Journal of Biomedical Optics

SPIEDigitalLibrary.org/jbo

Discrimination of selected species of pathogenic bacteria using near-infrared Raman spectroscopy and principal components analysis

Fernanda SantAna de Siqueira e Oliveira
Hector Enrique Giana
Landulfo Silveira, Jr.

SPIE

Discrimination of selected species of pathogenic bacteria using near-infrared Raman spectroscopy and principal components analysis

Fernanda SantAna de Siqueira e Oliveira,[a] Hector Enrique Giana,[b] and Landulfo Silveira Jr.[c]
[a]Universidade Paulista—UNIP, Instituto de Ciências e Saúde, Rod. Pres. Dutra, km 157, São José dos Campos, SP, Brazil 12240-420
[b]Laboratório Oswaldo Cruz, Pça. Cândida Maria Sawaya Giana, 128, São José dos Campos, SP, Brazil 12243-003
[c]Universidade Camilo Castelo Branco—UNICASTELO, Biomedical Engineering Institute, Parque Tecnológico de São José dos Campos, Rod. Pres. Dutra, km 138, São José dos Campos, SP, Brazil 12247-004

Abstract. A method, based on Raman spectroscopy, for identification of different microorganisms involved in bacterial urinary tract infections has been proposed. Spectra were collected from different bacterial colonies (Gram-negative: *Escherichia coli*, *Klebsiella pneumoniae*, *Proteus mirabilis*, *Pseudomonas aeruginosa* and *Enterobacter cloacae*, and Gram-positive: *Staphylococcus aureus* and *Enterococcus* spp.), grown on culture medium (agar), using a Raman spectrometer with a fiber Raman probe (830 nm). Colonies were scraped from the agar surface and placed on an aluminum foil for Raman measurements. After preprocessing, spectra were submitted to a principal component analysis and Mahalanobis distance (PCA/MD) discrimination algorithm. We found that the mean Raman spectra of different bacterial species show similar bands, and *S. aureus* was well characterized by strong bands related to carotenoids. PCA/MD could discriminate Gram-positive bacteria with sensitivity and specificity of 100% and Gram-negative bacteria with sensitivity ranging from 58 to 88% and specificity ranging from 87% to 99%. [DOI: 10.1117/1.JBO.17.10.107004]

Keywords: near-infrared Raman spectroscopy; discriminant analysis; Mahalanobis distance; principal components analysis.

Paper 12254P received Apr. 25, 2012; revised manuscript received Sep. 5, 2012; accepted for publication Sep. 11, 2012; published online Oct. 9, 2012.

1 Introduction

Rapid identification of infecting microorganisms, such as bacteria and fungi, is a challenging task. The time required for identification of infectious microorganisms is an important issue in hospitalized patients.[1] Although in recent years, techniques have been developed that allow direct identification of microorganisms, culturing followed by observations of morphological and biochemical reactions on different substrates is still recognized as the gold standard for the microorganism identification.

The implementation of traditional techniques based on isolation and further identification of the most clinically relevant species by biochemical tests is usually laborious, requiring a pure microbial culture and, in some cases, takes three to five days as it involves a combination of several methods in order to correctly identify the microorganism.[2] Sometimes, two or more tests are necessary to obtain the correct identification. Such limitations can impact the patient's clinical diagnosis. In emergency cases, where action should be taken immediately and before microbial identification is complete, it is common for empiric administration of broad-spectrum antibiotics to occur. As a side effect, this practice may lead to undesired resistance to antimicrobial agents, often delaying and ultimately disturbing the resolution of the infection.[3] Kerremans et al.[4] showed that shortening the time needed to identify the infectious agent and allowing administration of more specific drugs—would increase the chances of cure with reduced costs and with lower morbidity and mortality rates.

Researchers have proposed optical spectroscopic techniques for rapid and accurate identification of relevant microorganisms as these require no reagents or dyes, a minimum of sample preparation, hence allowing identification using small amounts of biomass.[5–8] These techniques can provide information on the biochemical constitution of various complex biological systems, including bacteria and fungi, enabling the collection of spectral information even with reduced culture time.[9] Among them, Raman spectroscopy has the ability to measure the molecular vibrational energies of cellular constituents, such as proteins, carbohydrates, lipids, and nucleic acids from membrane cell wall and components, allowing information on the molecular structure of cells and thereby discrimination of microbial species.[6,7,10–13] It is important to evaluate differences in the species influenced by the culture medium, incubation time, and the thickness/depth of the colony, since studies have shown that these factors influenced the spectra obtained.[14] Given the strong interest in rapid methods for microbiological analysis, Raman spectroscopy has proved to be an important tool in the study of the structural and chemical characteristics of microorganisms in order to provide identification in a faster and less expensive manner and even directly on the culture plate.[15]

The objective of this study was to evaluate the ability of Raman spectroscopy to identify, *in vitro*, the spectral differences of seven clinically relevant species of bacteria commonly found in urinary tract infections, previously identified by conventional biochemical methods. To do so, it has been necessary to develop a classification model to discriminate these species according to the spectral differences between them using the multivariate

Address all correspondence to: Landulfo Silveira Jr., Universidade Camilo Castelo Branco—UNICASTELO, Biomedical Engineering Institute, Parque Tecnológico de São José dos Campos, Rod. Pres. Dutra, km 138, São José dos Campos, SP, Brazil 12247-004. Tel: +55(12)3905-4401; E-mail: landulfo.silveira@unicastelo.br

0091-3286/2012/$25.00 © 2012 SPIE

statistical method of PCA (Principal Components Analysis) and the (statistical) Mahalanobis distance.

2 Materials and Methods

2.1 *Selection of the Culture Media*

We conducted a preliminary study to choose the agar that presented the least interference in the spectra of bacteria. Four culture media [MacConkey, Mueller Hinton with 5% sheep blood (blood), Mueller Hinton and cysteine lactose electrolyte deficient (CLED) agars]) were tested and one strain of *Escherichia coli* was grown on each of these. Following growth, agars and colonies were subjected to Raman spectroscopy as follows:

1. agar without bacteria growth;
2. agar with bacteria growth, but in areas with no colonies;
3. samples of colonies of *E. coli* grown on each medium.

For 1 and 2, one fragment of approximately 2×2 mm^2 of each agar was removed and evaluated spectrally. For 3, a small amount of biomass of *E. coli* grown on each medium was collected using a bacteriological inoculating disposable loop (10 μL capacity) and placed on a piece of reflective aluminum foil. Also, a spectrum of agar without additives (Biobrás Diagnósticos, MG, Brazil) was taken for comparison purpose.

2.2 *Origin and Types of Bacteria Samples*

We used 38 clinical strains obtained from urine delivered to the municipal laboratory health service of Jacareí (SP, Brazil) for microbiological analysis. The strains used in this study belong to the Gram-negative species: *Escherichia coli* ($n = 5$), *Klebsiella pneumoniae* ($n = 7$), *Enterobacter cloacae* ($n = 5$), *Proteus mirabilis* ($n = 6$) and *Pseudomonas aeruginosa* ($n = 5$) and Gram-positive species: *Staphylococcus aureus* ($n = 5$) and *Enterococcus* spp. ($n = 5$). The strains of both Gram-positive and Gram-negative bacteria were grown and stored on Mueller Hinton agar in a refrigerator at 4°C to 8°C, except for *Enterococcus* spp., which was grown in blood agar due to difficulty in growing it on Mueller Hinton agar.

A day before spectral analysis, the strains were inoculated in agar plates, incubated at 35 ± 2°C for 16 to 18 h and verified for signs of contamination. Agar plates were then safely packed in a biohazard box, cooled to 4 to 8°C with ice and transported (about 30 min.) to the Laboratory of Biomolecular Spectroscopy (BioSpecLab) at UNICASTELO. At the time of spectral analysis, a small volume of the biomass (approx. 10 μL) was carefully removed from the colonies, with the aid of a disposable loop, and transferred by touching the loop to the aluminum foil.

2.3 *Raman Measurements*

Raman spectra were acquired using a dispersive Raman spectrometer (Lambda Solutions, Inc., Massachusetts, model P1) composed of a diode laser (830 nm) as excitation source, coupled to an optical fiber cable "Raman probe," an imaging spectrograph with diffraction grating of 1200 lines/mm coupled to a "back thinned, deep-depleted" CCD camera of 1340×100 pixels, thermoelectrically cooled to −75°C. Spectrometer resolution was about 2 cm^{-1}. The resulting laser output power was about 300 mW at the Raman probe. The collection of the light scattered by the sample was performed by the Raman probe coupled to the spectrometer. The acquisition and storage of spectra is performed by a microcomputer using the software RamanSoft (Lambda Solutions, Inc., MA) via USB connection, which controlled the exposure time of the detector and the number of acquisitions per sample and stored the spectra for further analysis and interpretation.

The exposure time for obtaining the spectra was set to 5 s, and accumulation was set to 10 (total 50 s collecting time). All strains were scanned in five replicates (each replicate was considered a sample in the discrimination model), each one collected from a random place on the agar surface, allowing evaluation of the reproducibility of the spectra. Two spectra were not considered due to CCD saturation. A total of 182 spectra were obtained from all 38 samples. Spectra were preprocessed to remove background fluorescence using a 7th order polynomial fitted over the spectral range and subtracted from the gross spectrum,[16] normalized to the intensity of the Raman peak at 1453 cm^{-1} and then averaged and plotted with the aim of identifying spectral differences between all species that could be relevant for identification. Polynomial filtering has as an advantage very easy implementation and low spectral distortion.[16,17]

2.4 *Bacteria Discrimination Using PCA*

In order to develop a classification model to discriminate between the Raman spectra of the clinical species, according to similarity in the relevant spectral characteristics related to bacteria species, spectra were submitted to PCA. PCA is a statistical tool that can be used to analyze data of a multivariate nature, which transforms a set of inter-correlated original variables into a new set of uncorrelated variables called principal component (PCs) spectral vectors. With PCA, it is possible to identify patterns in the dataset and to highlight similarities and differences according to similarities and differences in the samples and group together the variables that are highly correlated. Through linear transformation of the data to a new coordinate system, the largest variance of all projections will be positioned as the first coordinate (principal component 1—PC1), the second largest variance, orthogonal to the first one, will be the second coordinate (principal component 2—PC2) and so on. Each principal component has a weight score (SC), which is the intensity of each principal component to recover the original data.

For separation of the bacteria species with similar spectral information, the Mahalanobis distance was applied to the SCs that showed higher differences among groups. Mahalanobis distance can be used to separate or discriminate groups according to differences in the selected variables (in our case the SC) and also to check the consistency in the distribution of the clustered points of an experiment. Mahalanobis distance takes into account the covariance matrix of the dataset instead of only the distance of a point from the centroid of the distribution, allowing examination of an eventual ellipsoid shape of the data cloud.

The PCA was calculated through a routine written under MATLAB 6 using all spectra of different strains. The resulting variable PC was plotted in Excel in order to identify main peaks related to bacteria type. The bacteria species grouping was done by evaluating the binary plot of the first five SC ($SC1 \times SC2$, $SC1 \times SC3, \ldots, SC4 \times SC5$) and calculating the mean Mahalanobis distance between the groups. The routine for

calculating the Mahalanobis distance was adapted from the original developed by Galhanone (unpublished data).

3 Results and Discussion

3.1 *Selection of Culture Medium*

Initially, the influence of different culture media (agar) on the Raman spectra using one specie (*E. coli*) was verified in order to determine which medium was most suitable. The spectra of MacConkey, blood, Mueller Hinton and CLED agars with and without *E. coli* growth, as well as a spectrum for pure agar, are presented in Fig. 1. Also, the spectra of *E. coli* grown in each one of the selected media are presented in Fig. 2. It was found that the Raman spectra of the different agars [Fig. 1(a)] showed some bands in the same positions as pure agar [Fig. 1(b)] and other bands in different positions and different intensities. These differences can be explained by the composition of each medium, since different substances are added to agar with specific objectives, such as to inhibit the growth of some species, reveal the presence of metabolic products, classify bacteria according to the growth characteristics, differences in manufacturers and lots, etc. Table 1 shows the main Raman peaks of the four agar spectra and the tentative assignment for the most relevant bands of agar and media.

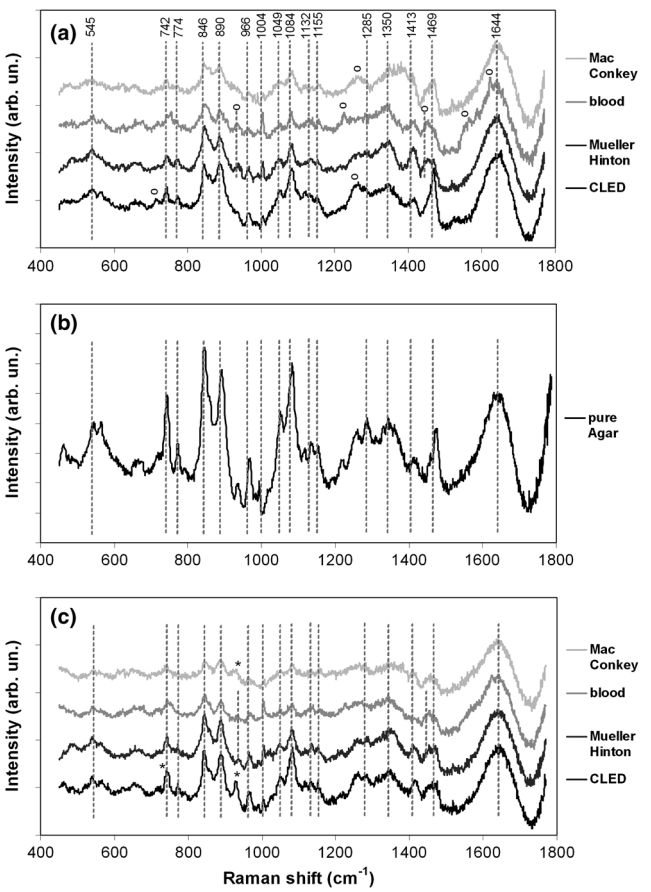

Fig. 1 Normalized Raman spectra of fragments of the four types of agar: (a) without bacterial growth; (b) reference spectrum of agar without additives, (c) with *E. coli* growth. The symbol ° represents different peaks in each of the agars and are suggestive of differences in the composition of the culture media and the symbol * represents Raman bands suggestive of products of bacterial metabolism. Spectra were displaced for better viewing.

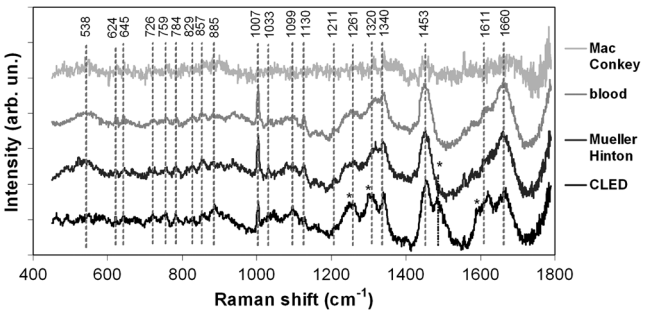

Fig. 2 Raman spectra of *E. coli* grown in four different types of culture media. The symbol * represents Raman bands of CLED agar that appeared in the bacteria spectrum probably due to medium's intake by the microorganism. Spectra were displaced for better viewing.

The spectrum of MacConkey agar [Fig. 1(a)] presented an intense peak in 1270 cm^{-1} when compared with the spectra of agar without additive [Fig. 1(b)], which can be attributed to lactose present in this medium. CLED agar has a different composition, with the addition of various substances that can produce colonies colored according to microbial metabolism. This pigment fluoresces and may influence the spectral analysis. This medium has peaks with higher intensity at 890, 1084, and 1469 cm^{-1} compared to the other media. It also has lactose in its composition, as implied by the presence of the characteristic peak at 1270 cm^{-1}. Mueller Hinton agar is a medium commonly employed to perform antibiograms, with basic chemical constitution and nutrients that do not interfere with the antimicrobial's effectiveness, making it an option for use in optical techniques. This agar presents peaks at 935, 1004, and 1450 cm^{-1}, which can be assigned to the protein extract.[5,14,22]

Although there are only small amounts of additives in the blood agar, hemoglobin can influence the bacterial spectral pattern. The peaks in the region of 1004, 1226, 1450, 1562, and 1624 cm^{-1} may be attributed to hemoglobin and other blood proteins.[6,7]

With bacterial growth [Fig. 1(c)], the MacConkey agar presented bands of lower intensity in the whole spectral range because of the strong fluorescence generated by the marker phenol red, indicating fermentation of lactose (producing lactic acid) and changing the color of the agar. This coloration can be transmitted to bacterial colonies and can decrease the signal-to-noise ratio of the Raman spectrum. Blood agar had small features from hemoglobin and decreased signal-to-noise ratio. The CLED agar had spectral features in the 700 to 1000 cm^{-1} region, with intense peaks in the same positions as the pure agar and bands of lower peak intensity (around 1460 m cm^{-1}) that may be attributed to changes in the polarizability of the agar molecule induced by changes in the pH.

The spectra of *E. coli* after growth in each one of the culture medium (Fig. 2) presented features depending on the medium used. As observed in the spectrum of the medium after culture, the spectrum of *E. coli* grown in MacConkey presented a few spectral features with a very low signal-to-noise ratio due to the strong fluorescence background, which decreases the Raman signal quality. The spectra of *E. coli* grown in blood and Mueller Hinton agars presented similar spectral features, since these media presented similar constitution. A small feature can be seen at 1611 cm^{-1}. The *E. coli* grown in CLED presented strong peaks at 886, 1480, and 1620 cm^{-1}, which are the same features appearing in the spectrum of CLED agar. This reflects the

Table 1 Peak positions and tentative attributions for Raman spectra of culture media and bacteria used in the study based on recent literature.[5,8,14,16,17,18,19,20,21]

Raman shift (cm^{-1})	Attribution	Origin
538	COC glycosidic ring deformation	Bacteria
545	CCC deformation	Agar
624	CC twisting—tryptophan	Bacteria
645	Tyrosine	Bacteria
672	Valine	Bacteria
721	Agar	Agar
726	Adenine ring stretching; peptidoglycan	Bacteria
742	CC skeletal deformation—galactose ring	Agar
759	Tryptophan	Bacteria
774	CC skeletal deformation—galactose ring	Agar
784	Cytosine, uracil (ring stretching)	Bacteria
814	C–O–P–O–C—RNA binding	Bacteria
829	Tyrosine	Bacteria
846	CC deformation/OCO wagging/CH vibrations with C–OH	Agar
857	C–O–C stretching of glycosidic linkage (saccharides); C–C proline stretching and CCH deformation ring breathing of tyrosine (protein); teicuronic acid of Gram positives cell wall	Bacteria
890	CCH deformation	Agar
907	C–O–C stretching of glycosidic linkage (saccharides); teicuronic acid of Gram positives cell wall	Bacteria
930[b]	Saccharides	Agar
935	C–C skeletal stretching—protein	Culture media
938	C–C stretching (amide III)—protein	Bacteria
966	CCH deformation	Agar
951	Phenylalanine, proline	Agar
1004[a]	C–C skeletal stretching of aromatic ring phenylalanine—protein	Agar
1007	C–C skeletal stretching of aromatic ring—phenylalanine/tyrosine	Bacteria
1020[b]	Bacteria metabolism (?)	Culture media
1049	C–O exocyclic stretching	Agar
1030–1130	Carbohydrates, mainly C–C, C–O, C–O–H stretching	Bacteria
1033	C–H in-plane deformation—phenylalanine/proline (proteins); C–O and C–C stretching—saccharides	Bacteria
1084	COH deformation/CCO stretching	Agar
1099	Nucleic acids (PO_2^-symmetrical stretching); C–C and C–O–C skeletal stretching—glycosidic linkage of saccharides	Bacteria
1130	C–C skeletal stretching acyl (trans conformation) of lipids; C–O and C–C—saccharides stretching	Bacteria
1132	COH sugar deformation	Agar
1155	COH deformation	Agar

Table 1 (*Continued*).

Raman shift (cm^{-1})	Attribution	Origin
1156	C–C and C–N—protein stretching; C–O and C–C—saccharides	Bacteria
1162	–C–C– conjugated stretching—carotenoids	Bacteria
1173	C–H wagging of tyrosine; guanine and cytosine; fatty acids (stearic acids)	Bacteria
1211	Amide III; C–C tyrosine stretching, phenylalanine, tryptophan (protein)	Bacteria
1226[a]	C–H stretching—hemoglobin	Agar
1261	C–N e N–H stretching (amide III); thymine and adenine (ring breathing); CH_2 lipids deformation; saccharides	Bacteria
1270[a]	Lactose	Agar
1285	CH_2 twisting	Agar
1293	CH_2 fatty acids deformation; cytosine	Bacteria
1320	C–N and N–H stretching (amide III); CH_2 and CH_3—protein deformation; guanine breathing ring	Bacteria
1340	CH_2 and CH_3—fatty acids and protein deformation; N–H stretching (amide III); C–C stretching—tryptophan; adenine, guanine (ring breathing)	Bacteria
1350	CH_2 wagging	Agar
1382	$–COO^-$ symmetric and asymmetric stretching—peptidoglycan	Bacteria
1413	CH deformation	Agar
1421	Adenine, guanine, –C–O vibration modes—peptidoglycan	Bacteria
1450[a]	CH_2/CH_3 deformation—proteins	Agar
1453	CH_2 and CH_3 deformations—lipids and proteins	Bacteria
1469	CH_2 deformation	Agar
1525	–C–C conjugated stretching—carotenoids	Bacteria
1557	Tryptophan; exopolysaccharides	Bacteria
1562[a]	CH_2 stretching—hemoglobin	Agar
1578	Adenine, guanine (ring stretching); –C–O vibration modes—peptidoglycan	Bacteria
1611	C–C ring stretching—phenylalanine, tyrosine and tryptophan	Bacteria
1624[a]	C–C asymmetric stretching—hemoglobin	Agar
1644	C–O stretching (amide I)—protein, O–H wagging (water)	Agar
1660	C–O stretching (amide I); C–C stretching—lipids	Bacteria

[a]Peaks related to the differences in the composition of the culture media.
[b]Peaks that appeared in the culture media after bacteria growth.

interference of the medium's composition in the bacteria spectrum; this might reflect the agar intake by the microorganism.

Mueller Hinton agar, which consists of basic nutrients and no additives that could significantly influence the spectral pattern, proved to be the best option for use in Raman spectroscopy. This agar has been used in several studies,[14,22] almost certainly due to this characteristic.

3.2 *Spectra of Bacteria*

In order to develop an optical technique for bacteria identification, we obtained the spectra of clinical strains of seven different bacteria species (Fig. 3). These Raman spectra showed peaks at positions characteristic of the bacteria's biochemical constitution, with peak assignments of these bands shown in Table 1. By analyzing Fig. 3, it can be seen that there is similarity between the spectra of different species, represented by the peaks in the same positions and same intensity of all bacteria analyzed, namely: 538, 624, 645, 726, 784, 814, 829, 857, 907, 938, 1007, 1033, 1099, 1130, 1211, 1261, 1320, 1340, 1421, 1453, 1578, 1611, and 1660 cm^{-1}. This reflects a general similarity in the biochemical composition of all species, in agreement with the results observed in other studies.[12]

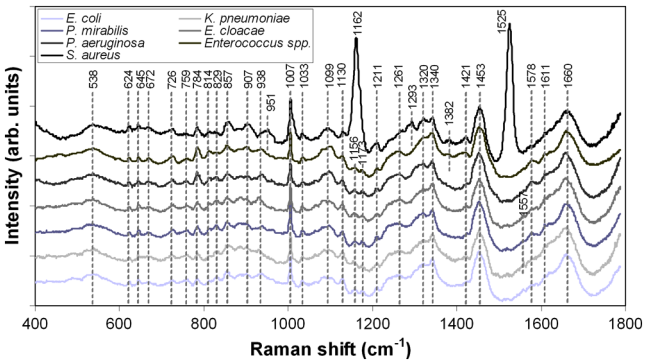

Fig. 3 Mean spectra of all bacteria (Gram-positive and Gram-negative) used in the discrimination model.

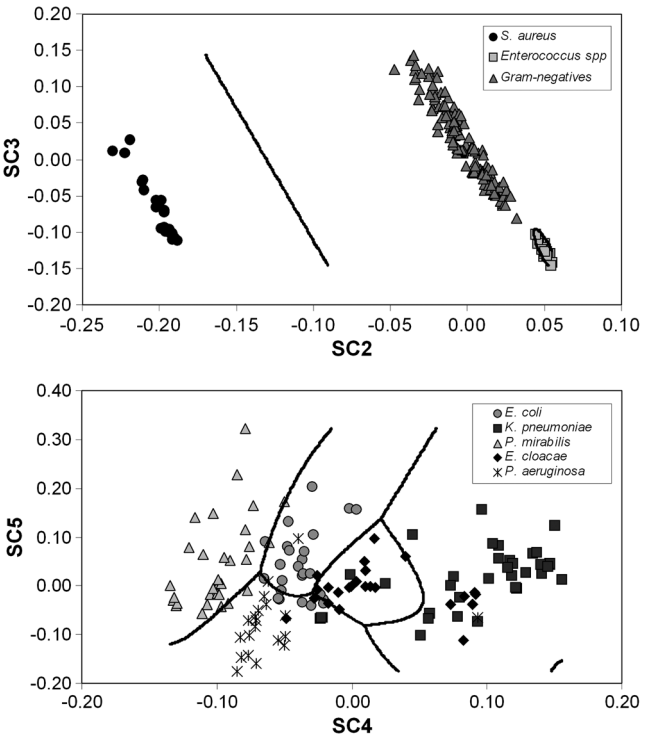

Fig. 4 PCA scores for bacteria discrimination: a) SC2 × SC3 for discriminating Gram-positive and b) SC4 × SC5 for discriminating Gram-negative.

It was observed that the peak at 1382 cm^{-1} appears only in Gram-positive strains (*S. aureus* and *Enterococcus* spp.), while the peaks at 726, 857, 907, 1340, and 1421 cm^{-1} are present in both groups but with increased intensity in Gram-positive species, which could indicate a higher concentration of the cell wall constituents peptidoglycan and teichoic or teichuronic acids,[18,23] being thicker in Gram-positive species[23] as described in a study developed by Jarvis et al.,[12] who presented the peak at 1340 cm^{-1} as being important in distinguishing between the species (Gram-positive and Gram-negative). *S. aureus* shows very characteristic peaks at 951, 1162, 1293, and 1525 cm^{-1}, which can be attributed to pigments of the carotenoid family, characteristic of this type of microorganism.[19,24] The species *Enterococcus* spp., *P. aeruginosa,* and *S. aureus* showed a slightly more intense peak at 784 cm^{-1} when compared to other species. The species *Enterococcus* spp. and *S. aureus* showed a peak at 1007 cm^{-1} of higher intensity when compared to other species. The peak at 759 cm^{-1} had reduced intensity in *S. aureus* and *Enterococcus* spp. The peak at 1578 cm^{-1} appeared more intense in *Enterococcus* spp., *P. aeruginosa*, and *P. mirabilis*. The species of *E. cloacae* and *K. pneumoniae* showed a peak at 1557 cm^{-1} of higher intensity. The peak at 1130 cm^{-1} appeared more intense in *E. coli* and *P. mirabilis* compared to the other species. The differences in peak intensities are related to differences in the biochemical constitution of bacteria and are in accordance with recent literature.[6,12,14,24,25]

3.3 *Bacteria Discrimination Using PCA*

In order to develop a routine for classification of bacterial species due to Raman spectral characteristics (inter-group differences and intra-group similarities), all spectra were submitted to PCA. Following PCA, group discrimination by means of Mahalanobis distance, related to the bacteria species, was applied to the principal component scores (SCs). To do so, all the normalized spectra were submitted to PCA calculation. The output variables (PC, SC and latent variable—LV) were then used to correlate the spectral differences with the bacterial species. The LV of the dataset indicated that principal components PC1 to PC5 are responsible for 98.8% of all spectral variations (PC1 = 90.3%; PC2 = 5.0%; PC3 = 2.3%; PC4 = 0.9%; PC5 = 0.3%); these PCs were then used in the discrimination model. The scores of principal components PC1 to PC5 (SC1 to SC5) were binary plotted in combination to determine which SC would provide the best discrimination.

Figure 4 presents binary plots of the principal components SC2 × SC3 and SC4 × SC5, which were found to carry most of the spectral variations between groups and show the highest relevance to bacterial discrimination. In both cases, groups were separated by the average Mahalanobis distance between specie. The score SC2 × SC3 was able to discriminate Gram-negative from Gram-positive bacteria, including the species. The score SC4 × SC5 was able to discriminate Gram-negative bacteria.

In terms of spectral information, the principal components PC2 and PC3 vectors (not shown) have spectral features related to Raman peaks of carotenoids, which can be used to separate *S. aureus* from other micro-organisms. Gram-negative bacteria were classified using SC4 and SC5. The PC4 and PC5 vectors (not shown) presented Raman peaks at around 900, 1000, 1100, 1200, 1300, 1400, 1470, 1530, and 1600 cm^{-1}, and the discrimination can be related to differences in the biochemical constitution of cell walls (peptidoglycan and teichoic or teichuronic acids) and other cell structure constituents of *E. coli*, *K. pneumoniae*, *P. mirabilis, E. cloacae*, and *P. aeruginosa*, which enabled differentiation of bacteria species through Raman spectroscopy.

Table 2 lists the rate of successful classification of clinical strains through PCA and Mahalanobis distance analyses. The percentage of correct discrimination of Gram-positive bacteria, or the sensitivity, was 100%, including the differentiation of species. The percentage of true negative, or the specificity, was also 100%, including the differentiation of species. The sensitivity values for the discrimination of Gram-negative bacteria was 64.0% for *E. coli*, 58.3% for *E. cloacae*, 88.5% for *K. pneumoniae*, 87.1% for *P. mirabilis,* and 85.0% for *P. aeruginosa*. The specificity values for each species were 95.4%, 98.1%, 93.0%, 99.0%, and 87.0%, respectively. The overall accuracy of the discrimination model was 83%.

The high specificity of the discrimination model is a positive factor for the practical application of bacterial identification

Table 2 Results of classification of the Raman spectra of clinical strains of bacteria using PCA/Mahalanobis distance

Biochemical identification (gold-standard)	PCA/Mahalanobis identification						
	Gram-positives		Gram-negatives				
	S. aureus	*Enter.* spp.	*E. coli*	*E. cloacae*	*K. pneum.*	*P. mirab.*	*P. aerug.*
S. aureus (22)	22	0	0	0	0	0	0
Enteroccus spp. (25)	0	25	0	0	0	0	0
E. coli (25)	0	0	16	0	0	1	8
E. cloacae (24)	0	0	1	14	6	0	3
K. pneumoniae (35)	0	0	0	2	31	0	2
P. mirabilis (31)	0	0	2	0	0	27	2
P. aeruginosa (20)	0	0	2	0	1	0	17
Sensitivity	100%	100%	64%	58%	88%	87%	85%
Specificity	100%	100%	95%	98%	93%	99%	87%

through Raman spectroscopy. For a disease of infectious origin, the correct identification of the microbial agent is important, but the correct exclusion of a microorganism from being involved in the process is essential for an appropriate therapeutic approach, avoiding administration of antimicrobials to patients without such infectious agents.

In this study, we found that the bacteria's spectrum can be influenced by the medium used in the culture. *E. coli* spectra presented different features depending on the medium that could be important in the PCA discrimination algorithm, leading to errors in grouping the bacterium genus if its influence is not considered (mainly when using CLED, with strong peaks at 886, 1480, and 1620 cm^{-1}). To minimize the impact of such differences, a standardization of the medium used in each microorganism is needed, so that the PCA algorithm could handle with these medium-related Raman bands.

A previous study done by Giana et al.[26] showed that three pathogenic bacteria strains (*E. coli*, *E. faecalis*, and *S. aureus*) could be selectively discriminated though an algorithm based on PCA and Euclidean distance applied to fluorescence spectra using selected excitation wavelength in the visible region. Buijtels et al.[27] evaluated differences in the Raman spectra of species of *Mycobacterium tuberculosis* and nontuberculous mycobacteria using hierarchical cluster analysis (HCA) for identification purposes, reaching sensitivity of 95%. Maquelin et al.[28] employed Raman spectroscopy and HCA for typing different species of *Acinetobacter*, concluding that the grouping by Raman/HCA is correlated to the results found using molecular biology techniques. Kirschner et al.[11] discriminated enterococci at the species level by using vibrational techniques correlated to the molecular biology techniques: the results of phenotypic method were correlated to the Fourier-Transform Infrared (FT-IR) spectroscopy with discrepancies for certain strains; genotypic methods confirmed the results obtained by FT-IR. In this work, we showed that it is possible to increase the number of strains from three to seven and use the vibrational technique Raman spectroscopy directly in the bacteria colony, grouping these strains according to spectral differences using PCA. Studies are under way to increase the number of strains to the 15 most common found in the urinary tract, together with an automated spectral collection and processing, being closer to the clinical setup.

Advantages of the Raman technique include there being no need for further sample preparation, that results can be obtained rapidly through Raman probes with small amounts of biomass, and the reproducibility. These characteristics are important for an early identification of clinical strains, allowing a most appropriate treatment for the patient and reducing the occurrence of bacterial resistance due to inaccurate empirical therapy, usually employed due to the delay in identification of microorganisms by classical microbiological techniques. This would reduce the morbidity and mortality related to serious infections. Thus, Raman spectroscopy could be a promising technique as a rapid method for bacterial species identification in microbiology.

4 Conclusion

Raman spectroscopy was able to identify spectral differences related to the biochemical content of seven different clinical species of bacteria involved in urinary tract infections (*S. aureus*, *Enterococcus* spp., *E. coli, K. pneumoniae*, *E. cloacae*, *P. mirabilis,* and *P. aeruginosa*), which were grown on Mueller Hinton and blood agars. A discriminant model based on PCA and Mahalanobis distance applied to Raman spectra collected *in vitro* was able to group those samples with good sensitivity and high specificity, as well as separate two Gram-positive species according to differences in the cell wall components (*Enterococcus* spp.) and presence of carotenoids (*S. aureus*), with 100% sensitivity and specificity. The Raman spectral information proved to be reproducible using a small amount of biomass, demonstrating it to be a fast, low cost, and easily implemented technique for an early identification of pathogenic microorganisms.

Acknowledgments

L. Silveira Jr. gratefully acknowledges FAPESP (São Paulo Research Foundation) (Proc. No. 2009/01788-5) for their partial financial support.

References

1. E. H. Ibrahim et al., "The influence of inadequate antimicrobial treatment of bloodstream infections on patient outcomes in the ICU setting," *Chest* **118**(1), 146–155 (2000).
2. H. K. Walker, W. D. Hall, and J. W. Hurst, *Clinical Methods: The History, Physical, and Laboratory Examinations*, pp. 999–1004, Butterworths, Boston (1990).
3. D. I. Andersson and D. Hughes, "Antibiotic resistance and its cost: is it possible to reverse resistance?," *Nat. Rev. Microbiol.* **8**(4), 260–271 (2010).
4. J. J. Kerremans et al., "Rapid identification and antimicrobial susceptibility testing reduce antibiotic use and accelerate pathogen-directed antibiotic use," *J. Antimicrob. Chemother.* **61**(2), 428–435 (2008).
5. K. Maquelin et al., "Identification of medically relevant microorganisms by vibrational spectroscopy," *J. Microbiol. Meth.* **51**(3), 255–271 (2002).
6. K. Maquelin et al., "Prospective study of the performance of vibrational spectroscopies for rapid identification of bacterial and fungal pathogens recovered from blood cultures," *J. Clin. Microbiol.* **41**(1), 324–329 (2003).
7. L. P. Choo-Smith et al., "Investigating microbial (micro)colony heterogeneity by vibrational spectroscopy," *Appl. Environ. Microbiol.* **67**(4), 1461–1469 (2001).
8. W. E. Huang et al., "Single-cell Raman spectral profiles of Pseudomonas fluorescens SBW25 reflects in vitro and in planta metabolic history," *Microb. Ecol.* **53**(3), 414–425 (2007).
9. W. H. Nelson, R. Manoharan, and J. F. Sperry, "UV resonance Raman studies of bacteria," *Appl. Spectrosc.* **27**(1), 67–124 (1992).
10. D. Naumann et al., "FT-IR spectroscopy and FT-Raman are powerful analytical tools for the non-invasive characterization of intact microbial cells," *J. Mol. Struct.* **347**, 399–405 (1995).
11. C. Kirschner et al., "Classification and identification of Enterococci: a comparative phenotypic, genotypic, and vibrational spectroscopic study," *J. Clin. Microbiol.* **39**(5), 1763–1770 (2001).
12. R. M. Jarvis and R. Goodacre, "Discrimination of bacteria using surface-enhanced Raman spectroscopy," *Anal. Chem.* **76**(1), 40–47 (2004).
13. R. M. Jarvis and R. Goodacre, "Characterization and identification of bacteria using SERS," *Chem. Soc. Rev.* **37**(5), 931–936 (2008).
14. G. J. Puppels et al., "Raman spectroscopic method for identification of clinically relevant microorganisms growing on solid culture medium," *Anal. Chem.* **72**(1), 12–19 (2000).
15. W. E. Huang et al., "Raman microscopic analysis of single microbial cells," *Anal. Chem.* **76**(15), 4452–4458 (2004).
16. L. Silveira et al., "Discrete wavelet transform for denoising Raman spectra of human skin tissues used in a discriminant diagnostic algorithm," *Instrum. Sci. Technol.* **38**(4), 268–282 (2010).
17. C. A. Lieber and A. Mahadevan-Jansen, "Automated method for subtraction of fluorescence from biological Raman spectra," *Appl. Spectrosc.* **57**(11), 1363–1367 (2003).
18. J. R. Goodwin, "Vibrational Microspectroscopy of Bacterial Colonies," (MS Thesis), pp. 36–46 School of Chemical and Physical Sciences at Queensland University of Technology, Brisbane (2006).
19. J. Popp et al., "Micro-Raman spectroscopic identification of bacterial cells of genus Staphylococcus and dependence on their cultivation conditions," *Analyst* **130**(11), 1543–1550 (2005).
20. Z. Movasaghi, S. Rehman, and I. U. Rehman, "Raman spectroscopy of biological tissues," *Appl. Spectrosc. Rev.* **42**(5), 493–541 (2007).
21. J. Popp et al., "The influence of intracellular storage material on bacterial identification by means of Raman spectroscopy," *Anal. Bioanal. Chem.* **397**(7), 2929–2937 (2010).
22. D. Naumann et al., "Infrared spectroscopy, a tool for probing bacterial Peptidoglycan. Potentialities of infrared spectroscopy for cell wall analytical studies and rejection of models based on crystalline chitin," *Eur. J. Biochem.* **125**(3), 505–515 (1982).
23. K. Maquelin et al., "Raman spectroscopic typing reveals the presence of carotenoids in Mycoplasma pneumonia," *Microbiol.* **155**(6), 2068–2077 (2009).
24. S. Efrima et al., "Surface-enhanced Raman spectroscopy as a tool for probing specific biochemical components in bacteria," *Appl. Spectrosc.* **58**(1), 33–40 (2004).
25. M. L. Paret et al., "Biochemical characterization of Gram-positive and Gram-negative plant-associated bacteria with micro-Raman spectroscopy," *Appl. Spectrosc.* **64**(4), 433–441 (2010).
26. H. E. Giana et al., " Rapid identification of bacterial species by fluorescence spectroscopy and classification through principal components analysis," *J. Fluoresc.* **13** (6), 489–493 (2003).
27. P. C. Buijtels et al., " Rapid identification of mycobacteria by Raman spectroscopy," *J. Clin. Microbiol.* **46**(3), 961–965 (2008).
28. K. Maquelin et al., "Rapid epidemiological analysis of Acinetobacter strains by Raman spectroscopy," *J. Microbiol. Meth.* **64**(1), 126–131 (2006).