Research Article

Attenuated Total Reflectance Fourier Transform Infrared Spectroscopy and Artificial Neural Networks Applied to Differentiate Escherichia coli $papG^+/papG^-$ Strains

Łukasz Lechowicz,1 Wioletta Adamus-Białek,2 and Wiesław Kaca1

1 Department of Microbiology, Jan Kochanowski University in Kielce, Swietokrzyska Street, 25–406 Kielce, Poland
2 Chair of Environmental Protection and Modeling, Jan Kochanowski University in Kielce, Swietokrzyska Street, 25–406 Kielce, Poland

Correspondence should be addressed to Łukasz Lechowicz; lechowiczlukasz@gmail.com

Received 1 August 2012; Accepted 29 October 2012

Academic Editor: Feride Severcan

Copyright © 2013 Łukasz Lechowicz et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Fimbriae are an important pathogenic factor of Escherichia coli during development of urinary tract infections. Here, we describe a new method for identification of Escherichia coli $papG^+$ from $papG^-$ strains using the attenuated total reflectance Fourier transform infrared Spectroscopy (ATR FT-IR). We applied artificial neural networks to the analysis of the ATR FT-IR results. These methods allowed to discriminate $E. coli$ $papG^+$ from $papG^-$ strains with accuracy of 99%.

1. Introduction

Some bacteria species have the ability to adhere to the surface of human cells. The adhesion plays an important role in pathogenesis, enabling bacteria colonization of the host and protecting them from the immune system. Bacterial adhesins are mostly proteins incorporated in a fimbriae [1]. A key phase in urinary tract infections involves colonization of the urethra and attaching bacterial cells to the urothelium [2]. Uropathogenic Escherichia coli strains form several types of fimbriae. The most important are class II P fimbriae, which play a significant role in the pyelonephritis and inflammation of the urinary bladder. These fimbriae increase the virulence of uropathogenic E. coli (UPEC) at different stages of pathogenesis. Strains possessing P fimbriae persist in the urinary tract for a long time; they spread more efficiently and induce inflammation. E. coli strains with P fimbriae induce a bacteriuria and can enter the bloodstream [3]. Studies in mice have shown that the presence of P fimbriae increases the levels of interleukin IL-6 and IL-8 in the urine, and also causes influx of polymorphonuclear leukocytes [4, 5]. P fimbriae are formed by six protein subunits: PapD, PapC, PapG, PapF, PapK, and PapH. PapG subunit allows bacteria binding to epithelial cells by attaching to glycolipid surface receptors containing two molecules of galactose.

The attenuated total reflectance Fourier transform infrared spectroscopy (ATR FT-IR) is a technique which is used to obtain an infrared spectrum of absorption scattering of various types of samples. This technique was used to detect and differentiate many bacterial species [6]. The major problem during the application of ATR FT-IR for microbiological diagnostics is a necessity to analyze large amounts of data. The usage of artificial neural networks (ANNs) eliminates this problem. During data analysis, artificial neural networks modify their structure, which allows them to minimize classification error.

The aim of this study was to use ATR FT-IR and artificial neural networks (ANNs) for differentiation of Escherichia coli $papG^+$ from $papG^-$ strains.

2. Materials and Methods

The materials used in the study were 63 uropathogenic Escherichia coli strains isolated from urine samples. The bacteria were stored in a collection deposited in the Chair of...
Environmental Protection and Modeling, Jan Kochanowski University in Kielce. The strains were previously characterized for the presence of the papG gene. Presence of the gene was confirmed in 26 strains (about 41% of all strains) [7]. The bacterial IR spectra were measured using a “Spectrum 400” spectrometer (Perkin Elmer). MS Excel and Statistica 10 applications were used for data processing and the design of artificial neural networks.

The bacteria were incubated on Luria-Bertani Agar at 37°C for 24 hours. After this time, the bacteria were stored at +4°C. A single typical colony was used to measure the bacterial IR spectrum. The colony was taken directly from Petri dishes and placed on the spectrometer crystal. The single colony was scanned 30 times, and, then, the results were averaged. For each bacterial strain, 10 to 15 samples were collected. The data obtained have been preprocessed according to the procedure proposed by [8]. Briefly, baseline of each spectrum was corrected automatically by spectrometer software. Next, first derivative of the spectrum was calculated using Savitzky-Golay algorithm (5 points smoothing). Finally, smoothed spectrum was normalized to the range from 0 to 1. Chi-square test and P value were used to choose predictors of the characteristics (the wavenumbers).

### 3. Results

The obtained bacterial IR spectra are presented in Figure 1. Several artificial neural networks were designed for the data analysis. The neural networks were built with 19 neurons in the input layer, a variable number of neurons in the hidden layer and 2 neurons in the output layer. The set of input data was randomly divided into three subsets: a training set (70% of spectra), a test set (15% of spectra), and a validating set (15% of spectra). The training subset was used for the network learning. The other two sets were used to validate the networks.

The designed networks differentiate bacterial strains with variable accuracy. Among the networks recognizing *E. coli* *papG*+ strains, the best results were reached by a network constructed from 5 neurons in the hidden layer (Table 1). In the validating set, this network correctly classified about 99% of the bacterial spectra. The results obtained by ATR FT-IR method are corresponding to the results obtained by PCR [7]. In all cases, activation function of hidden neurons was hyperbolic tangent. Using other types of activation functions (e.g., exponential or logistic function) yielded slightly worse results (data were not shown).

### 4. Discussion

Bacterial adhesion is one of the most important factor of UTI. The fimbiae formation is an essential features of bacterial pathogenesis. The P fimbiae are among that frequently recorded as surface factor of *E. coli* UTI strains. Our previous work divided collection of 63 *E. coli* strains isolated from UTI patients on *papG*+ and *papG*− genotypes. In the presented studies, ATR FT-IR spectra of bacterial colonies were analyzed by neural networks. The ANNs identified two

<table>
<thead>
<tr>
<th>ID</th>
<th>Topology</th>
<th>Quality (learning)</th>
<th>Quality (testing)</th>
<th>Quality (validation)</th>
<th>Learning algorithm</th>
<th>Error function</th>
<th>Activation function (hidden neurons)</th>
<th>Activation function (output neurons)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MLP 19-5-2</td>
<td>97.02%</td>
<td>93.51%</td>
<td>99.07%</td>
<td>BFGS 89</td>
<td>SOS</td>
<td>Tanh</td>
<td>Linear</td>
</tr>
<tr>
<td>2</td>
<td>MLP 19-6-2</td>
<td>86.50%</td>
<td>87.04%</td>
<td>82.40%</td>
<td>BFGS 40</td>
<td>SOS</td>
<td>Tanh</td>
<td>Linear</td>
</tr>
<tr>
<td>3</td>
<td>MLP 19-7-2</td>
<td>94.04%</td>
<td>90.74%</td>
<td>89.81%</td>
<td>BFGS 55</td>
<td>SOS</td>
<td>Tanh</td>
<td>Linear</td>
</tr>
<tr>
<td>4</td>
<td>MLP 19-8-2</td>
<td>96.42%</td>
<td>92.59%</td>
<td>88.89%</td>
<td>BFGS 103</td>
<td>SOS</td>
<td>Tanh</td>
<td>Linear</td>
</tr>
</tbody>
</table>

distinct E. coli groups with $\text{pap}G^+$ or $\text{pap}G^-$ genotypes with accuracy of 99%. That division was only possible with the help of designed artificial networks.

There is no assurance that the E. coli $\text{pap}G^+$ strains form the fimbriae in the applied culture conditions even it has all genes required for this process. It is not possible to perform direct interpretation of bacterial IR spectra. Especially, it is not possible to connect the presence of specific peak in IR spectrum with presence of specific protein in bacterial cell. The ATR FT-IR technique enables identifying the type and amount of functional groups in the sample. The IR bacterial spectrum allows getting to know a general molecular profile of the bacteria. For this reason, we assume that the spectra of E. coli $\text{pap}G^+$ strains are slightly different from the spectra of E. coli $\text{pap}G^-$ strains. The differences depend not only on the presence of the fimbriae, but also on other chemical compounds (i.e., proteins, fatty acids, polysaccharides, and nucleic acids). The chemical components responsible for varied spectra remain to be elucidated.

In conclusion, we showed that the ATR FT-IR and ANNs can be used for differentiation of E. coli strains.

Acknowledgments

The authors thank Dr. Marek Kwinkowski (Department of Microbiology, UJK) and Dr. Mariusz Urbaniak (Institute of Chemistry, UJK) for technical support. This work was supported by the ESF Human Capital Operational Program Grant 6/1/8.2.1/POKL/2009 and Grant 2011/01/D/NZ7/00107 supported by National Science Centre, Poland.

References
