

# Use of Artificial Neural Networks and a Gamma-Concept-Based Approach To Model Growth of and Bacteriocin Production by *Streptococcus macedonicus* ACA-DC 198 under Simulated Conditions of Kasserì Cheese Production<sup>∇</sup>

Panayiota Poirazi,<sup>3,†</sup> Frédéric Leroy,<sup>2,†</sup> Marina D. Georgalaki,<sup>1</sup> Anastassios Aktypis,<sup>1</sup>  
Luc De Vuyst,<sup>2</sup> and Effie Tsakalidou<sup>1\*</sup>

Laboratory of Dairy Research, Department of Food Science and Technology, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece<sup>1</sup>; Research Group of Industrial Microbiology and Food Biotechnology, Department of Applied Biological Sciences and Engineering, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium<sup>2</sup>; and Computational Biology Laboratory, Institute of Molecular Biology and Biotechnology, Foundation of Research and Technology-Hellas, Vassilika Vouton, P.O. Box 1385, 711 10 Heraklion, Greece<sup>3</sup>

Received 21 July 2006/Accepted 20 November 2006

Growth of and bacteriocin production by *Streptococcus macedonicus* ACA-DC 198 were assessed and modeled under conditions simulating Kasserì cheese production. Controlled fermentations were performed in milk supplemented with yeast extract at different combinations of temperature (25, 40, and 55°C), constant pH (pHs 5 and 6), and added NaCl (at concentrations of 0, 2, and 4%, wt/vol). The data obtained were used to construct two types of predictive models, namely, a modeling approach based on the gamma concept, as well as a model based on artificial neural networks (ANNs). The latter computational methods were used on 36 control fermentations to quantify the complex relationships between the conditions applied (temperature, pH, and NaCl) and population behavior and to calculate the associated biokinetic parameters, i.e., maximum specific growth and cell count decrease rates and specific bacteriocin production. The functions obtained were able to estimate these biokinetic parameters for four validation fermentation experiments and obtained good agreement between modeled and experimental values. Overall, these experiments show that both methods can be successfully used to unravel complex kinetic patterns within biological data of this kind and to predict population kinetics. Whereas ANNs yield a better correlation between experimental and predicted results, the gamma-concept-based model is more suitable for biological interpretation. Also, while the gamma-concept-based model has not been designed for modeling of other biokinetic parameters than the specific growth rate, ANNs are able to deal with any parameter of relevance, including specific bacteriocin production.

Lactic acid bacteria (LAB) are extensively used in the production of fermented foods for their contribution to shelf life, texture, and sensory properties (3). LAB are food grade and safe microorganisms that produce a variety of antimicrobial agents, including bacteriocins (5). Bacteriocins are peptides or proteins with antibacterial activity against bacteria closely related to the producer strain, including several spoilage bacteria and food-borne pathogens, such as *Listeria monocytogenes*, *Staphylococcus aureus*, and *Clostridium botulinum*. They are therefore of great interest to the food industry as natural food preservatives (22, 23, 28).

Among spoilage bacteria, *Clostridium (tyro)butyricum* is considered to be responsible for the so-called “late blowing” that is a serious defect especially in hard and semihard cheeses such as Emmental, Gruyère, Grana, Edam, Gouda, and Kasserì,

where *Clostridium* growth is favored by the rather high pH of these cheeses (27, 33). Clostridial spoilage is currently prevented by the use of the chemical nitrate, which is reduced to the antibacterial nitrite during cheese ripening. Bacteriocins are an interesting alternative to nitrate (16). Nisin, a lantibiotic originally added to processed cheese to prevent clostridial spoilage, is currently the only bacteriocin used as an effective biopreservative in dairy and nondairy foods (4).

*Streptococcus macedonicus* ACA-DC 198, a strain isolated from naturally fermented Greek Kasserì cheese (31), produces a lantibiotic, named macedocin, that shows inhibitory activity against *C. tyrobutyricum* (9, 10). Macedocin is active at pH values between 4.0 and 9.0, withstands sterilization, and is sensitive to rennet only after 24 h of incubation (10). Remarkably, macedocin is only produced when *S. macedonicus* ACA-DC 198 is grown in milk (10, 32). The producer strain itself is a moderate acidifier and displays low proteolytic and lipolytic activities, which may contribute to the development of cheese flavor (9).

The main factors influencing the growth, stability, and metabolism of microorganisms in foods are nutrient availability, temperature, pH, and water activity (2, 20). As these factors

\* Corresponding author. Mailing address: Laboratory of Dairy Research, Department of Food Science and Technology, Agricultural University of Athens, Iera Odos 75, 118 55 Athens, Greece. Phone: 30 210 529 4661. Fax: 30 210 529 4672. E-mail: et@aua.gr.

† P.P. and F.L. contributed equally to this work.

∇ Published ahead of print on 8 December 2006.

may vary extensively during food production and storage, mathematical models are needed to predict microbial behavior in the food matrix. So far, in food microbiology, mathematical models have been mainly developed to simulate the responses of undesirable microorganisms (i.e., pathogenic and spoilage microorganisms) to environmental factors (21). In recent years, however, there has been an increasing interest in the modeling of beneficial food microorganisms, including modeling of the functionality of bacteriocin-producing LAB to predict bacteriocin production and activity under food production and processing conditions (13).

Microbial growth may be predicted on the basis of a set of environmental factors by combining kinetic descriptions of the effect of individual factors (e.g., temperature) on biokinetic parameters (e.g., maximum specific growth rate), for instance, by applying the gamma concept (14, 30, 34, 36). The latter approach states that the effects of environmental factors on the maximum specific growth rate of LAB are generally additive, implying that no interaction effects occur. The gamma concept is also applicable under dynamic conditions such as a decreasing pH (15, 17). In addition, the possible use of artificial neural networks (ANNs) in the field of predictive microbiology has recently inspired several studies. While many of these studies used ANNs to investigate food-borne pathogens or spoilage microorganisms (12, 26, 35), few focused on the modeling of beneficial microorganisms (6, 18). The attractiveness of ANNs as empirical modeling schemes lies in their ability to extract, with high accuracy and irrespective of the degree of nonlinearity between system variables, the intrinsic relationships between independent (explanatory) and dependent (response) variables through training of the network on examples representing the phenomenon to be modeled (1). Specifically in food microbiology, ANNs can be used to describe more accurately the interacting effects of environmental factors such as temperature, pH, and NaCl in comparison with classical predictive microbiology models (8).

The aim of the present study was to investigate, in a kinetic way, the influence of temperature, pH, and salt concentration conditions prevailing during Kasseri cheese production on the growth of *S. macedonicus* ACA-DC 198, as well as on the production of macedocin. ANN models were contrasted with an approach based on the gamma concept.

#### MATERIALS AND METHODS

**Strain and media.** *S. macedonicus* ACA-DC 198, isolated from naturally fermented Kasseri cheese, belongs to the ACA-DC culture collection of the Laboratory of Dairy Research (Department of Food Science and Technology, Agricultural University of Athens, Athens, Greece). It was stored at  $-80^{\circ}\text{C}$  in skim milk (10%, wt/vol; Oxoid, Basingstoke, Hampshire, United Kingdom) containing glycerol (10%, vol/vol; Sigma-Aldrich, St. Louis, MO) as a cryoprotectant. Before experimental use, the strain was subcultured twice (inoculum, 1%, vol/vol) in skim milk (10%, mass/vol) containing yeast extract (0.3%, wt/vol; Oxoid) at  $37^{\circ}\text{C}$  for 18 h. Enumeration of *S. macedonicus* ACA-DC 198 bacteria was performed on M17 agar (Biokar Diagnostics, Beauvais, France) after incubation at  $37^{\circ}\text{C}$  for 24 h.

**Quantitative determination of bacteriocin.** The well diffusion assay was used for determination of macedocin activity (29) with *Lactococcus lactis* LMG 6890<sup>T</sup> as the indicator strain. Briefly, 15 ml of M17 agar containing 0.1% (vol/vol; initial population density,  $10^5$  CFU/ml) of a fresh culture of the sensitive strain was poured into a petri dish and wells 5 mm in diameter were made in the solidified medium. The wells were filled with 50  $\mu\text{l}$  of serial twofold dilutions of *S. macedonicus* ACA-DC 198 cell-free culture supernatant in 50 mM sodium phosphate buffer (pH 6.5). Activity was expressed in arbitrary units (AU) corresponding to

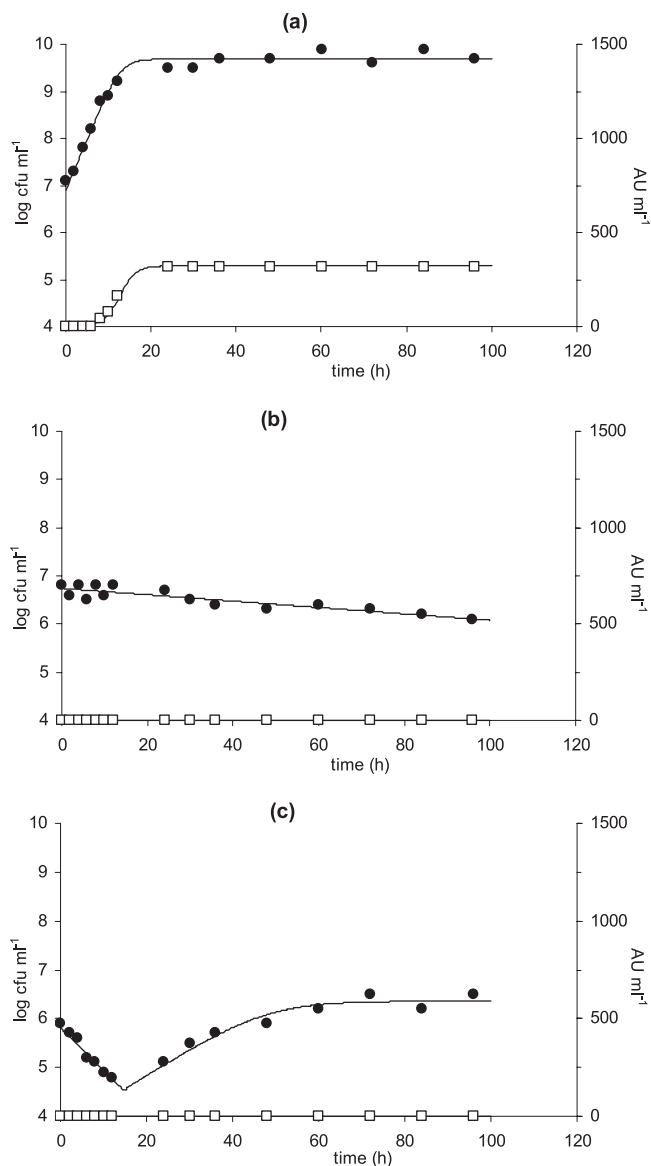


FIG. 1. Modeling of cell growth (closed circles, in log CFU per milliliter) and bacteriocin activity (open squares, in AU per milliliter) by *S. macedonicus* ACA-DC 198 during milk fermentations at  $25^{\circ}\text{C}$ , pH 6.0, and 0% (wt/vol) NaCl (a);  $25^{\circ}\text{C}$ , pH 5.0, and 4% (wt/vol) NaCl (b); and  $40^{\circ}\text{C}$ , pH 6.0, and 4% (wt/vol) NaCl (c). Lines represent the primary model.

50  $\mu\text{l}$  of the highest dilution causing a clear zone of inhibition of the indicator organism.

**Fermentation experiments.** A 2-liter glass fermentor (BioFlo 3000C; New Brunswick Scientific Co., New Brunswick, NJ) with temperature and pH controls was used in batch fermentations in order to study the kinetics of growth of and bacteriocin production by *S. macedonicus* ACA-DC 198. The fermentor contained 1.5 liters of fermentation medium (skim milk, 10%, wt/vol; yeast extract, 0.03%, wt/vol) and was sterilized at  $121^{\circ}\text{C}$  for 10 min. When NaCl was added, the salt solution was sterilized separately and added aseptically to the fermentor. The pH was continuously controlled at a preset value within 0.05 U of the set pH point by automatic addition of 5 N NaOH. The temperature stayed within 0.1 $^{\circ}\text{C}$  of the set point. The fermentor was inoculated with 1% (vol/vol; initial population density,  $10^6$  CFU/ml) of an exponentially growing culture of *S. macedonicus* ACA-DC 198 obtained by propagating a fresh culture twice at  $37^{\circ}\text{C}$  for 12 h. Fermentations were performed without aeration. Slow agitation (200 rpm) was maintained to keep the fermentation broth homogeneous.

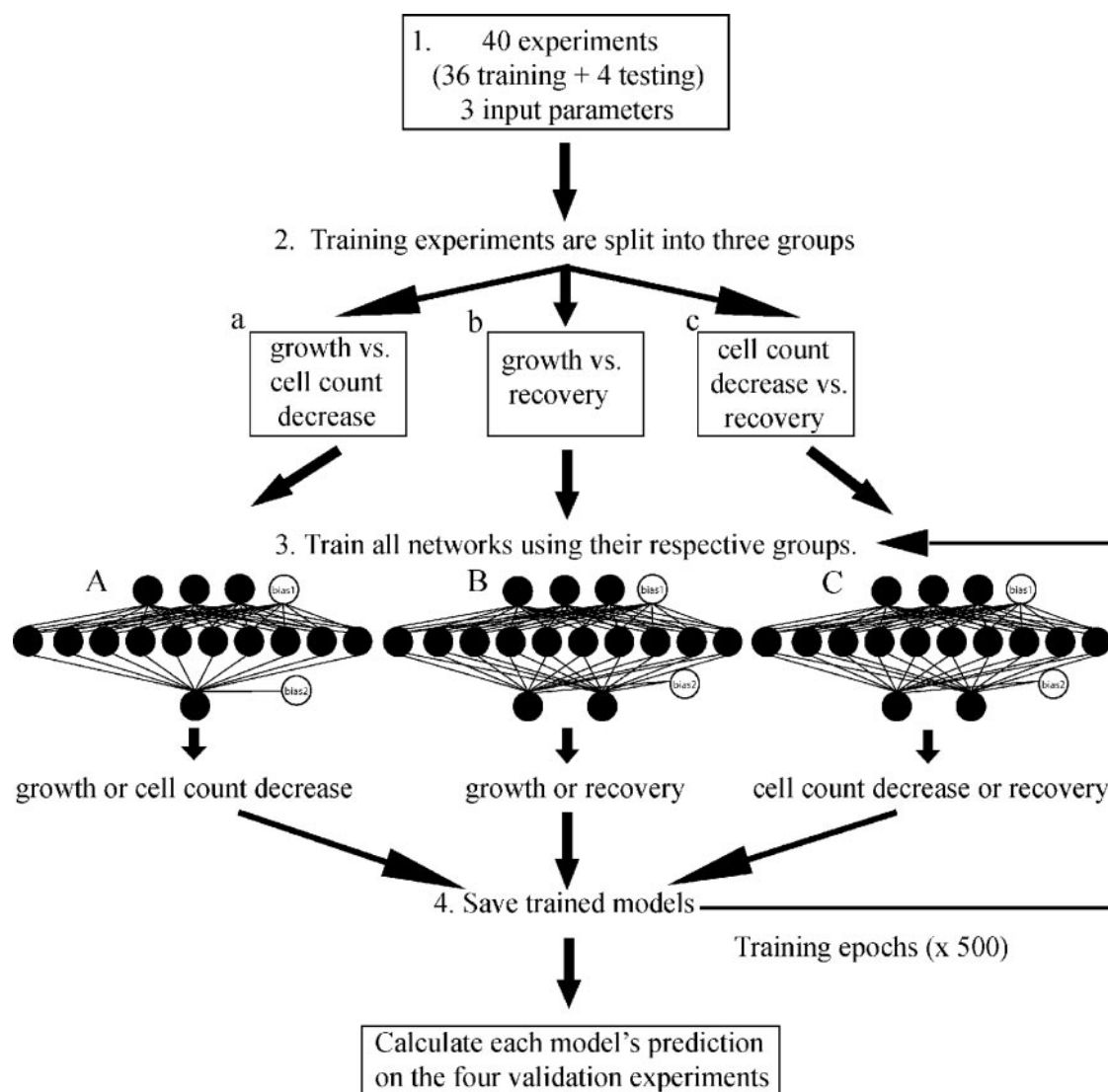


FIG. 2. Flow chart of the training and validation procedures of the ANN models used to predict growth parameters ( $\mu_{\max}$  and  $k_d$ ).

To train the ANN and to construct the gamma-concept-based model, 18 fermentations were performed at 25, 40, and 55°C; constant pHs of 5 and 6; and 0, 2, and 4% (wt/vol) added NaCl. All 18 fermentations were performed in duplicate. To validate the models, four additional fermentations were performed as follows: (i) 30°C, pH 5.5, and 3% (wt/vol) NaCl; (ii) 35°C, pH 5.5, and 0% (wt/vol) NaCl; (iii) 40°C, pH 5, and 3.5% (wt/vol) NaCl; and (iv) 40°C, pH 6, and 3.5% (wt/vol) NaCl.

Additionally, batch fermentations on a small scale (100 ml) were performed to determine the boundary conditions of temperature, pH, and NaCl concentration for *S. macedonicus* ACA-DC 198 growth. Therefore, *S. macedonicus* ACA-DC 198 was grown in skim milk (10%, wt/vol) supplemented with 0.3% (wt/vol) yeast extract, in a free-pH mode, as follows: (i) at 15, 25, 35, 45, and 55°C and an initial pH of 6.5; (ii) at an initial pH of 5.0, 5.6, 6.3, 6.8, 7.4, 8.6, 9.5, or 10.5 and 37°C; and (iii) at 0, 2.5, 3.0, 3.5, 4.5, or 6.5% (wt/vol) NaCl, 37°C, and an initial pH of 6.5.

To take into account natural variations in the 18 duplicated experiments, each of the total of 36 fermentations was used as a separate input instead of taking the average of the replicates. The advantage of this approach is that the method "learns" the expected variation between repetitive experiments and is thus expected to have better prediction performance.

**Sampling and analysis.** Samples were aseptically withdrawn from the fermentor over a period of 96 h. Growth of *S. macedonicus* ACA-DC 198 was assessed by measuring the cell number as CFU per milliliter. The cell number was

determined by 10-fold dilutions in sterile saline, followed by plating on M17 agar (Biokar). For determination of bacteriocin activity, cells were removed by centrifugation (10,000  $\times$  g, 15 min, 4°C). Bacteriocin activity in the cell-free culture supernatant was determined as described above.

**Primary modeling.** The biokinetic parameters used to describe cell growth, i.e., the maximum specific growth rate ( $\mu_{\max}$ , per hour) and the specific cell count decrease rate ( $k_d$ , per hour), were calculated in two ways.

The first method made use of linear regression. The  $\mu_{\max}$  was calculated by integrating the microbial growth equation  $dX/dt = \mu X$ , where  $X$  is the absolute cell count (in CFU per milliliter) at each time point  $t$  (in hours) and  $\mu$  (per hour) is the specific growth rate. The value of  $\mu_{\max}$  is thus given by the slope of the curve  $\ln(X/X_0) = \mu_{\max}(t - t_0) + C$  at its linear part. A similar procedure was used to calculate  $k_d$  in the case of a decrease in cell counts. During the lag phase, which varied between 0 and 2 h, the value of  $\mu$  was set equal to 0.

The second method was based on nonlinear regression with the logistic growth model, as shown by representative examples in Fig. 1. Biokinetic parameters were calculated through fitting of the growth data with the logistic model for growth (14),  $dX/dt = \mu_{\max}(1 - X/X_{\max})X$  (in the case of growth; Fig. 1a), with  $dX/dt = -k_d X$  (in the case of a decrease in cell counts; Fig. 1b), or with a combination of both (in the case of an initial decrease in cell counts followed by recovery; Fig. 1c).  $X_{\max}$  is the maximum achievable cell count (in CFU per milliliter) under a certain set of conditions. This approach focuses on the growth curve as a whole. The equations were solved numerically by Euler integration in

Microsoft Excel and by minimizing the residual sum of squares between model data and experimental data with the solver function.

Specific bacteriocin production  $k_B$  (in mega-AU [MAU] per CFU) was calculated by a similar approach, i.e., by fitting the bacteriocin data with the following growth-based equation:  $dB/dt = k_B dX/dt$ , where  $B$  is the bacteriocin activity (in MAU per liter). Linear regression, as was done for  $\mu_{\max}$  (see above), was not applied for  $k_B$ , considering the lack of suitable data points in the linear part and the large uncertainty about the bacteriocin measurements.

**Secondary modeling. (i) Gamma-concept-based model.** The gamma-concept-based model was used for secondary modeling of  $\mu_{\max}$ . The  $\mu_{\max}$  of an experiment, including growth and/or recovery, can be expressed as a function of an environmental factor,  $z$ , if all other environmental factors are kept constant, as  $\mu_{\max} = (\mu_{\max})_{\text{opt}} \gamma_z$ .

The equations used to quantify the effects of the environmental factors temperature ( $\gamma_T$ ) (25), pH ( $\gamma_{\text{pH}}$ ) (25), and salt concentration ( $\gamma_{\text{NaCl}}$ ) (17) on  $\mu_{\max}$  were

$$\gamma_T = \frac{(T - T_{\min})^2 (T - T_{\max})}{(T_{\text{opt}} - T_{\min})[(T_{\text{opt}} - T_{\min})(T - T_{\text{opt}}) - (T_{\text{opt}} - T_{\max})(T_{\text{opt}} + T_{\min} - 2T)]}$$

$$\gamma_{\text{pH}} = \frac{(\text{pH} - \text{pH}_{\min})(\text{pH} - \text{pH}_{\max})}{(\text{pH} - \text{pH}_{\min})(\text{pH} - \text{pH}_{\max}) - (\text{pH} - \text{pH}_{\text{opt}})^2}$$

$$\gamma_{[\text{NaCl}]} = (1 - [\text{NaCl}]/[\text{NaCl}]_{\max})^n$$

where the min, max, and opt subscripts indicate the minimum, maximum, and optimal values of the parameters studied, respectively, whereas  $n$  is a dimensionless curve-fitting coefficient. The above equations contain two or three parameters, with the minimum and maximum values being based on the small-scale experiments. Finally, the individual gamma functions can be combined as follows:

$$\mu_{\max} = (\mu_{\max})_{\text{opt}} \gamma_T \gamma_{\text{pH}} \gamma_{\text{NaCl}}$$

with, in this specific case,  $(\mu_{\max})_{\text{opt}}$  the value of  $\mu_{\max}$  under optimal conditions of temperature, pH, and salt concentration.

For the secondary modeling of a cell count decrease, being dependent on temperature and salt concentration, no adequate approach was available. Hence, a more empirical method was followed, i.e.,  $k_d = C(1 + a_1 \text{NaCl} + a_2 \text{NaCl}^2)e^{bT} + R$ , where  $C$  (per hour),  $a_1$  ( $[\text{NaCl}]^{-1}$ ),  $a_2$  ( $[\text{NaCl}]^{-2}$ ),  $b$  (per degree Celsius), and  $R$  (per hour) are fitting coefficients.

**(ii) Artificial neural networks. Models for maximum specific growth rate ( $\mu_{\max}$ ) prediction.** An aggregate of three feedforward ANNs was used whose combined outputs were able to discriminate among regular growth, a cell count decrease, and an initial cell count decrease followed by recovery. Each fully connected model comprised an input layer with three nodes corresponding to the temperature, pH, and NaCl parameters; a hidden layer with 10 sigmoidal neurons (using a logarithmic sigmoid transfer function); and an output layer with one or two linear neurons. Bias nodes were also included in the input and hidden layers. The network weights were adapted by using the Levenberg-Marquardt backpropagation algorithm within the Neural Network toolbox of the commercially available software package MATLAB (The MathWorks, Inc., <http://www.mathworks.com>). Each network was trained to discriminate between two different growth conditions: regular growth versus a cell count decrease (network A), regular growth versus an initial cell count decrease followed by recovery (network B), and a cell count decrease versus an initial cell count decrease followed by recovery (network C). The network outputs consisted of the maximum growth-cell count decrease-recovery rates corresponding to each fermentation input.

The three networks were trained on their respective training sets 500 times with random initial conditions in order to average out any variability in the estimated  $\mu_{\max}$  and  $k_d$  values (also see Discussion), thus giving rise to a total of 1,500 ANN models. A flow chart of the training and validation procedures is shown in Fig. 2. After training, the four validation experiments were individually passed through all models and their respective  $\mu_{\max}$  and/or  $k_d$  values were estimated. First, the growth condition that was predicted by the majority of the 1,500 models was selected as the most likely correct prediction. Then, the respective  $\mu_{\max}$  and/or  $k_d$  value for this condition was given by the average output of the model (A, B, or C) that had the minimum standard deviation over the 500 trials.

**Model for specific bacteriocin production ( $k_B$ ) prediction.** As the gamma-concept-based modeling approach has been designed for secondary modeling of growth parameters and not for other kinetic patterns such as bacteriocin production, modeling of  $k_B$  was done exclusively with ANNs. A simple feedforward,

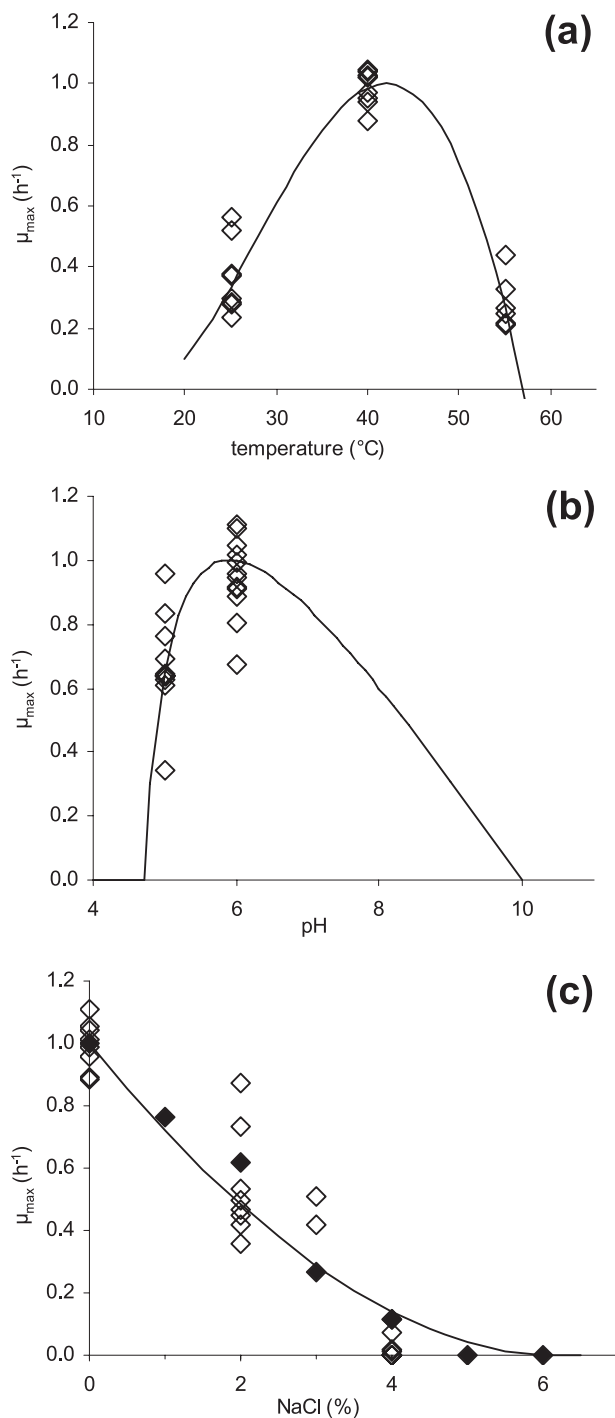


FIG. 3. Modeling of  $\mu_{\max}$  (per hour) as a function of temperature ( $^{\circ}\text{C}$ ) (a), pH (b), and NaCl concentration (percent, wt/vol) (c). The values have been normalized as  $\mu_{\max}/(\mu_{\max})_{\text{opt}}$  for each factor. Full diamonds indicate small-scale experiments. Lines represent the secondary model.

fully connected ANN model was used to predict  $k_B$  by using information about the temperature, pH, and NaCl concentration parameters. The model consisted of an input layer with three nodes, a hidden layer with 10 sigmoidal neurons (using a logarithmic sigmoid transfer function), and an output layer with one linear neuron. Bias nodes were also included in the input and hidden layers. The network weights were adapted by using the backpropagation algorithm. For any



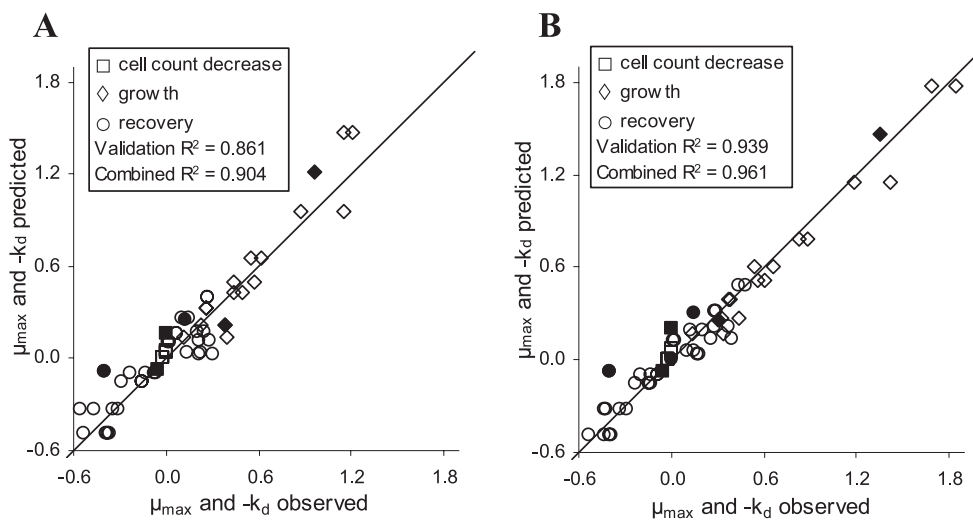


FIG. 4. Observed versus predicted values of  $\mu_{\max}$  (growth and recovery; positive values) and  $-k_d$  (cell count decrease; negative values) per hour, as generated by the gamma-concept-based model, by the linear regression method (A) and the logistic growth model (B), for the estimation of  $\mu_{\max}$  and  $k_d$ . Open symbols correspond to experiments used to construct the model; filled symbols correspond to the validation experiments.

given set of input parameters, the ANN estimated the specific bacteriocin production constant. The model was trained by using the 36 fermentations (18 by 2). Training was repeated 500 times with random initial conditions, and the resulting models were saved. For each validation fermentation, the predicted  $k_B$  value was given by the average of the 500 models' output.

The ANN software code used in this work can be downloaded from <http://www.imbb.forth.gr/people/poirazi/software.html>.

## RESULTS

**Primary modeling.** The data for growth of *S. macedonicus* ACA-DC 198 and bacteriocin activity were successfully modeled on the basis of the logistic growth equation and the model for growth-associated bacteriocin production (Fig. 3 to 6). In addition, the values of  $\mu_{\max}$  and  $k_d$  were also obtained by linear

regression (Fig. 4 and 5). Depending on the environmental conditions, three distinct cases were observed with respect to growth patterns: (i) growth, (ii) a cell count decrease followed by recovery (Fig. 1). Note that in the latter two cases, no bacteriocin activity could be detected because of a lack of growing cells or insufficient growth, respectively. Out of 18 experiments (performed in duplicate), eight cases showed regular growth (at 25°C, pHs 5.0 and 6.0, and NaCl at 0 and 2% [wt/vol]; at 40°C, pHs 5.0 and 6.0, and NaCl at 0 and 2% [wt/vol]), two cases exhibited a cell count decrease (at 25°C, pHs 5.0 and 6.0, and 4% [wt/vol] NaCl) and eight cases showed an initial cell count decrease followed by recovery (at 40°C, pHs 5.0 and 6.0, and NaCl at 4% [wt/vol]

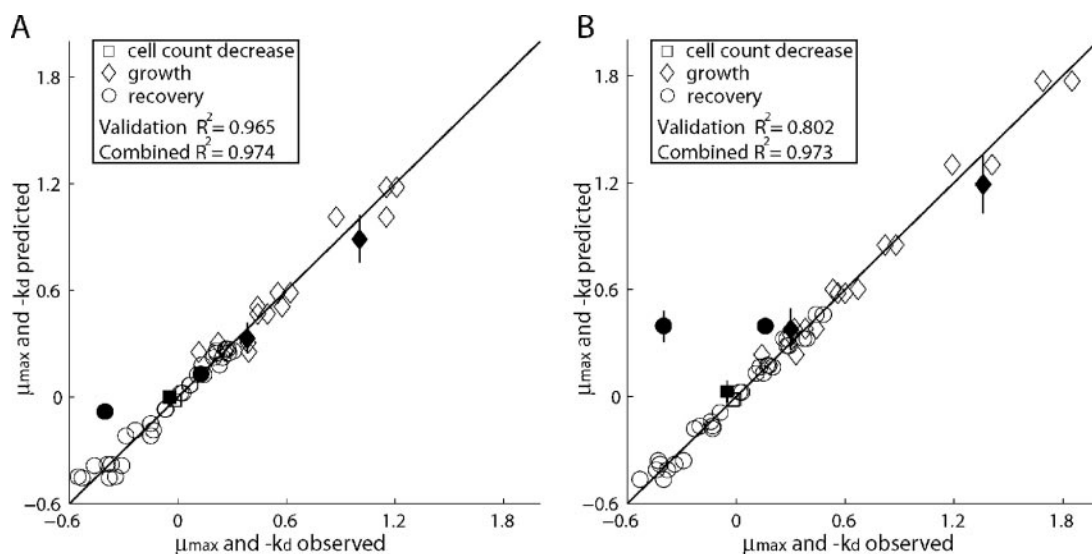


FIG. 5. Observed versus predicted values of  $\mu_{\max}$  (growth and recovery; positive values) and  $-k_d$  (cell count decrease; negative values) per hour, as generated by the ANN models, by the linear regression method (A) and the logistic growth model (B), for the estimation of  $\mu_{\max}$  and  $k_d$ . Open symbols correspond to experiments used to construct the model; filled symbols correspond to the validation experiments. Error bars on the predictions denote standard deviations.

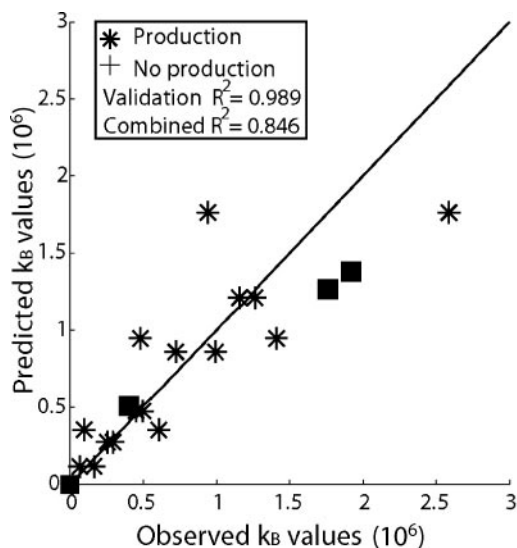


FIG. 6. Observed versus predicted values of  $k_B$  (in MAU per CFU) as generated by the ANN model. Open symbols correspond to experiments used to construct the model; filled symbols correspond to the validation experiments.

and at 55°C, pHs 5.0 and 6.0, and NaCl at 0, 2, and 4% [wt/vol]). At 55°C, a reduction of living cells was observed during the first hours of fermentation, ranging from <1 log CFU/ml (0 and 2% [wt/vol] NaCl, pHs 5 and 6) to 2 log CFU/ml (4% [wt/vol] NaCl, pHs 5 and 6). Thereafter, *S. macedonicus* ACA-DC 198 recovered either fully [at 0 and 2% (wt/vol) NaCl] or partially [at 4% (wt/vol) NaCl] but no further growth or bacteriocin production was observed over a period of 100 h. Recovery seemed to be caused by (combinations of) unfavorable growth conditions, i.e., a high temperature (55°C) and/or high salt concentrations (>2% [wt/vol] NaCl), leading to an initial cell count decrease, after which adapted cells managed to begin population growth. By species-specific PCR (24), it was verified that in the case of recovery the population was not a contaminant but indeed consisted of *S. macedonicus* ACA-DC 198.

In all cases, bacteriocin production paralleled growth (Fig. 1), confirming that bacteriocin production is limited to the active growth phase. When cell growth leveled off, bacteriocin production stopped and remained constant, indicating the absence of bacteriocin-inactivating factors.

**Secondary modeling. (i) Gamma-concept-based model.** The values of  $\mu_{\max}$  and  $k_d$  obtained from the logistic primary model (Fig. 1) were expressed as a function of temperature, pH, and NaCl concentration and modeled with the corresponding functions (Fig. 3). As calculated from the secondary model, growth was fastest in the absence of salt at 42°C and pH 5.9. The fitted minimum and maximum cardinal values for the temperature and pH profiles equaled 15 and 57°C and pHs 4.7 and 10.0, respectively. The maximum salt concentration for growth was set at 6.0 g liter<sup>-1</sup> with a fitting coefficient of 2.0. The kinetic data concerning the individual effects of temperature, pH, and salt concentration were combined in a gamma-concept-based approach.

The actual values of  $\mu_{\max}$  and  $k_d$  were plotted against the

predicted ones (Fig. 4). For the data used to construct the model, correlation coefficients ( $R^2$ ) of 0.915 (linear regression) and 0.970 (logistic growth model) were obtained. For the validation set, these values were 0.861 and 0.939, respectively. The combined correlation coefficient, which is taken on the entire data set (training and validation sets combined), is reported as a measurement of overall performance (0.904 for linear regression and 0.961 for the logistic growth model).

**(ii) Artificial neural networks. Prediction of maximum specific growth rate ( $\mu_{\max}$ ).** With the three types of ANN models, it was possible to discriminate with very high accuracy between the different fermentation conditions that resulted in regular growth, a cell count decrease, or an initial cell count decrease followed by recovery.

The actual versus predicted  $\mu_{\max}$  and  $k_d$  values, as well as the combined correlation coefficient, were again determined for all of the experimental conditions tested (Fig. 5). Because of the nature of the neural network models, the correlation coefficient ( $R^2$ ) achieved on the training set (0.978 for linear regression and 0.991 for the logistic growth model) is in general higher than the one achieved on the validation set (0.965 for linear regression and 0.802 for the logistic growth model). The combined correlation coefficients were similar for linear regression (0.974) and the logistic growth model (0.973).

**Prediction of specific bacteriocin production ( $k_B$ ).** With ANNs, the values of specific bacteriocin production were successfully modeled for all fermentations (Fig. 6). The correlation coefficient equaled 0.825 for the training fermentations, 0.989 for the four validation experiments, and 0.848 for the combined set of data, including both training and validation experiments.

## DISCUSSION

*S. macedonicus* ACA-DC 198, isolated from naturally fermented Greek Kasserri cheese (31), produces an anticlostridial when grown in skim milk supplemented with nitrogen sources (10). This functional property may constitute an ecological advantage for the producing strain during milk fermentation. To elucidate whether this strain can grow and/or survive and produce the lantibiotic macedocin under the harsh conditions (i.e., salt in moisture up to 5%, wt/vol, and scalding of the cheese curd at 75°C) prevailing during Kasserri cheese production, we performed a series of simulation fermentations. The results obtained for both growth and survival, as well as bacteriocin production, were used to develop two different predictive models.

Two types of computational methods were developed to model the population behavior of *S. macedonicus* ACA-DC 198, namely, an ANN model and a model based on the gamma concept approach. The models were constructed with a total of 36 fermentations and were able to predict the  $\mu_{\max}$  and  $k_d$  of four validation experiments. The ANN model accurately discriminated between the different fermentation conditions that resulted in regular growth, a cell count decrease, or an initial cell count decrease followed by recovery. Moreover, the correlation coefficient between predicted and actual growth parameters was generally higher for the ANN model than for the gamma concept model. The ANN model's generalization performance (based on the set of validation experiments) was

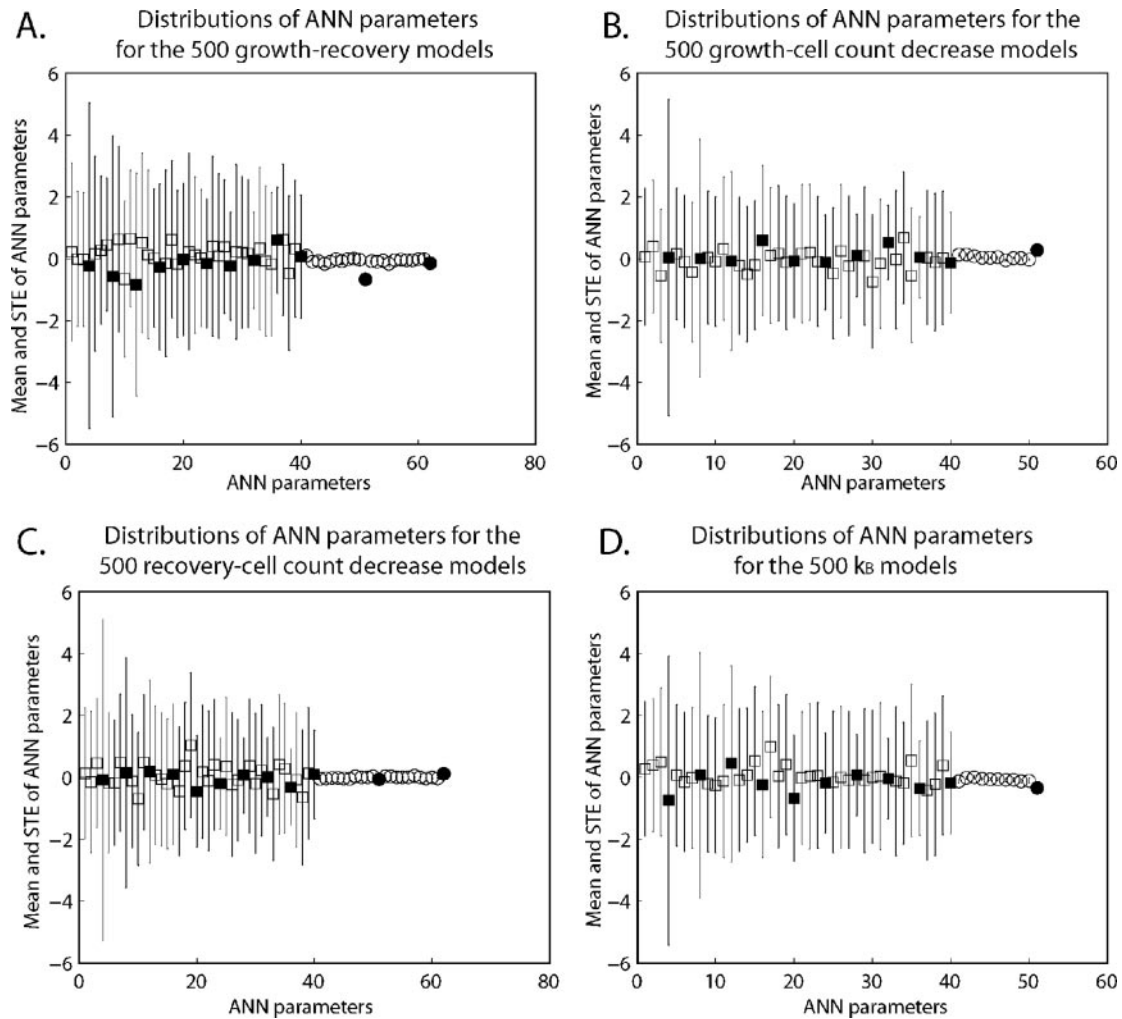


FIG. 7. Mean and standard error (STE) values for the parameters of the four different ANN models taken over 500 training repetitions by the logistic growth model for the estimation of  $\mu_{\max}$ ,  $k_{cb}$ , and  $k_B$ . (A to C) Parameter layer variability in ANN growth models. (D) Parameter variability in the bacteriocin production model. In all panels, squares represent hidden layer parameters and circles denote output layer parameters. Filled symbols depict bias weights, and open symbols represent the remaining network weights. Note that variability is largest in the hidden layer parameters which correspond to nodes with sigmoidal activation functions. Parameters in the output layer that fit into linear nodes show negligible variability over the 500 runs.

higher when the linear regression method was used ( $R^2 = 0.965$ ), even though the method based on the logistic growth model ( $R^2 = 0.802$ ) is a more sophisticated way of fitting bacterial growth curves in general. For the gamma-concept-based model, on the contrary, this result was better for the approach based on the logistic growth model ( $R^2 = 0.939$ ) than for the linear regression method ( $R^2 = 0.861$ ). A possible explanation for this discrepancy might be that the logistic growth model provides a smoother fit of the growth curves, thus reducing variations that may appear in repetitive experiments. While these variations may be due to measurement error, as suggested by the fact that reducing them improves the prediction of the gamma concept approach, they could also reflect natural variations that should be taken into account. Since the ANN models provide stochastic modeling of the bacterial growth parameters—as opposed to the gamma concept method—allowing some variation in repetitive experiments would reduce the probability of overfitting the training

fermentations, thus enhancing the generalization performance. However, one should keep in mind that ANNs require a large number of training samples to achieve optimal generalization performance, and in the present study only 36 samples (18 samples in duplicate) were used for training. To estimate the parameter uncertainty in our ANN models, we report the standard error for each weight of all ANN models trained with the logistic growth model data set. Parameter distributions are particularly useful for single-layer ANNs where there is no trial-dependent variability in the features detected by each network node and weight variability reflects parameter accuracy. However, for two-layer ANNs, each hidden layer neuron may learn a different feature when initialized with different weights, thus resulting in great parameter variability, which does not, however, reflect model uncertainty. Indeed, as depicted in Fig. 7, hidden layer parameters vary significantly over the 500 trained models of each type. However, the output layer parameters are very similar in all models. This variability is due

to (i) the nonspecificity of features learned by the different hidden layer neurons on each randomly initialized run and (ii) the nonlinear nature of the sigmoidal transfer functions used in the hidden layer of the ANN models, where combinations of different weight values can result in the same hidden layer outcome. This is not true for the output nodes of the models, which implement linear transfer functions. In this case, even small variations in the weights would induce large differences in the model predictions. Thus, the insignificant variability in the output layer parameters (Fig. 7, circles), along with the small standard deviations in the predictions of these models (Fig. 5), indicates that true parameter uncertainty is very small and supports the robustness and accuracy of our models.

In addition to the results obtained for growth behavior, the ANN model was able to accurately predict the specific bacteriocin production ( $k_B$ ) values for the four validation experiments. This is very important, since no satisfactory gamma-concept-based model is available to tackle this problem successfully. Hence, ANN models allow prediction of both growth-related and non-growth-related parameters.

Despite the somewhat lower correlations obtained and the difficulty in modeling parameters other than the maximum specific growth rate, gamma concept models provide an interesting complementary approach to study and predict the growth behavior of LAB (15, 17). They clearly present the effect of individual environmental parameters and yield interesting and easily interpretable information based on cardinal values. In this way, information about the optimum, minimum, and maximum values of an environmental factor (e.g., temperature) on the maximum specific growth rate is obtained. Moreover, they provide information on the relative importance of the environmental factors under a certain set of experimental conditions. The gamma values actually indicate, when multiplied by 100, the percentage of growth inhibition caused by each factor.

The present study showed that ANNs, although less suitable for biological interpretation, outperform a more traditional model based on the gamma concept in describing the growth behavior and bacteriocin production capacity of *S. macedonicus* ACA-DC 198. These findings are in agreement with other modeling studies that contrasted ANNs with traditional regression techniques (6, 7, 8, 11, 19). Geeraerd et al. (8) showed that the ANN model describing growth parameters as a function of temperature, pH, and NaCl concentration was far more accurate than the polynomial or any other linear modeling technique. Similarly, ANN predictive growth models of *Leuconostoc mesenteroides* showed a higher degree of accuracy and bias compared with response surface methodology (7). This can be explained by the flexible basis functions used in ANN modeling compared to the fixed basis functions of a polynomial or any other linear modeling technique (12). However, when choosing a predictive model it is important to ensure not only high performance but also interpretability and reduced complexity (number of parameters) without loss of accuracy. In this study, the aggregate ANN models used for modeling bacterial growth and bacteriocin production had a total of 175 and 51 modifiable parameters (weights), respectively; the model based on the gamma concept for bacterial growth only required a total of 14 parameters.

Because of its thermophilic character and bacteriocin pro-

duction, *S. macedonicus* ACA-DC 198 is of particular interest for application as a bioprotective culture in the production of hard-cooked cheeses. Its bacteriocin-producing ability in the mesophilic temperature range offers potential uses in other cheese varieties as well. The presence of the bacteriocin upon prolonged incubation suggests that the protective effect may be present during the ripening period as well. All of these properties may make a macedocin producer applicable in the prevention of clostridial spoilage, which has to be validated in challenge tests. Because of its moderate acidification and low proteolytic and lipolytic activities, the best strategy for cheese making is to use the strain as a bioprotective adjunct starter in addition to a classical technological starter culture. The aim of the present study was to reveal the kinetic behavior of the strain as a function of environmental factors (temperature, pH, and salt concentration). Future research will have to expand the present findings to the complexity of mixed starter cultures and actual cheese making.

#### ACKNOWLEDGMENTS

The present work was cofinanced by the European Social Fund and National Resources EPEAEK and YPEPTH. We also acknowledge the financial support of the Research Council of the Vrije Universiteit Brussel, the Fund for Scientific Research-Flanders (FWO), and the Institute for the Encouragement of Scientific Research in the Industry (IWT). Frédéric Leroy was supported by a postdoctoral fellowship of the FWO, and Panayiota Poirazi was supported by the EMBO Young Investigator Program.

#### REFERENCES

1. Basheer, I., and M. N. Hajmeer. 2000. Artificial neural networks: fundamentals, computation, design and application. *J. Microbiol. Methods* **43**:3-31.
2. Buchanan, R. L. 1993. Predictive food microbiology. *Trends Food Sci. Technol.* **4**:6-11.
3. Cogan, T. M., and C. Hill. 1993. Cheese starter cultures, p. 193-255. *In* P. F. Fox (ed.), *Cheese: chemistry, physics and microbiology*, vol. 1, second edition. Chapman & Hall, Ltd., London, United Kingdom.
4. Delves-Broughton, J., P. Blackburn, R. J. Evans, and J. Hugenholtz. 1996. Applications of the bacteriocin, nisin. *Antonie Leeuwenhoek* **69**:193-202.
5. De Vuyst, L., and E. J. Vandamme. 1994. Antimicrobial potential of lactic acid bacteria, p. 91-142. *In* L. De Vuyst and E. J. Vandamme (ed.), *Bacteriocins of lactic acid bacteria: microbiology, genetics and applications*. Blackie Academic and Professional, London, United Kingdom.
6. García-Gimeno, R. M., C. Hervás-Martínez, and M. I. de Sílóniz. 2002. Improving artificial neural networks with a pruning methodology and genetic algorithms for their application in microbial growth prediction in food. *Int. J. Food Microbiol.* **72**:19-30.
7. García-Gimeno, R. M., C. Hervás-Martínez, R. Rodríguez-Pérez, and G. Zurera-Cosano. 2005. Modelling the growth of *Leuconostoc mesenteroides* by artificial neural networks. *Int. J. Food Microbiol.* **105**:317-332.
8. Geeraerd, A. H., C. H. Herremans, C. Cenes, and C. Van Impe. 1998. Application of artificial neural networks as a non-linear modular modeling technique to describe bacterial growth in chilled food products. *Int. J. Food Microbiol.* **44**:49-68.
9. Georgalaki, M. D., P. Sarantinopoulos, E. S. Ferreira, L. De Vuyst, G. Kalantzopoulos, and E. Tsakalidou. 2000. Biochemical properties of *Streptococcus macedonicus* strains isolated from Greek Kasser cheese. *J. Appl. Microbiol.* **88**:817-825.
10. Georgalaki, M. D., E. Van den Berghe, D. Kritikos, B. Devreese, J. Van Beeumen, G. Kalantzopoulos, L. De Vuyst, and E. Tsakalidou. 2002. Macedocin: a food grade lantibiotic produced by *Streptococcus macedonicus* ACA-DC 198. *Appl. Environ. Microbiol.* **68**:5891-5903.
11. Hajmeer, M. N., I. A. Basheer, and Y. M. Najjar. 1997. Computational neural networks for predictive microbiology. II. Application to microbial growth. *Int. J. Food Microbiol.* **34**:51-66.
12. Jeyamkandan, S., D. S. Jayas, and R. A. Holley. 2001. Microbial growth modelling with artificial neural networks. *Int. J. Food Microbiol.* **64**:343-354.
13. Leroy, F., B. Degeest, and L. De Vuyst. 2002. A novel area of predictive modelling: describing the functionality of beneficial microorganisms in foods. *Int. J. Food Microbiol.* **73**:251-259.
14. Leroy, F., and L. De Vuyst. 1999. Temperature and pH conditions that prevail during fermentation of sausages are optimal for production of the antilisterial bacteriocin sakacin K. *Appl. Environ. Microbiol.* **65**:974-981.



15. Leroy, F., and L. De Vuyst. 2003. A combined model to predict the functionality of the bacteriocin-producing *Lactobacillus sakei* strain CTC 494. *Appl. Environ. Microbiol.* **69**:1093–1099.
16. Leroy, F., and L. De Vuyst. 2004. Functional lactic acid bacteria starter cultures for the food fermentation industry. *Trends Food Sci. Technol.* **15**:67–78.
17. Leroy, F., and L. De Vuyst. 2005. Simulation of the effect of sausage ingredients and technology on the functionality of the bacteriocin-producing *Lactobacillus sakei* CTC 494 strain. *Int. J. Food Microbiol.* **100**:141–152.
18. Lopes, M. F., C. I. Pereira, F. M. Rodrigues, M. P. Martins, M. C. Mimoso, T. C. Barros, J. J. Figueiredo Marques, R. P. Tenreiro, J. S. Almeida, and M. T. Barreto Crespo. 1999. Registered designation of origin areas of fermented food products defined by microbial phenotypes and artificial neural networks. *Appl. Environ. Microbiol.* **65**:4484–4489.
19. Lou, W., and S. Nakai. 2001. Artificial neural network-based predictive model for bacterial growth in a simulated medium of modified-atmosphere-packed cooked meat products. *J. Agric. Food Chem.* **49**:1799–1804.
20. McDonald, K., and D. W. Sun. 1999. Predictive food microbiology for the meat industry: a review. *Int. J. Food Microbiol.* **52**:1–27.
21. Nauta, M. J. 2002. Modeling bacterial growth in quantitative microbiological risk assessment: is it possible? *Int. J. Food Microbiol.* **73**:297–304.
22. Nettles, C. G., and S. F. Barefoot. 1993. Biochemical and genetic characteristics of bacteriocins of food-associated lactic acid bacteria. *J. Food Prot.* **56**:338–356.
23. O'Sullivan, L., R. P. Ross, and C. Hill. 2002. Potential of bacteriocin-producing lactic acid bacteria for improvements in food safety and quality. *Biochimie* **84**:593–604.
24. Papadelli, M., E. Manolopoulou, G. Kalantzopoulos, and E. Tsakalidou. 2003. Rapid detection and identification of *Streptococcus macedonicus* by species-specific PCR and DNA hybridisation. *Int. J. Food Microbiol.* **81**:231–239.
25. Rosso, L., J. R. Lobry, S. Bajard, and J. P. Flandrois. 1995. A convenient model to describe the combined effects of temperature and pH on microbial growth. *Appl. Environ. Microbiol.* **61**:610–616.
26. Siripatrawan, U., J. E. Linz, and B. R. Harte. 2004. Rapid method for prediction of *Escherichia coli* numbers using an electronic sensor array and an artificial neural network. *J. Food Prot.* **67**:1604–1609.
27. Steffen, C., P. Eberhard, J. O. Bosset, and M. Rüegg. 1993. Swiss-type varieties, p. 83–110. In P. F. Fox (ed.), *Cheese: chemistry, physics and microbiology*, vol. 2, second edition. Chapman & Hall, Ltd., London, United Kingdom.
28. Stiles, M. E. 1996. Biopreservation by lactic acid bacteria. *Antonie Leeuwenhoek* **70**:331–345.
29. Tagg, J. R., and A. R. McGiven. 1971. Assay system for bacteriocins. *Appl. Microbiol.* **21**:943–946.
30. te Giffel, M. C., and M. H. Zwietering. 1999. Validation of predictive models describing the growth of *Listeria monocytogenes*. *Int. J. Food Microbiol.* **46**:135–149.
31. Tsakalidou, E., E. Zoidou, B. Pot, W. Ludwig, L. Wassil, L. A. Devriese, G. Kalantzopoulos, K. H. Schleifer, and K. Kersters. 1998. Identification of streptococci from Greek Kasserli cheese and description of *Streptococcus macedonicus* sp. nov. *Int. J. Syst. Bacteriol.* **48**:519–527.
32. Van den Berghe, E., G. Skourtas, E. Tsakalidou, and L. De Vuyst. 2006. *Streptococcus macedonicus* ACA-DC 198 produces the lantibiotic, macedocin, at temperature and pH conditions that prevail during cheese manufacture. *Int. J. Food Microbiol.* **107**:138–147.
33. Walstra, P., A. Noomen, and T. J. Geurts. 1993. Dutch-type varieties, p. 39–82. In P. F. Fox (ed.), *Cheese: chemistry, physics and microbiology*, vol. 2, second edition. Chapman & Hall, Ltd., London, United Kingdom.
34. Wijtzes, T., F. M. Rombouts, M. L. Kant-Muermans, K. van 't Riet, and M. H. Zwietering. 2001. Development and validation of a combined temperature, water activity, pH model for bacterial growth rate of *Lactobacillus curvatus*. *Int. J. Food Microbiol.* **63**:57–64.
35. Zurera-Cosano, G., R. M. Garcia-Gimeno, M. R. Rodriguez-Perez, and C. Hervas-Martinez. 2005. Validating an artificial neural network model of *Leuconostoc mesenteroides* in vacuum packaged sliced cooked meat products for shelf life estimation. *Eur. Food Res. Technol.* **221**:717–724.
36. Zwietering M. H., J. C. de Wit, and S. Notermans. 1996. Application of predictive microbiology to estimate the number of *Bacillus cereus* in pasteurised milk at the point of consumption. *Int. J. Food Microbiol.* **30**:55–70.