CB substrates only required first derivative and mean centering as preprocessing steps. Generating efficient PLS-DA models for wood substrates was a greater challenge, and a third preprocessing step (standard normal variate transformation) was required to achieve the desired discrimination on these substrates. Moreover, classifications according to the type of HEM were achieved. PLS-DA models for the investigated HEMs on the three substrates tested (general PLS-DA model) allowed for discrimination even in the presence of highly interfering and complex substrates, although the model required 12 LVs to account for 80% of the variance. In general, QCL spectroscopy was demonstrated to be useful for detecting HEMs on non-ideal, low reflectivity substrates and for discriminating the analytes from highly interfering substrates when coupled with chemometrics tools such as PLS-DA analysis. Finally, PLS models demonstrated the capability of predicting the surface concentrations of HEMs on the substrates tested using a maximum of 8 latent variables (LVs) to obtain R² values greater than 0.9.

REFERENCES


PROJECT 3: SPECTROSCOPIC DETECTION OF BACTERIA USING TRANSMISSION MODE QCL SPECTROSCOPY

Transmission mode QCLS of microorganisms.

ABSTRACT
Development of capabilities for detection, identification and discrimination of microorganisms, such as bacteria, is at the very top of priorities of defense, security agencies, food technology industries, health related agencies and pharmaceutical and biotechnology industries. This research involves detection of bacteria using spectroscopic techniques such as Fourier transform infrared (FT-IR) spectroscopy and quantum cascade laser spectroscopy (QCLS). Zinc selenide disks were used as support media for bacterial samples. Three strains of bacteria were analyzed: *Escherichia coli*, *Staphylococcus epidermis* and *Bacillus thuringiensis*. Partial least squares combined with discriminant analysis was employed as chemometric tool for classification method to differentiate between bacterial strains using transmission mode QCLS and compared with FT-IR spectra as reference. Spectral differences of the bacterial membrane were used to determine that these microorganisms were present in the samples analyzed.

**Index Headings:** transmission mode quantum cascade laser spectroscopy, Fourier transform infrared spectroscopy, bacteria, zinc selenide substrates, partial least squares-discriminant analysis.
INTRODUCTION

Biological contamination represents a hazardous threat for food-related industries and plays an important role in pharmaceutical and biotechnology clean-room production environments. In health related fields, fast identification of microorganisms is necessary, because the time required for the identification of bacteria and viruses is an important determinant of infection-related mortality rates of hospitalized patients. The development of new techniques for the identification of microorganisms, which include molecular methods, such as mass spectrometry, electrospray ionization and matrix-assisted laser desorption ionization, Fourier transform infrared (FT-IR) spectroscopy, Raman spectroscopy and surface enhanced Raman scattering (SERS) has been addressed in several related publications [1-6]. Vibrational spectroscopy, in particular, has gained considerable attention because generally speaking, there is no need to add chemical dyes or labels for identification provides a definitive means for identifying the species involved and is able to distinguish between target microorganisms and the substrates they are adhered to. In principle, any technique that can be used to obtain vibrational data from substrates (IR, Raman, etc.) can be applied to the study of microorganisms. In addition, there are a number of techniques which have been specifically developed to study the vibrations of molecular species at interfaces based on IR spectroscopy in several of its modalities, including mid-IR (MIR) absorption/reflection in which the vibrational spectrum of a molecule is considered to be a unique physical property. The characteristics of the absorbed/reflected light depend on the molecules found within the sample and the environment in which the molecules are found. However, simplicity of sample preparation and speed of analysis are important advantages when vibrational methods are compared to classical microbial identifications tools [7]. Other advantages of vibrational spectroscopy are that it is a rapid, specific, and a noninvasive analytical method.

Several methods can be used for biological analyses of culture grown colonies to differentiate bacterial spores from other bacteria and from each other. Some of these require special equipment which is not present in an average biological laboratory, whereas others rely on sophisticated electronic equipment and techniques. These types of analyses are usually time consuming and labor intensive, and require expertise in
sample preparation, as well as delay in identification due to the fact that spores, for example, have to be germinated and grown in culture media to sufficient cell numbers until they can be measured. In contrast to these traditional techniques, optical spectroscopy holds promise as a rapid, label-free analytical strategy, which is potentially labor free and time saving, and requires a minimum amount of training. In particular, if appropriate data processing, particularly in the event of exposure to pathogen microorganisms is detected within the first few days, the majority of victims can be treated successfully [5]. Cultures of microorganisms can be grown on any medium that is appropriate for their isolation and cultivation. Identification and characterization of bacteria starts by inspecting the colony morphology when the cells have been cultured in solid media, followed by microscopic analysis of gram-stained preparations [8].

As vibrational spectra of bacterial cells consist of signal contributions of all components in the cells, they reflect their overall molecular composition. Thus, the aim of the present research was to compare measured FT-IR spectra with transmission mode spectra obtained with quantum cascade laser spectroscopy (QCLS) operating in the MIR region, in order to identify and discriminate vegetative bacteria by applying established multivariate techniques to the spectra obtained. Based on the above considerations, this work addressed the demonstration of the capability of this spectroscopic technique in identification and discrimination of bacterial cells between strains such as Bacillus thuringiensis (Bt), Escherichia coli (Ec) and Staphylococcus epidermidis (Se). Research was also aimed at demonstrating that the technique is suitable for detection of microorganisms on MIR transparent substrates.

EXPERIMENTAL

Materials
Zinc selenide (ZnSe) discs obtained from Specac, Inc. (Swedesboro, NJ, USA) were used as support material for transferring samples of bacteria. ZnSe is chemically inert, non-hygrosopic and is very effective in many infrared optical applications due to its extremely low bulk losses of IR light by absorption, high resistance to thermal shock and stability in virtually all environments. It can be easily machined into IR optical devices and is transparent in a wide spectral range from the yellow region (visible) to the far IR.
**Preparation of bacterial samples**

Bacterial strains of Bt (ATCC # 35646), Ec (ATCC # 8789), and Se (ATCC # 2228), were provided by the Microbial Biotechnology and Bioprospecting Lab (Biology Department, University of Puerto Rico-Mayagüez, Mayagüez, PR, USA). Samples were prepared following the protocol reported by Padilla et al., [9]. In brief, pure cultures were grown using Miller modified Luria-Bertani (LB) agar and broth (Thermo-Fisher Scientific, Waltham, MA, USA). After satisfactory growths were achieved in agar, colonies of Bt, Ec and Se were isolated on LB plates and inoculated into 5.0 mL LB broth or tryptic soy broth (TSB) and allowed to grow overnight. Se and Bt were placed in an orbital shaker at 32 °C (~120 rpm) for 24 h and cultured for 72 h. Ec was placed in an orbital shaker at 37 °C until growth of 5 h post inoculation. Sub-cultures were diluted 1:50 in appropriate media and centrifuged at 5K x g for 5 min at 4 °C until pellet formation. Pellets of bacteria were washed once with 20 mL of 1% phosphate buffered saline solution (PBS) to remove growth media. The pellets were resuspended in 4.0 mL of PBS (Ec and Se) or distilled water (Bt) before allowing them to grow for 72 h at 32°C and at 250 rpm to induce spores production. Harvested spores were stored at 4 °C in distilled, deionized H₂O until spectroscopic measurements were made. Serial dilution followed by plating on LB agar in order to count bacteria suspended in PBS. Plates were incubated at 37 °C for ~24 h in order to form colonies. Concentration of the stock suspension of bacteria was determined by back calculating the dilution series.

**Setup**

A MIR spectrometer (LaserScope™, Block Engineering, LLC, Marlborough, MA, USA) based on QCL technology, was used for data acquisition in the MIR: 875-1405 cm⁻¹ of the bacterial suspension samples. QCL are different from traditional semiconductor laser diodes, which use p-n junctions for light emission. Instead, QCLs have multiple active regions, which are composed of a multilayered semiconductor material structure, specially designed to have the appropriate electronic bands [10]. Figure 5.1 shows a schematic diagram of the QCL spectrometer set up for transmission/absorption measurements, rather than the more common reflectance mode setup.
A key feature of the QCL spectrometer is the high spectral radiance or brightness of the QCL source, which results in high signal to noise ratios (SNRs) and excellent quality data from samples compared to FT-IR equipped with conventional thermal sources particularly when analyzing thick, highly diffuse, or very small samples. However, FT-IR spectrometers are important in obtaining reference spectra of bacterial samples. A bench FT-IR interferometer, model IFS 66v/S FT-IR (Bruker Optics, Billerica, MA, USA) was used to characterize the microorganisms of interest and to provide reference spectra to assist in spectral assignment of bacteria studied.

![Sample on surface](image)

**Fig. 1.** (a) Transmission mode QCLS setup. (b) Se deposited on ZnSe discs ready for spectroscopic analysis.

**Data acquisition and analysis**

Bacterial sample of 10 µL each with concentrations at 5.9x10^8/5.0x10^1, 3.5x10^7 and 9.5x10^7 colony forming units per milliliter (CFU/mL) for *Bt* vegetative cell/endospores, *Ec* and *Se* respectively, were deposited on ZnSe discs for MIR transmission measurements. Samples were dried at 40 °C in an incubator for 5-10 min prior to analysis. Four replicas of each kind bacteria sample were analyzed.

OPUS™ 6.0 spectroscopic suite (Bruker Optics) was used to analyze the data obtained. During the analysis performed, the spectra were normalized in order to allow a
proper comparison of the spectra. Partial least squares combined with linear discriminant analysis (PLS-DA) was used for classification of samples by reducing the number of variables and determining if group differences will dominate the total variability of the samples. PLS-DA is a supervised discriminant analysis methodology derived from PLS regression algorithm. PLS not only considers the variation in the original multidimensional dataset generated by analytical measurement (e.g., MIR spectra), but also simultaneously takes into account the variation in the original multidimensional value dataset (e.g., concentration). This algorithm was applied in the MATLAB™ computational environment (The MathWorks, Inc., Natick, MA, USA), using the PLS-Toolbox, v. 7.0.3 (Eigenvector Research, Inc., Wenatchee, WA, USA).

RESULTS AND DISCUSSION
FT-IR Detection of microorganisms
Conventional micro FT-IR in transmission mode was employed as a reference spectroscopic technique for establishing vibrational signatures of bacteria utilized in this study and is shown in Figure 5.2.

![MIR spectra of bacteria deposited on ZnSe discs by FT-IR.](image)

**Fig. 1.** MIR spectra of bacteria deposited on ZnSe discs by FT-IR.
**QCLS Detection and discrimination**

MIR spectra of bacteria acquired on the QCL spectroscopic system in the spectral region from 875 to 1400 are illustrated in Figure 5.3. It is possible to observe the different vibrational spectroscopic signatures of biological components of bacteria clearly. There is difference in the appearance of the QCL spectra in the region between 900 and 1200 cm\(^{-1}\) in which P–O–C and C–O–C stretches of oligo- and polysaccharides occur [11, 12]. However it is more practical to use chemometric tools to have a better spectral discrimination of bacteria.

![MIR spectra of bacteria deposited on ZnSe discs by QCLS.](image)

**Fig. 2.** MIR spectra of bacteria deposited on ZnSe discs by QCLS.

IR spectra were analyzed for identification of bacteria for bending vibrations of fatty acid, proteins, phosphate containing compounds and carbohydrates bands of the microbial cell wall [7]. Tentative band assignments used in bacterial identification of functional groups associated with major vibrational bands in the MIR following Padilla-Jimenez, et al., [13] are shown in Figs. 5.4, 5.5 and 5.6. Spectra in these figures display representative QCLS spectra in transmittance mode and in reflectance mode for each bacterium and a comparison with corresponding reference FT-IR spectra. A QCL based
MIR dispersive spectrometer: LaserScan™, model 710 (Block Engineering) operating in reflectance mode was used to measure the reflectance spectra of the bacterial species. The main differences found between the two QCLS modes arise from spectral artifacts caused from the gap junction between two adjacent QCL laser diodes that affected even more the absorption/transmission QCLS as can be observed in Figs. 4 to 6.

**Fig. 3.** Comparison between *Bt* spectra QCLS in transmittance mode and in reflectance mode with FT-IR.

**Fig. 4.** Comparison between *Ec* spectra QCLS in transmittance mode and in reflectance mode with FT-IR.
Two bacteria are gram-positive and one is gram-negative, according to cell membrane characteristics. Gram-positive bacteria contain teichoic acids that are covalently bound to the peptidoglycan whereas; gram-negative cells do not contain teichoic acids. One of the major components of the outer membrane of gram-negative bacteria is lipopolysaccharide [14]. Ribitol is a significant component of gram (+) bacterial membranes, for this reason in the MIR spectra measured on bacteria the characteristic bands of ribitol were identified. Vibrational modes at 1315, 1180, 1128, 1060, 1037, 953, 915, 889 cm\(^{-1}\) for \(Bt\) and 1357, 1282, 1211, 1189, 1123, 1025, 947, 915, 897 cm\(^{-1}\) for \(Se\) were observed that correspond to contributions of this component in a bacterial cell. The \(P=O\) asymmetric stretch of phosphodiesters in phospholipids on gram-negative bacteria was detected at 1240 and 1086 cm\(^{-1}\). The notable peaks that distinguish vegetative cell forming endospores for bacillus were not observed.

**PLS-DA multivariate analysis**

Figure 7, shows QCL spectra of bacteria from the three strains studied: \(Bt\), \(Ec\) and \(Se\). It was difficult to find differences in the raw QCL spectra between the different classes of bacteria at first sight. To prepare for applying multivariate analysis, data pre-treatments
were applied. In particular, normalization and smoothing were used. PLS-DA was employed as chemometric tool as classification method to differentiate bacterial species.

**Fig. 6.** QCL raw spectra normalized of bacteria showing reproducibility of MIR signatures: **a.** Bacillus thuringiensis (Bt); **b.** Escherichia coli (Ec); **c.** Staphylococcus epidermidis (Se).
Thirty QCL spectra of samples of each bacterium were obtained for a total of 90 total spectra. PLS-DA model was built on vibrational information in spectral region of 875-1400 cm\(^{-1}\). The spectral data were normalized and pretreated employing a second-order derivative (Savitzky-Golay) and improving the classification performance by reducing model complexity and achieving 100% of sensitivity and specificity for all bacteria. These values and other PLS-DA quantitative parameters such as root mean square error of cross validation (RMSECV) and root mean square error of prediction (RMSEP) with the lowest latent variable. These values are shown in Table 1, where important parameters such as, sensitivity and specificity are employed in PLS-DA. Sensitivity is defined as the estimated experimental percentage of correctly classified samples. Moreover, specificity is defined as the estimated experimental percentage of the samples that are rejected by the other classes in the model. Thus, a perfect class model has sensitivity and specificity values of 100 % [15].

<table>
<thead>
<tr>
<th>PLS-DA Parameter</th>
<th>Bt</th>
<th>Ec</th>
<th>Se</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specificity (CV)</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Sensitivity (PRED)</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Specificity (PRED)</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Class. Err (PRED):</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Num. LVs</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>RMSECV</td>
<td>0.158</td>
<td>0.139</td>
<td>0.169</td>
</tr>
<tr>
<td>RMSEP</td>
<td>0.138</td>
<td>0.136</td>
<td>0.162</td>
</tr>
</tbody>
</table>

Spectroscopic data were separated two groups. Approximately 70% of the sample spectra were randomly selected as training set for the calibration and cross-validation (CV) models. The remaining 30% of the spectra were used as an external test set employing all the spectral range. Models were built by a cross-validation procedure performed by using venetian blinds with 10 splits. This procedure was used to build a classification model for 90% of the spectra. Then, the remaining 10% were assessed to
determine the accuracy of the models. The discrimination models for each bacterium on ZnSe cell are shown in Figure 8, which illustrate the predicted cross-validated classes for each sample (as shown in the PLS-DA plots). Based on these results, all the bacteria classes were successfully classified with only four latent variables (LV). Sensitivity and specificity values of 100% for each bacterium and RMSECV and RMSEP values were below 0.2 for all the bacteria. These results corroborate that PLS-DA was highly efficient in discriminating between bacterial strains on ZnSe employing QCLS.

**Fig. 7.** PLS-DA plots for discriminating bacteria on ZnSe discs: (a). cross-validation (CV) predicted of Ec; (b). CV predicted of Bt; (c). CV predicted of Se; (d). Percent variation accounted for by each LV used in the model for each specimen.
CONCLUSIONS

Results obtained demonstrate that transmission mode QCLS, in the range of 875-1400 cm\(^{-1}\) produced high quality spectral information of bacterial species studied: Bt, Ec and Se. Results achieved are in excellent agreement with previously published QCL spectra obtained in reflectance mode and with reference spectra obtained in transmittance mode in the micro compartment of a bench FT-IR interferometer.

When transmittance mode QCL spectra were coupled to multivariate analysis (PLS-DA) a highly effective discrimination between bacterial strains deposited on ZnSe discs: Bt, Ec, and Se was achieved. PLS-DA classified correctly all bacterial samples using only four LV. Sensitivity and specificity values were 100% with very low values for RMSECV and RMSEP. The development of this new methodology for analysis of bacteria using transmittance mode QCLS provides a fast and accurate analysis in the detection of microorganisms accompanied to a great potential for discrimination between different and similar strains microorganisms as was demonstrated in this study.

REFERENCES

thuringiensis using Silver Nanoparticles Reduced with Hydroxylamine and with Citrate Capped Borohydride, Int. J. Spectrosc. 2011. 9.


X. STATUS AND OVERALL PROJECT PROGRESS: FINAL DEVELOPMENTS

The results of studies completed indicate that the experimental setups designed as part of this project can be applied in real world situations such as screening for explosives on complex substrates. The results of the studies also indicate that explosives, chemical warfare agents and toxic compounds can be strategically mixed with other compounds with similar chemical structure in order to confuse the authorities and make criminal or terrorist episodes. Therefore the detection of concealed hazardous compounds such as explosives, chemical and biological agents employed by terrorist are considered of high interest in CWMD research.

This project was also intended to expand the knowledge in the field of preparation of samples and standards based on thin films of HEM also HEM for both basic and applied research by depositing traces of target analytes on substrates of high interest using spin coating technology to develop the mentioned assemblies. The two main goals of this project are related to the study of solid-substrate interactions. The first and primary goal is to develop methodologies capable of producing solid samples deposited on substrates in controlled size and distribution modes deposited on selected surfaces to generate specimens that would reproduce real contamination samples. In a second closely related aim, the prepared samples will be used to analyze-substrate interactions, physical and chemical properties derived from these interactions, including, but not limited to, particle sizes and shapes, polymorphism at the sub-micro range, optical and spectroscopic properties, thermal properties, and others. At the micron and sub-micron ranges, the residence time of these adsorbate-substrate assemblies becomes an important property for study, since behavior at the microscopic scale is very different from bulk phase characteristics. Finally, the applications of these well-characterized samples include use in experiments that require fine control of the distribution of loadings of analytes on surfaces.

Use of multivariate analysis to enhance spectroscopic based threat detection within the scope of this project and its multiple sub-projects has been continuously developing. Use of Neural Network Analysis (NN) and other multivariate techniques for analyzing
large amounts of data will undoubtedly advance spectroscopic based remote detection of hazardous chemical and biological agents or their simulants.

Work on QCLS based detection of explosives is also being pursued in collaboration with Eos Photonics and Thermo-Fisher Scientific (formerly Ahura) for development explosives detection equipment. The commercial applications have been envisioned in the standards market for traditional nitroexplosives or high explosives (HE) and for cyclic organic peroxides (CAP) HME, including DADP, HMTD and TMDD.

The technologies under development are useful for gas phase detection of relatively high vapor pressure HME (such as 2,4-DNT, TATP and others), for chemical agents, for biological threat agents and for illicit drugs detection. They can also be used in Process Analytical Technology (PAT) applications in pharmaceutical and biotechnology industries. Among the applications for pharma and biotech are in tablet content uniformity verification and in cleaning validation of batch reactors, mixers, filters and dryers. Coupling with multivariate analysis routines in all of these applications will definitively enhance the sought results and drive low limits of detection to the lowest possible values.

In still another application of QCLS coupled to chemometrics, detection and classification of microorganisms, such as bacteria, viruses and yeasts can be addressed very successfully, shortening analysis time from hours to days to seconds. Defense and security as well as biotech, food industries and public health facilities and hospitals will benefit to a large extent from the methods and techniques under development at R3-C research and development component.

VII. LEVERAGING OF RESOURCES
The National Science Foundation Chemical Measurement and Imaging Program supports research focusing on chemically-relevant measurement science and imaging, targeting both improved understanding of new and existing methods and development of innovative approaches and instruments. For which the primary focus is on development of new instrumentation enabling chemical measurements likely to be of wide interest and utility to the chemistry research community. Some of our efforts are directed to targeting science and technology research programs such as this one.
Under the auspices of the Defense Threat Reduction Agency (DTRA) of the DoD and within the program Rapid Investment Funds, Eos Photonics (Lead) and Thermo-Fisher Scientific (Co-Lead) and University of Puerto Rico-Mayaguez (Academic Partner, S.P. Hernández-Rivera Other companies are also being considered as partners for collaborative and transitioning work.

Grants:

1. Project: Puerto Rico REU: Research Training in Cross-disciplinary Chemical Sciences
   - Program: Research Experiences for Undergraduates (REU)
   - Funding agency: NSF
   - Status: Approved, $450,000 3 years; co-Funded by DOD
   - PI: Dr. Osvaldo Cox
   - Co-PI: Dr. Beatriz Zayas
   - Research Mentors: Dr. Ajay Kumar, Dr. Mitk’El B. Santiago, Dr. Oliva M. Primera-Pedrozo, Dr. Jonathan Friedman, Dr. Shikha Raizada, Dr. Antonio Alegría, Dr. Carlos Cabrera and Dr. Samuel P. Hernández.
   - Description: The goal of this project is to build a cross-disciplinary research community in chemistry to increase the competitiveness of chemistry majors in Puerto Rico and thereby increase their preparedness and enthusiasm to continue to graduate school. The project will recruit students at Universidad Metropolitana, a Hispanic Serving Institution in San Juan, Puerto Rico and offer a program during the academic year followed by a summer component at four (4) partner institutions which are research intensive institutions.

2. Project: Acquisition of Confocal RAMAN-AFM-NSOM Imaging Spectroscopic System
   - Pending: Major Research Instrumentation, NSF 13-517
   - Funding agency: NSF
   - Status: Pending; Requested: $803,596
   - PI: Dr. Carlos Padín-Bibiloni, Chancellor, UMET-AGMUS
Co-PIs: Dr. Samuel P. Hernández-Rivera, Dr. Oliva M. Primera-Pedrozo, Dr. Mitk’El B. Santiago, Dr. Wilfredo Otaño,

Researchers: Dr. Francisco Marquez, Dr. Jaime Ramirez-Vick, Dr. Luis F. de la Torre

Description: Metropolitan University (UMET) and UPR-M proposes the acquisition of an equipment capable of performing atomic force microscopy (AFM), near-field scanning optical microscope (NSOM), Raman spectroscopy (RS), surface enhanced Raman spectroscopy (SERS) and material characterization, in general. The objective is to increase the participation of scientists from multiple universities in Puerto Rico in state-of-the-art research and to broadly impact the STEM pipeline education and training in spectroscopic techniques, in fiber optics in scientific instrumentation, in laser techniques, The requested instrumentation will enable a new network for cutting-edge materials science research that brings together young, competitive researchers and seasoned experienced investigators from four (4) different institutions across the island of Puerto Rico: UMET, University of Puerto Rico-Mayaguez (UPR-M), University of Puerto Rico-Cayey (UPR-C), and Universidad del Turabo (UT).


• Program: Defense University Research Instrumentation Program (DURIP) PA-AFOSR-2013-0001
• Funding agency: Department of Defense – DURIP
• Status: Approved, $225,000
• PI: Dr. Samuel P. Hernández-Rivera
• Co-PI: Dr. Oliva M. Primera-Pedrozo, Dr. Luis F. de la Torre
• Participants: Dr. Arturo Hernández-Maldonado, Dr. Wilfredo Otaño, Dr. Jaime Ramírez-Vick,
• Description: This proposal is aimed at strengthening research and education infrastructure capabilities at University of Puerto Rico – Mayaguez (UPR-M)
aligned with ARO goals and objectives. Equipment acquired through this grant will assist in developing local workforce by fostering the development of the doctoral program in Applied Chemistry at UPR-M, one of the nation’s leading Hispanic Serving Institutions producing graduates in science, mathematics, and engineering.

A second goal is to increase participation of students and scientists in state-of-art research and to impact STEM pipeline education and training in confocal Raman microscopy, atomic force microscopy (AFM), scanning near-field optical microscopy (SNOM), surface enhanced Raman spectroscopy, tip enhanced Raman spectroscopy (TERS) and in materials characterization in general. The requested instrumentation will greatly enhance current research capabilities in National Defense/Security in materials sciences, nanotechnology applications in sensors development, biophysical/cellular studies, regenerative medicine, anticancer drug delivery, and cancer studies, including new protocols for early and non-invasive diagnosis. This equipment will also enhance the research experiences of graduate and undergraduate students from underrepresented groups by exposing them to research equipment not currently available in any research or educational institution. Computer remote access will provide opportunity for successful sharing and maximizing resources. The project will create a “virtual laboratory” environment including implementation of tools and resources to build necessary infrastructure enabling online instrument access: network connection, multiuser account, instrument registration schedule and multiuser sharing screen.

4. Project: New Approaches to Cleaning Validation in Pharma & Biotech Industries

- Program: Alternate technologies for evaluation of equipment surface samples for Cleaning Validation
- Funding agency: Industry-University Consortium (INDINIV), Government of Puerto Rico
- Status: Pending, $80,000
- PI: Dr. Samuel P. Hernández-Rivera
- Co-PI: Dr. Leonardo C. Pacheco-Londoño, Dr. Pedro M. Fierro-Mercado,
Researchers: Mrs. Nataly Y. Galán-Freyle, MS; Miss. Amanda Figueroa-Navedo, UGS

Description: The proposed project focuses on the development, test and validation of new methodologies for analyzing trace level amounts of APIs, detergents and cleaning agents left as contamination residues on surfaces of batch reactors and handling equipment as part of cleaning protocols. The analytical methodologies to be developed will be robust enough for consistent every-day use and will be of wide capability of implementation in Pharma and Biotechnology industries. The analytical methods will be capable of providing highly cost effective alternatives to established methodologies of cleaning validation in these industries.

For direct examination of equipment surfaces a QCL spectrometer based methodology will be developed, tested and validated to substitute swab/HPLC based sample collection and analysis. Recent developments in QCL technology include size reduction, which has enabled the transition from table-top laboratory instruments to easy-to-handle, portable, and small instruments. Moreover, the augmented output power has enabled the use of QCL-based spectrometers in long-distance applications, making the detection of chemicals possible at several meters from the source. However, the majority of previous investigations are focused on the detection of chemicals deposited on nearly ideal, highly reflective substrates (such as metallic surfaces), and there are no published reports on spectral effects of real-world substrates (such as Teflon, silicones and polymer plastics) on the spectra of the analyzed targets.

For the particular cases that using a direct, remote sensed methodology is not viable (the field of view of the laser is obstructed), we propose two alternative, in situ, validate-CIP methodologies that will be based on the use of swabs or their equivalent. The first one envisions collecting the sample using swabs and analyzing with CE-QCLS. The other proposed methodology uses gold/silver nanoparticles (Au/Ag NP) coated swabs/filters to capture the APIs/detergents and use a portable Raman spectrometer and surface enhanced Raman scattering (SERS) to analyze the target chemicals residues.
VIII. PROJECT DOCUMENTATION AND DELIVERABLES

A. Peer Reviewed Archival Publications (including book chapters): 29


[22]. Félix-Rivera, H. and Hernández-Rivera, S.P., “Raman Spectroscopy Techniques for the Detection of Biological Samples in Suspensions and as Aerosol Particles:


B. Conference Proceedings: 9


C. Other Presentations
Invited Presentations: 4


Presentations: 17


Continued Education Courses:

IX. BIBLIOGRAPHY


