View Article Online / Journal Homepage / Table of Contents for this issue

Micro-Raman spectroscopic identification of bacterial cells of the genus *Staphylococcus* and dependence on their cultivation conditions

M. Harz,[a] P. Rösch,[a] K.-D. Peschke,[b] O. Ronneberger,[b] H. Burkhardt[b] and J. Popp*[a]

Received 1st June 2005, Accepted 31st August 2005
First published as an Advance Article on the web 30th September 2005
DOI: 10.1039/b507715j

Published on 30 September 2005. Downloaded by Michigan State University on 16/01/2016 13:34:20.

Microbial contamination is not only a medical problem, but also plays a large role in pharmaceutical clean room production and food processing technology. Therefore many techniques were developed to achieve differentiation and identification of microorganisms. Among these methods vibrational spectroscopic techniques (IR, Raman and SERS) are useful tools because of their rapidity and sensitivity. Recently we have shown that micro-Raman spectroscopy in combination with a support vector machine is an extremely capable approach for a fast and reliable, non-destructive online identification of single bacteria belonging to different genera. In order to simulate different environmental conditions we analyzed in this contribution different *Staphylococcus* strains with varying cultivation conditions in order to evaluate our method with a reliable dataset. First, micro-Raman spectra of the bulk material and single bacterial cells that were grown under the same conditions were recorded and used separately for a distinct chemotaxonomic classification of the strains. Furthermore Raman spectra were recorded from single bacterial cells that were cultured under various conditions to study the influence of cultivation on the discrimination ability. This dataset was analyzed both with a hierarchical cluster analysis (HCA) and a support vector machine (SVM).

Introduction

Fast and exact identification of microorganisms is becoming an important challenge in various fields of research and industry, *e.g.*, reliable and rapid methods are needed for the characterization of relevant microorganisms in medical diagnostics, pharmaceutical production or food processing technology. Routine bacterial identification of pathogenic microorganisms is largely based on a morphological evaluation of the microorganism, growing them in various media and a set of biochemical tests.[1] Consequently, this procedure takes at least a few hours or even longer than a day.[2] Therefore novel and fast analytical techniques are necessary, making an extensive, reliable and fast identification of microbes possible. Concerning these requirements, new analysis methods such as mass spectrometry, polymerase chain reaction (PCR), flow cytometry or fluorescence spectroscopy were developed.[1,2] Alternatively vibrational spectroscopic techniques such as Fourier transform infrared spectroscopy (FT-IR)[3–7] and Raman spectroscopy[8–11] in combination with chemometric procedures[12,13] or neural networks (NN)[14,15] are suitable tools for the rapid identification of bacteria.[5] The IR and vibrational spectra provide biochemical information on the molecular composition of the studied microorganisms.[3,8] Spectral signals can be used to differentiate between several microbial species and strains in bulk material or micro colonies.[16–21]

Recently we have shown that micro-Raman spectroscopy in combination with a support vector machine is an extremely capable approach for a fast and reliable, non-destructive online identification of single bacteria belonging to different genera.[22] It could be shown that for the identification of single bacterial cells, one Raman spectrum is sufficient if a powerful database and appropriate evaluation software are available. In contrast yeasts as eukaryotes differ from prokaryotes such as bacteria in their molecular composition. They exhibit certain organelles and therefore they show a spatial heterogeneity. Hence one spectrum is not enough to describe the whole yeast cell. Using a mean spectrum over several measured positions inside the yeast cell it is feasible to distinguish yeasts on a single cell level.[23]

With the help of this technique it is possible to identify both bacteria and yeasts on a single cell level by means of micro-Raman spectroscopy. By investigating microorganisms under real environmental conditions the origin of these cells is unknown, *e.g.* nutrition, temperature or cell age. In order to simulate this situation, in this contribution different *Staphylococcus* strains with varying cultivation conditions were analyzed in order to evaluate this method with a reliable database. The species *Staphylococcus cohnii*, *Staphylococcus epidermidis* and *Staphylococcus warneri* were grown under different cultivation conditions such as nutrition, temperature and culture age. This dataset is analyzed both with a supervised and an unsupervised technique.

In a first attempt, micro-Raman measurements were performed to obtain spectra from bulk material to prove the principal feasibility of the method. For a distinct

[a]*Institut für Physikalische Chemie, Friedrich-Schiller-Universität Jena, Helmholtzweg 4, 07743 Jena, Germany. E-mail: juergen.popp@uni-jena.de; Fax: +49 3641 948302; Tel: +49 3641 948320*
[b]*Lehrstuhl für Mustererkennung und Bildverarbeitung, Institut für Informatik, Albert-Ludwigs-Universität Freiburg, Georges-Koehler-Allee Geb. 052, D-79110 Freiburg, Germany*

View Article Online

Published on 30 September 2005. Downloaded by Michigan State University on 16/01/2016 13:34:20.

identification of the strains a chemometric data analysis was performed.

In further experiments micro-Raman spectra were recorded from single bacterial cells grown under the same conditions as bulk material for a discrimination at a species and strain level. Since signals of single bacteria have more distinctive variabilities[22–26] than those of bulk material a complete dataset including all possible variations of cultivation conditions is needed. A varying culturing environment leads to a variance in the biochemical composition of a microbial cell and could therefore affect the ability to discriminate and identify the investigated species.[27–29] Therefore Raman spectra were recorded from single bacterial cells that were grown under several cultivation conditions with respect to the nutrient medium, incubation temperature and culturing age. These spectra were used to discriminate and classify bacterial species and strains and to analyze the effect of growth dependent distributions of the biochemical compositions of bacterial cells on the distinction capability.

Experimental

Sample preparation

The microorganisms of the species *Staphylococcus cohnii* DSM 6669, DSM 6718, DSM 6719 and DSM 20260, *Staphylococcus warneri* DSM 20036 and DSM 20316 and *Staphylococcus epidermidis* DSM 1798 were obtained from DSMZ (Deutsche Sammelstelle für Mikroorganismen und Zellkulturen). *Staphylococcus epidermidis* ATCC 35984 (American Type Culture Collection) was purchased from the Institut für Infektionsbiologie, Universität Würzburg.

Bacterial cells were cultivated on CA (corynebacterium agar) or CASO (trypticase soy yeast extract medium) plates under varying cultivation conditions with respect to nutrient medium, growing time and temperature. Table 1 gives an overview of the investigated bacteria. In initial experiments the micro-Raman measurements were performed for bulk samples and single bacterial cells cultured on CA-agar or CASO-agar at a temperature of 37 °C following the recommended conditions listed in the DSMZ. For further single bacterial experiments the strains were grown on different media to that listed in the DSMZ. In an additional experiment bacteria were grown on the suggested media but under a different temperature of 30 °C instead of 37 °C.

For a Raman analysis the grown cells were taken from the agar plates and smeared by a diluting loop on a fused silica plate. Fig. 1 shows an image of single bacterial cells smeared on a fused silica plate. The dark roundly shaped features correspond to single bacterial cells. The light features are caused by the blurred extra cellular matrix of the bacteria.

Table 1 Selected bacteria strains for micro-Raman spectroscopy with sample labels

Sample	
S. cohnii subsp. cohnii DSM 6669	A
S. cohnii subsp. urealyticum DSM 6718	B
S. cohnii subsp. urealyticum DSM 6719	C
S. cohnii subsp. cohnii DSM 20260	D
S. warneri DSM 20036	E
S. warneri DSM 20316	F
S. epidermidis ATCC 35984	G
S. epidermidis DSM 1798	H

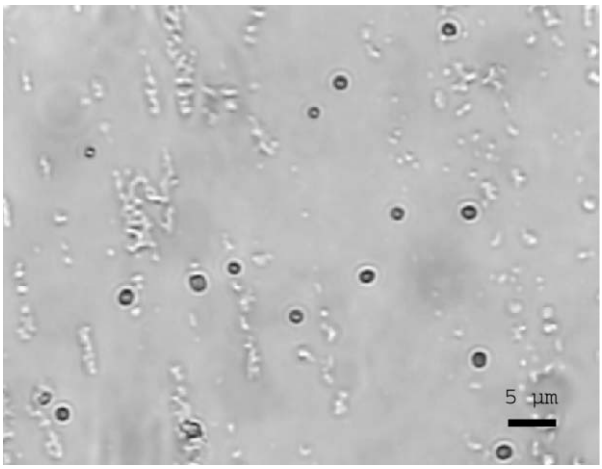

Fig. 1 Microscopic image of a smear of single bacterial cells of *S. cohnii* DSM 6669 on a fused silica plate. The dark roundly shaped features correspond to single bacterial cells and the light features to the blurred extra cellular matrix of the bacteria.

Spectroscopic instrumentation

The Raman spectra of the bacteria were recorded with a micro-Raman setup (HR LabRam invers system, Jobin Yvon, Horiba). Raman scattering was excited by a frequency doubled Nd:YAG laser at a wavelength of 532 nm with a laser power of about 10 mW incident on the sample. The laser beam was focused on individual cells or a bacterial layer by means of a Leica PLFluoar x100/0.75 microscope objective down to a spot diameter of approximately 0.7 µm which is sufficient to resolve single bacteria from the background. The dispersive spectrometer has an entrance slit of 100 µm, a focal length of 800 mm and is equipped with a grating of 300 lines mm^{-1}. The Raman scattered light was detected by a CCD camera operating at 220 K.

Raman spectra of bacterial bulk samples were measured with an integration time of about 10–190 s depending on the fluorescence background. Thereby for each spectrum a new area of the bulk sample was investigated. For the single cell analysis for every spectrum a new cell was measured and an acquisition time of 60 s was used. The data were acquired over the course of several days. For the calibration procedure of the spectrometer a routine check was performed on a daily basis using titanium dioxide as a reference control.

Chemotaxonomic analysis

An unsupervised classification method, hierarchical cluster analysis, was applied to analyze both the bulk spectra and the single bacterial spectra cultivated under recommended conditions. This was done with the program OPUS IDENT from Bruker. Spectral preprocessing was also performed with the OPUS software from Bruker (first derivative, 13 points). For a distinct identification of single bacteria grown under different

 This journal is © The Royal Society of Chemistry 2005

View Article Online

Published on 30 September 2005. Downloaded by Michigan State University on 16/01/2016 13:34:20.

cultivation conditions a supervised classification method, a support vector machine, was used.[22,30]

Results and discussion

Raman spectra from bulk samples

In a first attempt, Raman measurements within a multilayer region of the smear were recorded in order to obtain Raman bulk spectra. Raman spectra of bacterial bulk samples are often masked by the appearance of fluorescence. About 148 spectra of the bulk material were recorded. Fig. 2 shows one representative micro-Raman spectrum of a bacterial layer for each *Staphylococcus* strain. It can be seen that the spectra are very similar and reveal a high fluorescence background. The spectrum of *S. cohnii* DSM 6669 (A) exhibits more pronounced signals over the whole spectral region in contrast to the other spectra of *S. cohnii* DSM 6718 (B), *S. cohnii* DSM 6719 (C), *S. cohnii* DSM 20260 (D), *S. warneri* DSM 20036 (E), *S. warneri* DSM 20316 (F), *S. epidermidis* ATCC 35984 (G) and *S. epidermidis* DSM 1798 (H). In the Raman spectrum of *S. cohnii* DSM 6669 (A) there are three dominant bands at 746 cm^{-1}, 1125 cm^{-1} and 1575 cm^{-1}. The peak around 1125 cm^{-1} could be assigned to the ν (C–O–C) stretching vibration from symmetric glycosidic linkages[31] or to ν (C–N) and ν (C–C) stretching vibrations.[11] The peak at 1575 cm^{-1} might be due to the δ (NH) deformation vibration and the ν (CN) stretching vibration.[31,32] These bands are not so pronounced in the spectra of the other examined bacterial bulk samples. However the prominent C–H stretching band region in the range of 2700–3100 cm^{-1} with the ν (CH_2) asymmetric stretching vibration around 2935 cm^{-1} exhibits similar intensities for all investigated strains (Fig. 2). This major C–H stretching band and the peak in the region of 1440–1460 cm^{-1} corresponding to a δ (CH_2) scissoring vibration originate from the CH bindings in lipids, proteins and carbohydrates.[32] Another important peak in the range of 1650–1680 cm^{-1} could be attributed to the amide I band.[32] Table 2 provides an overview of the observed Raman bands of the investigated bacterial cells together with a tentative assignment.

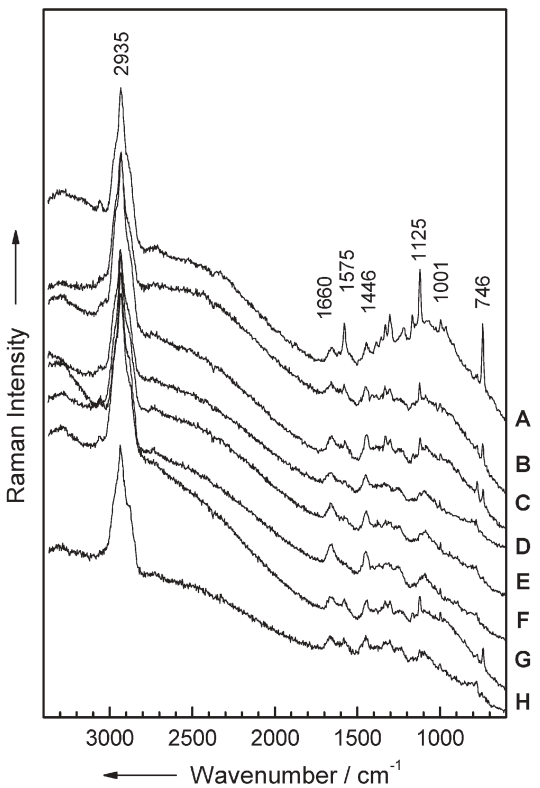

Fig. 2 Raman spectra of bacterial bulk samples of different *Staphylococcus* strains that were smeared on fused silica plates, recorded with an excitation wavelength of 532 nm. The investigated bacterial strains were *S. cohnii* DSM 6669 (A), *S. cohnii* DSM 6718 (B), *S. cohnii* DSM 6719 (C), *S. cohnii* DSM 20260 (D), *S. warneri* DSM 20036 (E), *S. warneri* DSM 20316 (F), *S. epidermidis* ATCC 35984 (G) and *S. epidermidis* DSM 1798 (H) cultured under the recommended conditions listed in the DSMZ.

Table 2 Raman bands observed in spectra of bacterial cells and tentative assignment

Wavenumber/cm^{-1}	Assignment[a]	Reference
~722	ρ (CH_2)	32
746		
~778	Cytosine, uracil ring stretching	11
850	ν (CC) Ring breathing	32
	ν (COC) 1,4-Glycosidic link	32
1001	ν (CC) Aromatic ring breathing (phenylalanine)	20
1125	ν (COC) Symmetric glycosidic link	31
	ν (C–N) and ν (C–C)	11
1200–1280	Amide III	31, 32
~1327	δ (C–H)	20
1440–1460	δ (CH_2) Scissoring	31, 32
1575	Amide II	31, 32
	Guanine, adenine ring stretching	11
1650–1680	Amide I	32
2935	ν (CH_2) Asymmetric	32
3060	ν (CH) Olefinic	32

[a] δ: Deformation vibration, ν: stretching vibration, ρ: rocking vibration.

In order to classify the analyzed bacteria an unsupervised method, hierarchical cluster analysis, was performed. The spectra were pretreated by calculating the first derivative of the Raman spectra to eliminate disturbing baseline contributions. Furthermore the spectral range for the classification was changed and then investigated in order to evaluate the most suitable classes. The best results were obtained by carrying out the hierarchical cluster analysis in the spectral range between 850–1750 cm^{-1} and 2650–3150 cm^{-1}. The spectral distances between every spectrum were calculated with the Normalization to Reproduction Level algorithm. This method is based on a vector normalization that is calculated separately for each spectral range. Furthermore several spectral ranges can be weighted independently with a reproduction level. Hence the spectral distances are expressed in units of the reproduction level. Ward's Technique was used to calculate the spectral distances between a newly created cluster and all the other spectra or clusters. With the help of this cluster algorithm the most similar groups are not clustered but feasible most homogeneous clusters are formed. Fig. 3 shows the dendrogram of the resulting classification of the bulk samples based on 137 spectra of the eight different strains. The labels A to H for each cluster correspond to the

View Article Online

Published on 30 September 2005. Downloaded by Michigan State University on 16/01/2016 13:34:20.

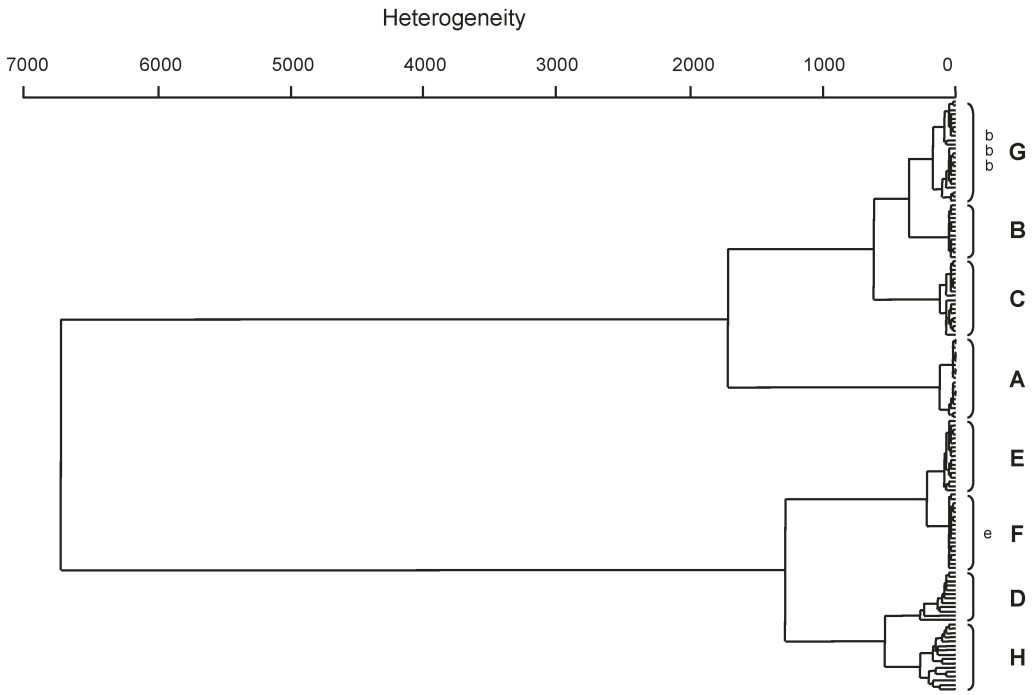

Fig. 3 Cluster analysis of the first derivative spectra using the wavenumber range 850–1750 cm^{-1} and 2650–3150 cm^{-1}. The dendrogram of Raman spectra resulting from HCA was calculated with the spectra of bacterial bulk samples of *S. cohnii* DSM 6669 (A), *S. cohnii* DSM 6718 (B), *S. cohnii* DSM 6719 (C), *S. cohnii* DSM 20260 (D), *S. warneri* DSM 20036 (E), *S. warneri* DSM 20316 (F), *S. epidermidis* ATCC 35984 (G) and *S. epidermidis* DSM 1798 (H). The bacterial bulk samples were cultivated under the recommended culture condition listed in the DSMZ.

one in Table 1. The smaller the spectral distances the more similar are the spectra. A discrimination on the strain level could be achieved. Nearly all spectra of one strain are classified to the same cluster and are well separated from the others. Altogether there is a distinction into two main clusters. One cluster consists of four strains that are *S. cohnii* DSM 6669 (A), *S. cohnii* DSM 6718 (B), *S. cohnii* DSM 6719 (C) and *S. epidermidis* ATCC 35984 (G). The other group is formed by *S. cohnii* DSM 20260 (D), *S. epidermidis* DSM 1798 (H), *S. warneri* DSM 20036 (E) and *S. warneri* DSM 20316 (F). Nevertheless the distinction into two clusters does not reflect the relationship between the species *S. cohnii*, *S. warneri* and *S. epidermidis* within subclusters. In Fig. 3 the incorrectly grouped spectra are displayed with small letters, where the small characters have the same attribution to the different strains as capital letters. Three spectra of *S. cohnii* subsp. urealyticum DSM 6718 (b) were incorrectly grouped to *S. epidermidis* ATCC 35984 (G). They were assigned to another species. Furthermore one spectra of *S. warneri* DSM 20036 (e) was not clustered to the equivalent strain but was grouped to the appropriate species. Nevertheless an identification accuracy of 97.0% (133/137) on the strain level could be obtained, while three of the four spectra were not clustered to the appropriate species leading to a distinction precision of 97.7% (134/137) on the species level.

Beside the hierarchical cluster analysis, a supervised method, the support vector machine, was applied to classify the analyzed bacteria. In contrast to the hierarchical cluster analysis, where spectra with spikes were excluded for the clustering, every measured spectrum was used for the classification by means of the support vector machine. Before classification *via* the support vector machine Raman spectra were preprocessed by running median filtering and by centering. The running median filter is realized as a window that moves over each spectrum while replacing the data point at the center of the window by the median calculated over all points that are covered by the window. The size of the window was chosen to be 9 spectral points. By the centering step the mean value of all feature points over all spectra was set to 0 and the standard deviation was set to 1. This was done to homogenize the distribution in all dimensions in the feature space. The classification is based on the wavenumber region of 330–3600 cm^{-1}. For the classification a nonlinear SVM with a Radial Basis Function (RBF)-Kernel was applied and for our data the one-*versus*-one test was used for training and classification. Only two parameters needed to be adjusted for SVM classification. Firstly, a value that controlled the local support of the radial basis function, called the gamma value and secondly, a cost value that punishes outliers during the training process. For the experiments the gamma value was set to 3.5×10^{-7} while the cost value was set to 2.6×10^{6}. For an estimation of the classification error probability of the final system (which used all N recorded spectra as training set) the leave-one-out error was chosen, which uses $N - 1$ samples as a training set. The reported "average recognition rate" is the arithmetic mean of the recognition rates for each strain and species and therefore equalizes the varying number of samples per strain and species in our database. Accomplishing a classification with 148 spectra of the eight different strains from three species, we obtain 141 correctly identified spectra on the strain level (average recognition rate: 94.9%) and 144 correctly identified spectra on the species level (average

 This journal is © The Royal Society of Chemistry 2005

View Article Online

Published on 30 September 2005. Downloaded by Michigan State University on 16/01/2016 13:34:20.

Table 3 Recognition rate for Raman spectra of bulk material

Sample	Total number of spectra	Number of wrongly classified strain spectra	Recognition rate for strains (%)	Number of wrongly classified species spectra	Recognition rate for species (%)
S. cohnii subsp. cohnii DSM 6669	20	2	90	1	98.9
S. cohnii subsp. urealyticum DSM 6718	16	1	93.8	0	100
S. cohnii subsp. urealyticum DSM 6719	18	1	94.4	0	100
S. cohnii subsp. cohnii DSM 20260	13	1	92.3	1	92.3
S. warneri DSM 20036	20	0	100	0	100
S. warneri DSM 20316	20	1	95	1	95
S. epidermidis ATCC 35984	25	0	100	0	100
S. epidermidis DSM 1798	16	1	93.8	1	93.8
Average recognition rate			**94.9**		**97.0**

recognition rate: 97.0%). All results are summarized in Table 3. It can be shown that the support vector machine for the classification of our dataset leads to a similar identification accuracy to a hierarchical cluster analysis. However by means of the support vector machine, spectra with spikes could also be used for identification because the data were median filtered.

Raman spectra from single cells

The Raman spectra of the bulk material result from the averaged signal over several bacteria. For the identification of single bacterial cells individual variations have to be taken into account. Therefore it is essential to investigate how different parameters influence the identification ability of microorganisms on a single cell level. Hence to create a reliable dataset the influence of several factors like nutrition, temperature and growing time need to be considered. Thus bacterial cells grown under several cultivation conditions with respect to the nutrient medium (CA-agar or CASO-agar), incubation temperature (30 °C or 37 °C) and culture age were investigated.

First, micro-Raman spectra of single bacteria grown under the DSMZ listed recommended conditions were recorded and used to discriminate and classify bacterial species and strains. Fig. 4 shows representative micro-Raman spectra of eight different strains. The Raman spectra of single cells show characteristic differences compared with the corresponding bulk spectra plotted in Fig. 2. They reveal a lower fluorescence background and additional signals in the range of 720–890 cm^{-1} and 950–1120 cm^{-1} due to the fused silica plate because of the very low sample volume of a single cell. Furthermore a peak at 1001 cm^{-1} corresponding to the ring breathing vibration of the amino acid phenylalanine[20] is more intense in the single cell spectra than in the bulk spectra. It is noteworthy that the three dominant bands at 746 cm^{-1}, 1125 cm^{-1} and 1575 cm^{-1} which appear in the bulk spectrum of *S. cohnii* DSM 6669 (A) (Fig. 2) are marginal or absent in the corresponding single bacterial spectrum (Fig. 4). However an additional band at 778 cm^{-1} that could be attributed to a cytosine and uracil ring stretching vibration[11] can be seen.

For a hierarchical cluster analysis the resulting dendrogram was calculated for the spectral range between 850–1750 cm^{-1} and 2650–3150 cm^{-1}. The spectral distances between every spectrum were calculated with the Normalization to Reproduction Level algorithm and the clustering was based on Ward's Technique. The spectra containing spikes were not considered for the classification. The dendrogram of the resulting classification of single bacteria spectra based on 143 spectra of eight different *Staphylococcus* strains cultivated under the same conditions as bacterial layers is shown in Fig. 5. There are two major clustering branches in the dendrogram. The first cluster consists of *S. warneri* DSM 20036 (E) and *S. warneri* DSM 20316 (F).

The second cluster is made up of two subgroups: one group consists of the four *S. cohnii* strains: *S. cohnii* DSM 6669 (A), *S. cohnii* DSM 6718 (B), *S. cohnii* DSM 6719 (C) and *S. cohnii* DSM 20260 (D) and the other subcluster contains *S. epidermidis* ATCC 35984 (G) and *S. epidermidis* DSM 1798 (H). Hence within this latter group, a clear separation

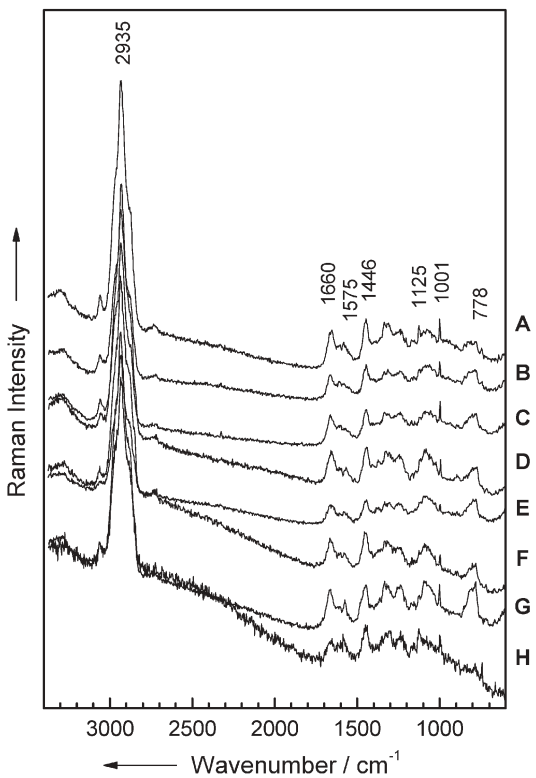

Fig. 4 Raman spectra of single bacterial cells of different *Staphylococcus* strains that were smeared on quartz plates recorded for an excitation wavelength of 532 nm. The analyzed bacterial strains were *S. cohnii* DSM 6669 (A), *S. cohnii* DSM 6718 (B), *S. cohnii* DSM 6719 (C), *S. cohnii* DSM 20260 (D), *S. warneri* DSM 20036 (E), *S. warneri* DSM 20316 (F), *S. epidermidis* ATCC 35984 (G) and *S. epidermidis* DSM 1798 (H) cultured under the recommended condition listed in the DSMZ.

View Article Online

Published on 30 September 2005. Downloaded by Michigan State University on 16/01/2016 13:34:20.

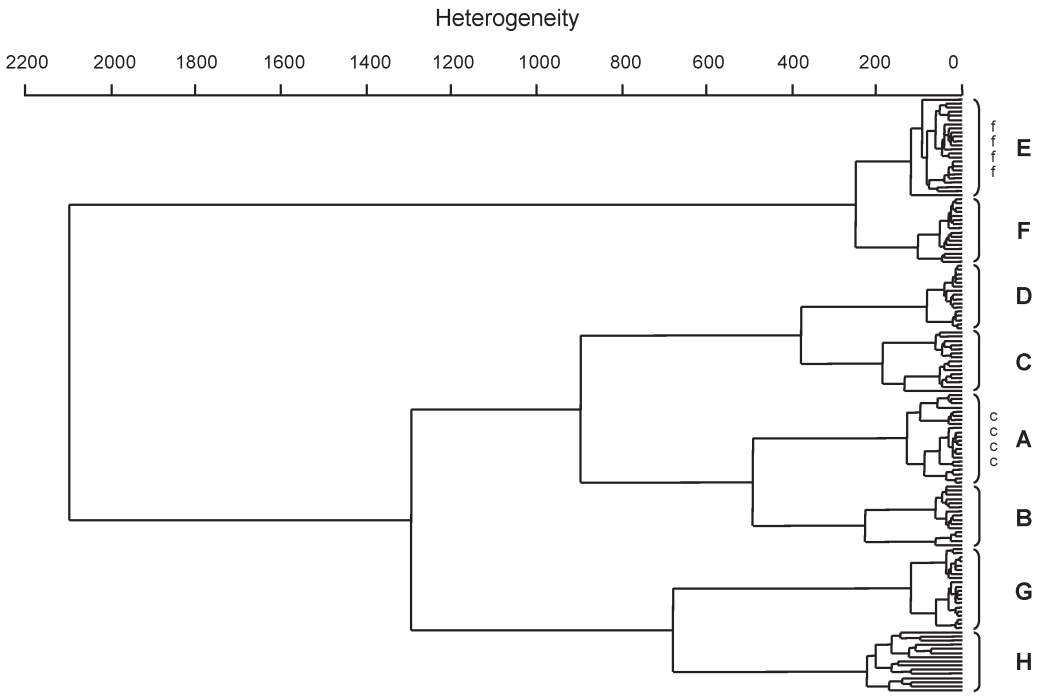

Fig. 5 Cluster analysis on the first derivative spectra for the wavenumber range 850–1750 cm^{-1} and 2650–3150 cm^{-1}. The dendrogram of Raman spectra resulting from HCA was determined with the spectra of single bacterial cells of *S. cohnii* DSM 6669 (A), *S. cohnii* DSM 6718 (B), *S. cohnii* DSM 6719 (C), *S. cohnii* DSM 20260 (D), *S. warneri* DSM 20036 (E), *S. warneri* DSM 20316 (F), *S. epidermidis* ATCC 35984 (G) and *S. epidermidis* DSM 1798 (H). The single bacterial cells were cultivated under the recommended culture conditions listed in the DSMZ.

between the species *S. cohnii* and *S. epidermidis* can be seen. The dendrogram exhibits a good separation of each strain and in contrast to the classification of the bulk samples also a discrimination of the species. As a result there is a distinction on the strain level and on the species level. The divisions into subclusters reflect the relationship between the species *S. cohnii*, *S. warneri* and *S. epidermidis* in contrast to the classification of the bulk spectra. This could be the consequence of the lower background signal of the single cell spectra compared to the bulk spectra. Nevertheless eight samples out of 143 were not clustered to the correct strain. In Fig. 5 the incorrectly grouped spectra are displayed with small letters, where the small characters have the same attribution to the different strains as capital letters. Four spectra of *S. warneri* DSM 20316 (f) were incorrectly grouped to *S. warneri* DSM 20036 (E) however they were assigned to the same species. Furthermore four spectra of *S. cohnii* subsp. urealyticum DSM 6719 (c) were not clustered to the correctt strain, but they were grouped to the appropriate species. Nevertheless an identification accuracy of 94.9% (135/143) on the strain level could be obtained while all spectra were clustered to the appropriate species leading to a distinction precision of 100% on the species level.

These experiments were performed on single bacterial cells cultivated under the recommended culture conditions listed in the DSMZ. However in real applications we do not know the origin of the bacteria. So it was necessary to analyze the effect of different growing circumstances to increase the possible variability of bacteria in authentic applications. The effect of culturing conditions on the identification capability of single bacterial cells for some representative examples was investigated. Therefore bacterial cells were cultivated on different media to those listed in the DSMZ; strains that are recommended to be grown on CA were grown in this experiment on CASO and *vice versa*. In an additional experiment, bacteria were grown under a lower temperature than that recommended from the DSMZ; here the strains were grown at 30 °C.

Fig. 6A shows for each *Staphylococcus* strain representative Raman spectra of *S. cohnii* DSM 20260 and of *S. warneri* DSM 20316 obtained after culturing under several growing conditions. The spectra of the single cells from cultures grown under a lower temperature (30 °C instead of 37 °C) exhibit nearly the same background signal with bands as broad as the bacterial spectra grown under 37 °C. The same effect is observable in the Raman spectra of bacteria grown on CASO-agar instead of a CA-agar and reversed respectively. In general Raman spectra recorded from cells cultured under different conditions exhibit nearly equivalent Raman bands. This is exemplarily shown in Fig. 6A for *S. cohnii* DSM 20260 and for *S. warneri* DSM 20316. All spectra exhibit a characteristic band at 778 cm^{-1}. This peak occurs almost always for *S. warneri* DSM 20036 and *S. warneri* DSM 20316 grown on CASO-agar at 37 °C and CA-agar at 30 °C. This peak cannot be found for the bacteria of *S. cohnii* DSM 6669, *S. cohnii* DSM 6719 and *S. cohnii* DSM 20260 under several growing circumstances.

Additionally the influence of the growing time on the Raman spectrum of single *S. epidermidis* ATCC 35984 cells was evaluated. The first Raman spectra were recorded after only 6 hours incubation time. New samples were prepared and measured in intervals of 6 hours to a total incubation time of

 This journal is © The Royal Society of Chemistry 2005

View Article Online

Published on 30 September 2005. Downloaded by Michigan State University on 16/01/2016 13:34:20.

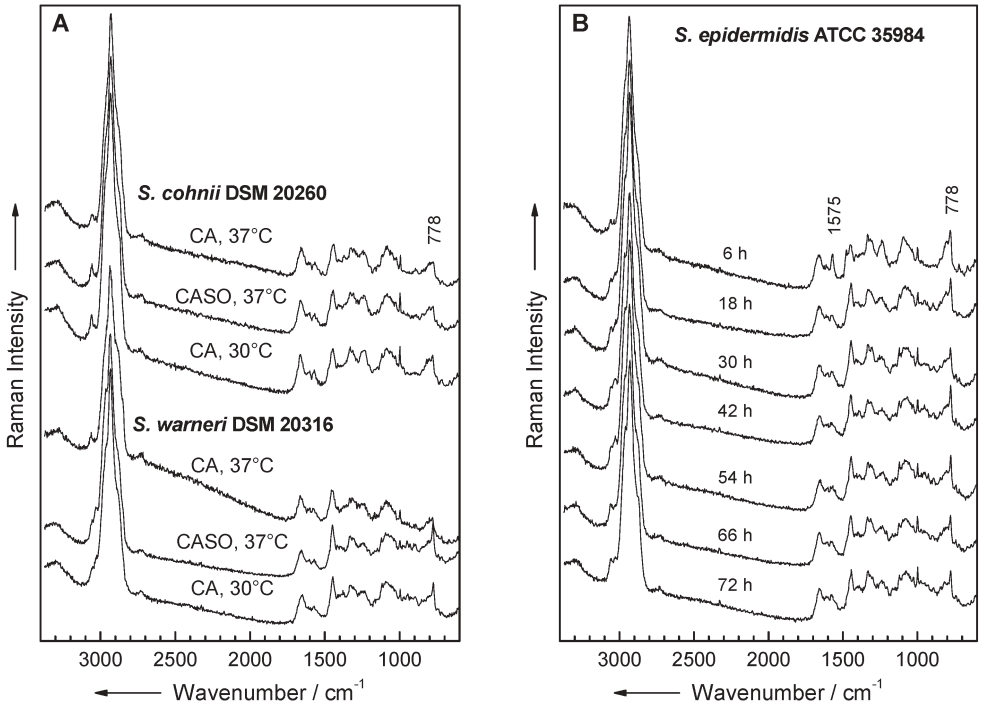

Fig. 6 A: Representative Raman spectra of single cells of *S. cohnii* DSM 20260 and *S. warneri* DSM 20316 obtained after growth on CA-agar or CASO-agar at 37 °C and CA-agar at 30 °C are shown. B: Raman spectra of single *S. epidermidis* ATCC 35984 cells cultured on CASO-agar at 37 °C were recorded for different growth times as indicated.

72 hours. In Fig. 6B representative Raman spectra of single *S. epidermidis* ATCC 35984 cells are displayed that were measured after growing on CASO-agar for different incubation times as indicated. Micro-Raman spectra recorded from cells cultured under these conditions exhibit nearly the same Raman bands, but there is one band in the region about 1575 cm^{-1} that decreases with increasing culturing time. This effect was also observable by examining single *S. epidermidis* ATCC 35984 cells grown on CA-agar at 37 °C and for a lower temperature of 30 °C, not shown here. In contrast to the other strains, every spectrum of *S. epidermidis* ATCC 35984 exhibits a band at 778 cm^{-1} independently of the culture conditions. It can be observed that the spectra of single bacterial cells from older cultures exhibit a better SNR. For an unambiguous analysis on a single cell level, these variations need to be taken into account by including these variations into the applied database for the chemotaxonomic identification.

When this complex data set was applied to the HCA, no clear separation of the species in separate clusters was possible, nor were clusters formed that are based on the same nutrient medium or incubation temperature. Recently we have shown that micro-Raman spectroscopy in connection with a support vector machine is able to identify single bacteria on a strain level.[22] Therefore a support vector machine was used to achieve a classification of single bacteria spectra grown under different culturing conditions. The classification was based on the same pretreatment of the spectra and the same parameters as were used for the identification of bulk material while the gamma value was set to 1×10^{-6} and the cost value to 1×10^{6}. Performing a classification with 1280 spectra of the eight different strains from three species we obtain 1248 correctly identified spectra on the strain level (average recognition rate: 94.1%) and 1266 correctly identified spectra on the species level (average recognition rate: 97.6%). The lowest recognition rate for strains is for *S. cohnii* DSM 6718 with 83.6% and for species for *S. cohnii* DSM 6718 with 90.2%. All results are summarized in Table 4. With these results it can be shown that micro-Raman spectroscopy in combination with support vector machines could be an appropriate approach for a rapid identification of single bacterial cells on the strain level.

Conclusions

In this contribution a chemotaxonomic study of a micro-Raman spectroscopic identification of bacterial bulk material and single cells of different species within the genus *Staphylococcus* is presented. In our experiments, spectra of bulk samples reveal a higher fluorescence background signal in comparison to the corresponding single cell spectra but nevertheless a distinction of strains is possible when applying a support vector machine and a hierarchical cluster analysis although the relationship on the species level is not reflected correctly with the latter method.

When going from a bulk environment to a single cell analysis, individual cell variations need to be considered in order to prove the ability to discriminate single bacteria on a strain level. For this reason, different culturing methods, which include various growing times, temperatures and different agar types, are used to maximize the variation in the single strains. When applying a hierarchical cluster analysis to single bacteria spectra where the strains were grown under one culture condition a distinction of the species is possible

View Article Online

Published on 30 September 2005. Downloaded by Michigan State University on 16/01/2016 13:34:20.

Table 4 Recognition rate for Raman spectra of single bacteria that were grown under different culture conditions

Sample	Total number of spectra	Number of wrongly classified strain spectra	Recognition rate for strains (%)	Number of wrongly classified species spectra	Recognition rate for species (%)
S. cohnii subsp. cohnii DSM 6669	62	0	100	0	100
S. cohnii subsp. urealyticum DSM 6718	61	10	83.6	6	90.2
S. cohnii subsp. urealyticum DSM 6719	63	8	87.3	2	96.8
S. cohnii subsp. cohnii DSM 20260	65	3	95.4	0	100
S. warneri DSM 20036	65	2	96.9	1	98.5
S. warneri DSM 20316	67	7	89.6	3	95.5
S. epidermidis ATCC 35984	785	2	99.7	2	99.7
S. epidermidis DSM 1798	112	0	100	0	100
Average recognition rate			**94.1**		**97.6**

and reflects the relationship of species and subspecies. After cultivating bacteria under various methods a hierarchical cluster analysis was not successful for classification. Therefore a support vector machine was used to identify the whole dataset. By applying 1280 spectra for the classification an average recognition rate of 94.1% on the strain level and 97.6% on the species level was achieved. These results make us hopeful that the combination of Raman spectroscopy with a support vector machine is an extremely capable method of identifying single bacteria of different cultivation conditions not only on the species level but also on the strain level.

For the identification of single bacteria by means of Raman spectroscopy and support vector machines for real applications the origin of the cells is often unknown. As could be shown, varying cultivation conditions have an influence on the Raman spectra, nevertheless for our dataset the identification of the single cell spectra with a support vector machine is suitable.

Acknowledgements

The funding of the research project FKZ 13N8365 and 13N8369 within the framework 'Biophotonik' from the Federal Ministry of Education and Research, Germany (BMBF) is gratefully acknowledged.

References

1 D. Ivnitski, I. Abdel-Hamid, P. Atanasov and E. Wilkins, *Biosens. Bioelectron.*, 1999, **14**, 599–624.
2 S. F. Al-Khaldi and M. M. Mossoba, *Nutrition*, 2004, **20**, 32–38.
3 D. Naumann, D. Helm and H. Labischinski, *Nature*, 1991, **351**, 81–82.
4 D. Helm, H. Labischinski, G. Schallehn and D. Naumann, *J. Gen. Microbiol.*, 1991, **137**, 69–79.
5 D. Naumann, S. Keller, D. Helm, C. Schultz and B. Schrader, *J. Mol. Struct.*, 1995, **347**, 399–405.
6 D. Naumann, in *Encyclopedia of Analytical Chemistry*, ed. R. A. Meyers, John Wiley & Sons, Chichester, 2000, pp. 102–131.
7 C. Kirschner, K. Maquelin, P. Pina, N. A. N. Thi, L. P. Choo-Smith, G. D. Sockalingum, C. Sandt, D. Ami, F. Orsini, S. M. Doglia, P. Allouch, M. Mainfait, G. J. Puppels and D. Naumann, *J. Clin. Microbiol.*, 2001, **39**, 1763–1770.
8 G. J. Puppels, W. Colier, J. H. F. Olminkhof, C. Otto, F. F. M. De Mul and J. Greve, *J. Raman Spectrosc.*, 1991, **22**, 217–25.
9 J. Greve and G. J. Puppels, in *Biomolecular Spectroscopy*, ed. R. E. Hester and R. J. H. Clark, John Wiley & Sons, Chichester, 1992, pp. 231–265.
10 S. J. Webb, *Phys. Rep.*, 1980, **60**, 201–24.
11 K. Maquelin, C. Kirschner, L. P. Choo-Smith, N. van den Braak, H. P. Endtz, D. Naumann and G. J. Puppels, *J. Microbiol. Methods*, 2002, **51**, 255–71.
12 B. K. Lavine, *Anal. Chem.*, 1998, **70**, 209R–228R.
13 E. C. Lopez-Diez and R. Goodacre, *Anal. Chem.*, 2004, **76**, 585–591.
14 R. Goodacre, R. Burton, N. Kaderbhai, E. M. Timmins, A. Woodward, P. J. Rooney and D. B. Kell, *NATO ASI Ser., Ser. 1*, 2000, **30**, 111–136.
15 R. Goodacre, E. M. Timmins, R. Burton, N. Kaderbhai, A. M. Woodward, D. B. Kell and P. J. Rooney, *Microbiology (Reading, UK)*, 1998, **144**, 1157–1170.
16 N. A. Ngo-Thi, C. Kirschner and D. Naumann, *J. Mol. Struct.*, 2003, **661–662**, 371–380.
17 N. A. Ngo Thi, C. Kirschner and D. Naumann, in *Spectroscopy of Biological Molecules: New Directions, 8th European Conference*, ed. J. Greve, G. J. Puppels and C. Otto, Kluwer Academic Publishers, Dordrecht, 1999, pp. 557–558.
18 L.-P. Choo-Smith, K. Maquelin, H. P. Endtz, H. A. Bruining and G. J. Puppels, in *Spectroscopy of Biological Molecules: New Directions, 8th European Conference*, ed. J. Greve, G. J. Puppels and C. Otto, Kluwer Academic Publishers, Dordrecht, 1999, pp. 537–540.
19 L. P. Choo-Smith, K. Maquelin, T. Van Vreeswijk, H. A. Bruining, G. J. Puppels, N. A. N. Thi, C. Kirschner, D. Naumann, D. Ami, A. M. Villa, F. Orsini, S. M. Doglia, H. Lamfarraj, G. D. Sockalingum, M. Manfait, P. Allouch and H. P. Endtz, *Appl. Environ. Microbiol.*, 2001, **67**, 1461–1469.
20 K. Maquelin, L.-P. Choo-Smith, T. van Vreeswijk, H. P. Endtz, B. Smith, R. Bennett, H. A. Bruining and G. J. Pupples, *Anal. Chem.*, 2000, **72**, 12–19.
21 P. Rösch, M. Schmitt, W. Kiefer and J. Popp, *J. Mol. Struct.*, 2003, **661–662**, 363–369.
22 P. Rösch, M. Harz, K.-D. Peschke, O. Ronneberger, H. Burkhardt, H.-W. Motzkus, M. Lankers, S. Hofer, H. Thiele and J. Popp, *Appl. Environ. Microbiol.*, 2005, **71**, 1626–1637.
23 P. Rösch, M. Harz, M. Schmitt and J. Popp, *J. Raman Spectrosc.*, 2005, **36**, 377–379.
24 K. C. Schuster, E. Urlaub and J. R. Gapes, *J. Microbiol. Methods*, 2000, **42**, 29–38.
25 K. C. Schuster, I. Reese, E. Urlaub, J. R. Gapes and B. Lendl, *Anal. Chem.*, 2000, **72**, 5529–5534.
26 W. E. Huang, R. I. Griffiths, I. P. Thompson, M. J. Bailey and A. S. Whiteley, *Anal. Chem.*, 2004, **76**, 4452–4458.
27 H. C. van der Mei, D. Naumann and H. J. Busscher, *Infrared Phys. Technol.*, 1996, **37**, 561–564.
28 D. Hutsebaut, K. Maquelin, P. De Vos, P. Vandenabeele, L. Moens and G. J. Puppels, *Anal. Chem.*, 2004, **76**, 6274–6281.
29 P. K. Sharma, A. Das, K. H. Rao, K. A. Natarajan and K. S. E. Forssberg, *Process Metall.*, 2001, **11B**, 589–598.
30 V. N. Vapnik, *The Nature of Statistical Learning Theory*, Springer Verlag, New York, 1995.
31 H. G. M. Edwards, N. C. Russell, R. Weinstein and D. D. Wynn-Williams, *J. Raman Spectrosc.*, 1995, **26**, 911–916.
32 A. C. Williams and H. G. M. Edwards, *J. Raman Spectrosc.*, 1994, **25**, 673–677.

 This journal is © The Royal Society of Chemistry 2005