POLISH JOURNAL OF FOOD AND NUTRITION SCIENCES

www.pan.olsztyn.pl/journal/
e-mail: joan@pan.olsztyn.pl

Pol. J. Food Nutr. Sci.
2007, Vol. 57, No. 3, pp. 301–306

IDENTIFICATION OF *LACTOBACILLUS* STRAINS AT THE SPECIES LEVEL USING FTIR SPECTROSCOPY AND ARTIFICIAL NEURAL NETWORKS

Bartłomiej Dziuba

Chair of Industrial and Food Microbiology, University of Warmia and Mazury in Olsztyn, Olsztyn

Key words: *Lactobacillus* strains, FTIR spectra, artificial neural networks

Fourier-transform infrared (FTIR) spectroscopy and artificial neural networks were used to identify bacteria of the genus *Lactobacillus* at the species level. A previously developed method for measuring FTIR spectra, and a strategy for their analysis provided the basis for selecting the FTIR spectra of the tested bacteria, and for creating a spectral library, as described elsewhere [Dziuba *et al*., 2007b]. In our previous study [Dziuba *et al*., 2007b] we demonstrated that the FTIR spectral characteristics of *Lactobacillus* strains based exclusively on the differentiation index D, calculated from the Pearson's correlation coefficient, and cluster analysis are not sufficient to describe the relationships between FTIR spectra and bacteria as molecular systems in a way that would permit their proper identification. Thus, research was launched in which the spectra collected in the above library were used for developing artificial neural networks. The practical value of these networks was verified based on the results of identification of 17 bacterial strains of known taxonomy as well as 7 strains isolated from dairy products and identified on the basis of their taxonomy and biochemical tests.

The application of artificial neural networks, *i.e.* the most advanced chemometric method, to analysis of FTIR spectra enabled correct identification of 93% of bacterial strains of the genus *Lactobacillus*.

INTRODUCTION

The latest methods for microorganism differentiation and identification are oriented towards simplicity, reducing the time needed to perform analysis, as well as towards high specificity and quality of results. One of the methods that meet these criteria is Fourier-transform infrared (FTIR) spectroscopy. Our previous studies [Dziuba *et al*., 2006, 2007a, b] show that FTIR spectra can be treated as specific to particular bacterial strains. They reflect the distinguishing features of all bacterial cell components, such as fatty acids, membrane and intracellular proteins, polysaccharides and nucleic acids [Naumann, 2000]. The differences between FTIR spectra of microorganisms are difficult to observe, so their analysis requires the use of statistical methods, which can be divided into supervised and unsupervised. Cluster analysis is the most popular among unsupervised methods. With the use of appropriate algorithms, cluster analysis allows to group bacteria based on the similarities between their FTIR spectra (Pearson's correlation coefficient). However, results obtained in this way are often subjective and dependent on the interpretation of complex dendrograms or dispersion diagrams. In addition, they are not always satisfactory [Johnsen & Nielsen, 1999; Timmins *et al*., 1998, Helm *et al*., 1991; Kummerle *et al*., 1998]. Among the most often applied supervised methods there are discriminant analysis, discriminant function analysis and canonical correlation analysis [Amiel *et al*., 2000; Beattie *et al*., 1998; Lefier *et al*., 1997]. The results of the identification of bacterial spectra obtained by the above methods are comparable to those obtained with the use of unsupervised methods.

Apart from classical methods, artificial neural networks are often applied for microorganism identification [Goodacare *et al*., 1998; Kirchner *et al*., 1999; Tintelnot *et al*., 2000; Udelhoven *et al*., 2000]. Among numerous reports on microorganism differentiation, classification and identification using FTIR spectroscopy, a few only deal with lactic acid bacteria [Amiel *et al*., 2001; Lefier *et al*., 2000; Weinrichter *et al*., 2001]. However, their authors have not applied artificial neural networks in their investigations.

In our previous report [Dziuba *et al*., 2007b] we demonstrated that the FTIR spectral characteristics of lactic acid bacteria based exclusively on the differentiation index D and cluster analysis are not sufficient to describe the relationships between FTIR spectra and bacteria as molecular systems (*i.e.* systems of a certain nature, composition, and complexity of chemical components) in a way that would permit their proper identification. FTIR spectra of bacteria are composed of hundreds or even thousands of overlapping bands that cannot be separated. It follows that their analysis requires the application of pattern recognition techniques, analyzing spectra as fingerprints. One of the most advanced and promising methods of this group are artificial neural networks, which were used in the present study.

Author's address for correspondence: Bartłomiej Dziuba, Chair of Industrial and Food Microbiology, University of Warmia and Mazury in Olsztyn, Cieszynski Square 1, 10-957 Olsztyn, Poland; e-mail: Dziuba@uwm.edu.pl

MATERIALS AND METHODS

Strains and growth conditions. Strains used in the study come from international collections, *i.e.* ATCC, LMG, NCFB as well as from the microbial strain collection of the Department of Industrial and Food Microbiology, University of Warmia and Mazury (UWM) in Olsztyn and the microbial strain collection created within the Inco-Copernicus IC15-CT98-0905 Project (ICCT), and seven strains isolated from environment. Forty five strains of *Lactobacillus* were used in the study.

Strains were cultured in solid media for 48±2 h under optimum conditions, *i.e.* mesophilic and thermophilic lactobacilli were cultured in the MRS medium at 30°C and 37°C, respectively.

Sample preparation. Small amounts of bacterial cells were taken with a loop from the agar surface and suspended in 200 μL of physiological saline. The suspension (120 μL) was evenly spread over the surface of a ZnSe window and dried for 30 min at 50°C until a transparent film was obtained. Three samples of two independent cultures (a total of six samples) were prepared for each strain. The experiment was conducted in three replications.

Spectroscopic measurements. All bacterial FTIR spectra were recorded between 4000 cm^{-1} and 500 cm^{-1} in transmission/absorbance mode on a Spectrum One spectrometer (Perkin-Elmer) equipped with a KBr beam-splitter and a DTGS (Deuterated TriGlycerine Sulfate) detector by co-addition and averaging of 64 scans. Spectral resolution was 4 cm^{-1}, encoding interval was 1 cm^{-1} and scanning speed was 0.2 cm/s. The background spectrum measurement was performed with a ZnSe window free of any sample.

In order to standardize data and reduce the noise, the following operations were carried out: baseline correction (5-point) at a wavelength of 4000 cm^{-1}, 3700 cm^{-1}, 2700 cm^{-1}, 1800 cm^{-1} and 750 cm^{-1} [Weinrichter *et al.,* 2001]; normalization to the value 1.0 in relation to the most intensive band (amides I, approx. 1640 cm^{-1}). To increase the resolution of the obtained spectra, their first derivatives (Savitzky-Golay algorithm, 9-point smoothing) and second derivatives (Savitzky-Golay algorithm, 9-point smoothing) were calculated.

All operations on spectra were performed using original Spectrum One 3.01.00 software (Perkin Elmer), and saved in ASCII format, compatible with Excel (Microsoft) and Statistica (Statsoft Poland).

Artificial neural networks. Two types of neural networks were used in the study: probabilistic neural networks (PNNs) and multilayer perceptrons (MLPs) [Bishop, 1995; Goodacre *et al.,* 1998; Speckt, 1990]. The first layer of the network was an input layer, to which the absorbance values at a given wavelength were introduced. The number of neurons in the input layer was equal to the number of data points of the spectrum or selected wave numbers. The spectral data were divided into three sets (learning, validation and testing) to include the spectra of all strains analysed. The training set was composed of 160 FTIR spectra, *i.e.* 16 spectra of each bacterial species. The validation set consisted of 40 FTIR spectra of 21 reference strains, *i.e.* 4 spectra of each bacterial species. The testing set included 21 spectra of reference strains, 17 spectra of bacteria of known taxonomy and 7 spectra of bacteria isolated from the environment. The hidden layer consisted of neurons whose number was equal to the number of learning examples in the case of PNNs and defined experimentally in that of MLPs.

In the output layer, the number of neurons was equal to the number of considered classes. For the identification of *Lactobacillus* strains at the species level, the output layer consisted of ten neurons with assigned classes. The bacterial spectra were binary encoded: *L. amylovorus*-(1,0,0,0,0,0,0,0,0,0); *L. acidophilus*-(0,1,0,0,0,0,0,0,0,0); *L. brevis*-(0,0,1,0,0,0,0,0,0,0); *L. rhamnosus*-(0,0,0,1,0,0,0,0,0,0); *L. casei*-(0,0,0,0,1,0,0,0,0,0); *L. delbrueckii*.-(0,0,0,0,0,1,0,0,0,0); *L.fermentum*-(0,0,0,0,0,0,1,0,0,0); *L. helveticus*-(0,0,0,0,0,0,0,1,0,0); *L. hilgardii*-(0,0,0,0,0,0,0,0,1,0); *L. plantarum*-(0,0,0,0,0,0,0,0,0,1}. Neurons organized in the layers, were connected in a way specific to a given type of artificial neural network. Each connection in the network was assigned weight, modified by the network during training. In addition, "thresholds" were set and used to compute the activation level of each neuron. In the output layer, if the threshold value on given neuron was exceeded, the case under consideration was classified to given species represented by this neuron.

The multilayer perceptrons used in the study consisted of three to four layers. Back-propagation and conjugate gradient descent algorithms [Bishop, 1995] were used in the learning process, which was conducted until the root mean square (RMS) error between the actual and desired outputs was lower than 0.01.

Probabilistic neural networks are a type of the so-called "Bayesian" networks [Bishop 1995; Speckt, 1990] that use kernel-based approximation to form an estimate of the probability density functions of classes in a classification problem. The PNN used in this study consisted of three layers (input, radial and linear) of classification neurons. The only control factor that needs to be selected for probabilistic neural network training is the smoothing factor (*i.e.* the radial deviation of the Gaussian function), which was set between 0.3 and 1.0. All statistical operations were done using Statistica ANN (Statsoft Poland).

RESULTS AND DISCUSSION

The spectral characteristics and results of differentiation (cluster analysis) of bacteria of the genus *Lactobacillus* are presented in our previous paper [Dziuba *et al*., 2007b]. However, cluster analysis did not allow to correctly differentiate between bacteria of the genus *Lactobacillus* at the species level. Thus, artificial neural networks were used for the identification of lactic acid bacteria in the present study. One of the advantages of ANNs is their capability to generalize knowledge to novel objects and previously unseen cases, not fed into the network in the training process. At the same time ANNs are able to memorize the acquired knowledge, which can be used at any moment without the need to feed the information again. Two types of neural networks were used in the study: probabilistic neural networks (PNNs) and multilayer perceptrons (MLPs).

Before the training process began, the digitized spectra put in calculation sheets with assigned classes had been divided into three sets: training, validation and testing. Appropriate

TABLE 1. Set of the best artificial neural networks used for identification of *Lactobacillus* species.

Spectrum range	Type	Structure			RMSEF			Quality				No. of epochs
		inlet	hidden	hidden 2	learning	validation	test*	learning	validation	test 1	test 2	
4000 cm^{-1}–500 cm^{-1}	MLP	236	37	-	0.006	0.01	0.040	1	1	1	0.91	5.8×10^3
W5W4W3W2	PNN	1041	160	-	0.009	0.015	0.055	1	1	1	0.91	-
W5W4W3	MLP	651	40	40	0.008	0.014	0.033	1	1	1	0.93	2×10^4
W5W4W2	MLP	36	39	-	0.012	0.019	0.045	1	1	1	0.89	7×10^3
W5W4W2	PNN	741	160	-	0.008	0.008	0.041	1	1	1	0.91	-
AL.GEN.	MLP	368	33	-	0.007	0.013	0.038	1	1	1	0.91	4.1×10^3

options of the software were used to this end. Only the spectra from the training set were used in the learning process. The spectra from the validation served to control the learning process and to evaluate the validity of network structure. The neural network performance, *i.e.* its capability to generalize the knowledge acquired during training, was tested on the basis of the spectra contained in the testing set (the cases from the testing set were not used in the learning process).

A total of 245 spectra were used during the development of artificial neural networks for the identification of bacteria of the genus *Lactobacillus* at the species level. The training set was composed of 160 FTIR spectra, while the validation set consisted of 40 FTIR spectra of 21 reference strains. The testing set included 21 spectra of reference strains, 17 spectra of bacteria of known taxonomy and 7 spectra of bacteria isolated from the environment. The number of spectra of each bacterial species used for training or validation had to be the same to maintain a balance in their representation. This is of primary importance with probabilistic neural networks determining the probability of occurrence of a given case. In order to reduce the number of input variables, the process of network training was followed by sensitivity analysis. This analysis allowed to estimate the usefulness of particular input variables, to determine which of them could be ignored and which contributed to problem solving to the greatest extent (651 variables for ANN no. 3 in Table 1). The sensitivity analysis was performed using appropriate option of the software. The selection of the variables was performed experimentally. The selected variables provided the basis for creating another neural network. The trained network enabled to obtain fully correct results.

The correctness of identification of *Lactobacillus* strains at the species level was verified according to a two-stage procedure. At the first stage artificial neural networks were tested based on the spectra of 21 reference strains. At the second stage the spectra of 17 strains of known taxonomy and 7 isolated strains, identified with biochemical tests (APICHL 50), were used. All neural networks were generated on the basis of the first derivatives of the FTIR spectra of the tested bacteria. Similarly as in the case of bacteria identification at the genus level [Dziuba *et al*., 2007a], the optimum artificial neural networks (Table 1) were selected as a consequence of searching for the most relevant input variables, network structure and training parameters. The ANNs presented in Table 1 correctly identified all reference strains of testing set I, whereas bacteria of testing set II were correctly identified in 89% to 93% of cases. The best results were achieved using MLP (network 3) for a combination of the spectral ranges: W5xW4xW3. The network correctly identified all reference strains (Table 2) and all strains of known taxonomy (Table 2, strains denoted with *). Among 7 isolated strains (Table 2, strains denoted with **), the network explicitly classified four of them. The affiliation of the following three strains was not determined unambiguously: *Lb. casei* IZ, *Lb. rhamnosus* IZ 01 and *Lb. rhamnosus* IZ 02. In the case of the first two strains the network could not decide between *Lb. casei* and *Lb. fermentum*, and between *Lb. rhamnosus* and *Lb. delbrueckii*, respectively. Taking into account the winning neuron only, their classification would be correct. The third strain, *Lb. rhamnosus* IZ 02, was wrongly identified.

The other networks (Table 3) were not able to correctly identify the following strains: *Lb. casei* IZ, *Lb. rhamnosus* IZ 01 and *Lb. rhamnosus* IZ 02. In addition, these networks (no. 1, 2, 4, 5, 6) made errors. For instance, MLP developed on the basis of a combination of the ranges W5xW4xW2, apart from the above strains, could not correctly identify the following ones: *Lb. acidophilus* 5e2, *Lb. rhamnosus* S11 and *Lb. delbrueckii* LY03.

It was found that differentiation of *Lactobacillus* strains at the species level based on analysis of individual spectral regions and the entire spectrum was limited. The present results confirmed the correctness of conclusions drawn from a comparison of bacterial spectra determined on the basis of the differentiation index D, but also proved the existence of such combinations of spectral ranges which permitted the most reliable identification of bacterial strains. The best results of differentiation of bacteria of the genus *Lactobacillus* at the species level were obtained for combinations of the ranges W4xW5xW3 of the first derivatives of spectra, *i.e.* for the polysaccharide region, the fingerprint region and the mixed region. Curk *et al*. [1994] and Weinrichter *et al*. [2001] also achieved the best results for the above combinations. However, these authors did not use artificial neural networks, and we demonstrated previously [Dziuba *et al*., 2007b] that the FTIR spectral characteristics of bacteria based exclusively on the differentiation index D and cluster analysis are not sufficient to describe the relationships between FTIR spectra and bacteria as molecular systems in a way that would permit their proper identification. That is why we used artificial neural networks, *i.e.* an advanced method for pattern recognition enabling to analyze spectra as fingerprints. The application of artificial neural networks enabled correct identification of 93% of bacterial strains of the genus *Lactobacillus*. This result could be probably improved if the number of strains was increased, preferably including the so-called typical strains. A promising solution is to develop multilevel artificial neural networks, forming a single structure. The networks organized in this way would enable identifying microorganisms at various taxonomic levels.

Table 2. Identity of the *Lactobacillus* species in the testing set as judged by MLP (651:40:40:10) analysis of their FTIR spectra.

Strain	Estimates from MLP									
	L .acidophilus	*L. amylovorus*	*L. brevis*	*L casei*	*L. delbrueckii*	*L. fermentum*	*L. helveticus*	*L. hilgardii*	*L. plantarum*	*L. rhamnosus*
Lactobacillus acidophilus LMG 9433	1.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.00
Lactobacillus acidophilus ICCT 4356	0.98	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lactobacillus acidophilus ICCT 5e2*	0.83	0.00	0.00	0.02	0.00	0.02	0.35	0.00	0.00	0.00
Lactobacillus acidophilus IZ**	0.94	0.00	0.00	0.00	0.03	0.00	0.02	0.00	0.00	0.00
Lactobacillus amylovorus LMG 9496	0.00	0.94	0.00	0.00	0.00	0.00	0.01	0.00	0.12	0.00
Lactobacillus brevis LMG 6906	0.00	0.00	0.98	0.00	0.01	0.00	0.00	0.00	0.00	0.00
Lactobacillus brevis ICCT 110	0.00	0.00	0.97	0.00	0.03	0.00	0.00	0.00	0.00	0.00
Lactobacillus brevis UWM 1*	0.00	0.00	0.90	0.00	0.00	0.00	0.00	0.00	0.14	0.00
Lactobacillus brevis UWM 2*	0.00	0.00	0.89	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Lactobacillus casei LMG 6904	0.02	0.00	0.00	0.97	0.00	0.01	0.00	0.00	0.00	0.00
Lactobacillus casei ICCT 370	0.05	0.00	0.00	0.93	0.00	0.01	0.00	0.00	0.00	0.00
Lactobacillus casei ICCT 541*	0.00	0.00	0.00	0.99	0.05	0.00	0.00	0.00	0.00	0.00
Lactobacillus casei IZ**	0.00	0.00	0.00	0.90	0.11	0.00	0.00	0.00	0.00	0.00
Lactobacillus casei IZ**	0.00	0.00	0.00	0.48	0.06	0.42	0.00	0.00	0.00	0.01
Lactobacillus delbrueckii ATCC 6901	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Lactobacillus delbrueckii LMG 7942	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.01	0.00
Lactobacillus delbrueckii ATCC 7830	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
Lactobacillus delbrueckii ATCC 9649	0.00	0.00	0.00	0.00	0.88	0.00	0.00	0.00	0.00	0.00
Lactobacillus delbrueckii ICCT T162*	0.00	0.00	0.00	0.00	0.97	0.00	0.00	0.00	0.04	0.00
Lactobacillus delbrueckii ICCT LY58*	0.00	0.00	0.00	0.00	0.98	0.00	0.00	0.00	0.03	0.00
Lactobacillus delbrueckii ICCT LY03*	0.00	0.00	0.00	0.00	0.89	0.00	0.00	0.00	0.00	0.00
Lactobacillus delbrueckii IZ**	0.00	0.00	0.07	0.00	0.45	0.00	0.00	0.00	0.00	0.00
Lactobacillus fermentum LMG 6902	0.00	0.00	0.02	0.00	0.00	0.98	0.00	0.00	0.00	0.00
Lactobacillus fermentum ICCT 02	0.00	0.00	0.01	0.06	0.00	0.89	0.00	0.00	0.00	0.00
Lactobacillus fermentum ICCT 6991*	0.32	0.00	0.00	0.17	0.00	0.61	0.00	0.00	0.00	0.00
Lactobacillus helveticus LMG 6413	0.05	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
Lactobacillus helveticus rosyjski	0.01	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00	0.01
Lactobacillus helveticus ICCT 122*	0.03	0.00	0.00	0.00	0.00	0.00	0.96	0.00	0.01	0.00
Lactobacillus helveticus ICCTCH-1-18*	0.05	0.00	0.00	0.00	0.00	0.00	0.93	0.00	0.00	0.00
Lactobacillus hilgardii LMG 6895	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.99	0.00	0.00
Lactobacillus hilgardii ICCT 113	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.99	0.00	0.00
Lactobacillus hilgardii ICCT 1031*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.93	0.00	0.00
Lactobacillus plantarum ATCC 8014	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.99	0.00
Lactobacillus plantarum LMG 6907	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
Lactobacillus plantarum UWM 3*	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.98	0.00
Lactobacillus plantarum UWM 4*	0.10	0.29	0.00	0.00	0.00	0.00	0.00	0.00	0.96	0.00
Lactobacillus plantarum UWM 98*	0.00	0.00	0.00	0.00	0.00	0.12	0.00	0.01	0.97	0.00
Lactobacillus plantarum UWM 65*	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.88	0.00
Lactobacillus plantarum IZ**	0.00	0.03	0.00	0.00	0.00	0.00	0.00	0.01	0.80	0.00
Lactobacillus rhamnosus ATCC 7469	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.99
Lactobacillus rhamnosus ICCT 574	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00
Lactobacillus rhamnosus ICCT S11*	0.00	0.00	0.00	0.00	0.10	0.00	0.00	0.00	0.00	0.83
Lactobacillus rhamnosus ICCT L*	0.00	0.00	0.00	0.00	0.04	0.00	0.00	0.00	0.00	0.87
Lactobacillus rhamnosus IZ 01**	0.00	0.00	0.00	0.00	0.45	0.00	0.01	0.00	0.00	0.53
Lactobacillus rhamnosus IZ 02**	0.00	0.00	0.11	0.00	0.52	0.00	0.00	0.00	0.12	0.43

* – strains with known taxonomy, ** – strains isolated from the medium

TABLE 3. Examples of misjudgement of the genus of lactic acid bacteria.

Range	Type	Strain	Estimates in output layer									
			L .acidophilus	*L. amylovorus*	*L. brevis*	*L casei*	*L. delbrueckii*	*L. fermentum*	*L. helveticus*	*L. hilgardii*	*L. plantarum*	*L. rhamnosus*
4000 cm^{-1}–500 cm^{-1}	MLP	*L. rhamnosus* S11	0.00	0.00	0.00	0.00	0.78	0.00	0.00	0.00	0.00	0.89
W5W4W3W2	PNN	*L. plantarum* UWM65	0.00	0.00	0.52	0.00	0.00	0.00	0.00	0.00	0.48	0.00
W5W4W2	MLP	*L. acidophilus* 5e2	0.31	0.00	0.00	0.00	0.98	0.00	0.00	0.00	0.02	0.01
		L. rhamnosus S11	0.00	0.00	0.00	0.00	0.81	0.00	0.00	0.00	0.00	0.99
		L. delbrueckii LY03	0.00	0.00	0.00	0.00	0.64	0.00	0.00	0.00	0.63	0.00
W5W4W2	PNN	*L. plantarum* UWM65	0.00	0.00	0.65	0.00	0.00	0.00	0.00	0.00	0.35	0.00
AL.GEN.	MLP	*L .acidophilus* 5e2	0.88	0.00	0.01	0.00	0.00	0.00	0.57	0.02	0.00	0.00

CONCLUSIONS

1. Combinations of the polysaccharide region (1200 cm^{-1} – 900 cm^{-1}) with the fingerprint region (900 cm^{-1} – 700 cm^{-1}) and the mixed region (1500 cm^{-1} – 1200 cm^{-1}) of spectra as well as their first derivatives were most useful for characterising the FTIR spectra of members of the genus *Lactobacillus*.

2. The application of artificial neural networks, *i.e.* the most advanced chemometric method, to analysis of FTIR spectra, enabled correct identification of bacteria of the genus *Lactobacillus* strains, belonging to particular species, in 93% of cases.

3. FTIR spectroscopy, combined with artificial neural networks, was found to be a rapid and accurate method for bacteria identification.

REFERENCES

1. Amiel C., Mariey L., Denis C., Pichon P., Travert J., FTIR spectroscopy and taxonomic purpose: Contribution to the classification of lactic acid bacteria. Le Lait, 2001, 81, 249–255.
2. Amiel C., Mariey L., Curk-Daubie M.C., Travert J., Potentiality of Fourier Transform Infrared Spectroscopy (FTIR) for discrimination and identification of dairy Lactic acid bacteria. Le Lait, 2000, 80, 445–459.
3. Beattie S.H., Holt C., Hirst D., Williams A.G., Discrimination among *Bacillus cereus*, *B. mycoides* and *B. thuringiensis* and some other species of the genus *Bacillus* by Fourier transform infrared spectroscopy. FEMS Microbiol. Lett., 1998, 164, 201–206.
4. Bishop C.M., Neural Networks for Pattern Recognition, 1995, Oxford Univ. Press, New York, pp. 364–368.
5. Curk M.C., Peladan F., Hubert J.C., Fourier transform infrared spectroscopy for identifying *Lactobacillus* species. FEMS Microbiol. Lett., 1994, 123, 241–248.
6. Dziuba B., Babuchowski A., Niklewicz M., Brzozowski B., FTIR spectral characteristics of lactic acid bacteria – a spectral library. Milchwissenschaft, 2006, 61, 146–149.
7. Dziuba B., Babuchowski A., Dziuba M., Nałęcz D., Identification of lactic acid bacteria using FTIR spectroscopy and artificial neural networks. Milchwissenschaft, 2007a, 62, 28–31.
8. Dziuba B., Babuchowski A., Dziuba M., Nałęcz D., Identification of lactic acid bacteria using FTIR spectroscopy and cluster analysis. Int. Dairy J., 2007b, 17, 183–189.
9. Goodacre R., Timmins E.M., Burton R., Kaderbhal N., Woodward A., Kell D.B., Rooney P.J., Rapid identification of urinary tract infection bacteria using hyperspectral whole-organism fingerprinting and artificial neural networks. Microbiology, 1998, 144, 1157–1170.
10. Helm D., Labischinski H., Nauman D., Elaboration of a procedure for identification of bacteria using Fourier-Transform IR spectral libraries: a stepwise correlation approach. J. Microbiol. Meth., 1991, 14, 127–142.
11. Johnsen K., Nielsen P., Diversity of Pseudomonas strains isolated with King's B and Gould's S1 agar determined by repetitive extragenic palindromic-polymerase chain reaction, 16S rDNA sequencing and Fourier transform infrared spectroscopy characterisation. FEMS Microbiol. Lett., 1999, 173, 155.
12. Kirchner C., Ngo Thi N.A., Naumann D., FT-IR microscopy: A rapid method for classifying microorganisms. 1999, *in*: Spectroscopy of Biological Molecules: New Directions (eds. J. Greve, G.J. Puppels, C. Otto). Kluwer Academic Publishers, Dordrecht, pp. 557–558.
13. Kummerle M., Scherer S., Seiler H., Rapid and reliable identification of fermentative yeasts by Fourier-transform infrared spectroscopy. Appl. Environ. Microbiol., 1998, 64, 2207–2214.
14. Lefier D., Hirst D., Holt C., Williams A.G., Effect of sampling procedure and strain variation in *Listeria monocytogenes* on the discrimination of species in the genus *Listeria* by Fourier transform infrared spectroscopy and canonical variates analysis. FEMS Microbiol. Lett., 1997, 147, 45–50.
15. Lefier D., Lamprell H., Mazerolles G., Evolution of *Lactococcus* strains during ripening in Brie cheese using Fourier transform infrared spectroscopy. Le Lait, 2000, 80, 247–254.
16. Nauman D., Infrared spectroscopy in microbiology. 2000, *in*: Encyclopedia of Analytical Chemistry (ed. R.A. Meyers). John Wiley & Sons, Ltd., Chichester, pp. 102–131.
17. Speckt D.F., Probabilistic neural networks. Neural Networks, 1990, 3, 109–118.
18. Timmins E.A., Quain D.E., Goodacre R., Differentiation of brewing yeast strains by pyrolysis mass spectrometry and Fourier transform infrared spectroscopy. Yeast, 1998, 14, 885–893.
19. Tintelnot K., Haase G., Seibold M., Bergmann F., Staemmler M., Franz T., Nauman D., Evaluation of phenotypic markers for selection and identification of *Candida dubliniens*. J. Clin. Microbiol., 2000, 46, 1599–1608.
20. Udelhoven T., Nauman D., Schmitt J., Development of a hierarchical classification systems with artificial neural networks and FTIR spectra fort he identification of bacteria. Appl. Spectroscopy, 2000, 54, 1471–1479.
21. Weinrichter B., Luginbuhl W., Rohm H., Jimeno J., Differentiation of facultatively heterofermentative Lactobacilli from plants, milk and hard type cheeses by SDS-PAGE, RAPD, FTIR, energy source utilisation and autolysis type. Lebens. Wiss. Technol., 2001, 34, 556–566.

Received November 2006. Revision received and accepted February 2007.

IDENTYFIKACJA BAKTERII Z RODZAJU *LACTOBACILLUS* NA POZIOMIE GATUNKU ZA POMOCĄ SPEKTROSKOPII FTIR I SZTUCZNYCH SIECI NEURONOWYCH

Bartłomiej Dziuba

Katedra Mikrobiologii Przemysłowej i Żywności, Uniwersytet Warmińsko-Mazurski w Olsztynie, Olsztyn

W pracy podjęto badania nad wykorzystaniem spektroskopii w podczerwieni z transformacją Fouriera (FTIR) oraz sztucznych sieci neuronowych (artificial neural networks) do identyfikacji bakterii z rodzaju *Lactobacillus* na poziomie gatunku. Metodyka pomiaru widm FTIR oraz strategia ich analizy wykorzystane zostały do selekcji widm FTIR badanych bakterii i utworzenia biblioteki widm, tak jak opisano w naszej poprzedniej pracy [Dziuba *et al*., 2007b]. W badaniach tych [Dziuba *et al.*, 2007b] stwierdziliśmy, że charakterystyka widm FTIR bakterii z rodzaju *Lactobacillus*, jedynie w oparciu o współczynnik różnicowania D, obliczony na podstawie korelacji Pearsona oraz analiza skupień, nie są wystarczające do takiego opisania zależności pomiędzy widmami FTIR a bakteriami jako systemami molekularnymi, by możliwa była ich poprawna identyfikacja. Podjęto więc badania, w których zgromadzone w bibliotece widma zostały wykorzystane do opracowania sztucznych sieci neuronowych. Użytkowa wartość sieci neuronowych została ustalona na podstawie wyników identyfikacji 21 szczepów bakterii o znanej taksonomii, oraz 7 szczepów wyizolowanych z produktów mlecznych i zidentyfikowanych na podstawie ich taksonomii i testów biochemicznych.

Zastosowanie sztucznych sieci neuronowych do analizy widm FTIR, jako najbardziej zaawansowanej metody chemometrycznej, pozwoliło poprawnie zidentyfikować w 93% szczepy bakterii z rodzaju *Lactobacillus.* W przypadku trzech szczepów sieć nie potrafiła jednoznacznie określić ich nazwy taksonomicznej. Wynik ten można poprawić poprzez zwiększenie liczby szczepów reprezentujących poszczególne gatunki, najlepiej o tak zwane szczepy typowe.

Przyszłościowym rozwiązaniem powyższego problemu jest zbudowanie wielopoziomowych sztucznych sieci neuronowych, tworzących jedną strukturę. Tak zorganizowane sztuczne sieci neuronowe identyfikowałyby mikroorganizmy na różnych poziomach taksonomicznych.