Detection and Discrimination of *Enterobacter sakazakii* Strains by Mid-infrared Spectroscopy and Multivariate Statistical Analyses

MENGSHE LIN\textsuperscript{a,*}, MURAD AL-HOLY\textsuperscript{b}, DEWI SETIADY\textsuperscript{c}, BARBARA A. RASCO\textsuperscript{c}

\textsuperscript{a} Food Science Program, 256 William Stringer Wing, University of Missouri, Columbia, Missouri 65211, USA;

\textsuperscript{b} Department of Clinical Nutrition and Dietetics, Faculty of Allied Health Sciences, Hashemite University, Zarqa-Jordan;

\textsuperscript{c} Department of Food Science and Human Nutrition, Box 646376, Washington State University, Pullman, WA 99164, USA;

* Corresponding author. Tel.: (573) 884-6718; fax (573) 884-7964. E-mail address: linme@missouri.edu.
Abstract

Enterobacter sakazakii, a member of the family Enterobacteriaceae, can cause rare but
life-threatening diseases such as meningitis in infants and neonates. Infant formula has
been implicated in several E. sakazakii outbreaks as the vehicle of contamination. Fourier
transform infrared (FT-IR) spectroscopy was used to detect and discriminate between
eight E. sakazakii strains (ATCC 29004, ATCC 12868, FSM 292, FSM 287, 3, 55, 2.39,
and 2.68), two E. cloacae strains (Rev 1210 and Rev 1343), three Escherichia coli strains
(ATCC 25922, K12 2B and B E4a), and two Klebsiella pneumoniae strains (Revco 41
and K1a). FT-IR vibrational combination bands reflect subtle compositional differences
in the cell membranes of E. sakazakii strains, especially in the region between 1200 and
900 cm\(^{-1}\) which contains absorption bands from carbohydrates. Two multivariate
statistical analyses including principal component analysis (PCA) and soft independent
modeling of class analogy (SIMCA) were used to analyze the spectral data. Distinctive
segregations among spectral sample clusters of E. sakazakii and other strains were
observed by PCA. Based on SIMCA analysis, 90% of E. sakazakii spectral samples were
correctly classified, while 88% of E. cloacae, 91% of E. coli, and 91% of K. pneumoniae
samples were correctly classified, suggesting that this technique may provide a tool to
detect E. sakazakii strains rapidly and accurately.

KEYWORDS: Enterobacter sakazakii; FT-IR; spectroscopy; PCA; SIMCA
INTRODUCTION

Enterobacter sakazakii, a member of the family Enterobacteriaceae, is a rod-shaped, motile, non-sporeforming, Gram negative, facultative anaerobe (1) and until 1980 was classified as an Enterobacter cloacae. E. sakazakii can be distinguished from E. cloacae strains by difference in biochemical features such as pigment production (2, 3). E. sakazakii can cause rare but life-threatening diseases such as meningitis, bacteremia, and necrotizing enterocolitis in infants and neonates (4).

The first case of neonatal meningitis caused by E. sakazakii was reported in two infants in 1958 in the United Kingdom (5). Wound infections in immunocomprised adults by this organism have also been reported (6, 7). Mortality rates of E. sakazakii infections can be as high as 80% (3). In 2002, the International Commission for Microbiological Specifications for Foods described this emerging pathogen as a “severe hazard for restricted populations, causing a life threatening or substantial chronic sequelae or illness of long duration” (8).

Infant formula has been implicated in both E. sakazakii outbreaks and sporadic cases as the vehicle of contamination (1, 3). This microorganism has also been isolated from various food sources including cheese, rice seed, cured meat, vegetables, as well as some environmental sources such as the guts of fly larvae (4, 9). Some E. sakazakii strains are able to produce viscous capsular material and forms a biofilm on the surface of materials, work surfaces and equipment commonly used for manufacturing infant formulas and foods including latex, polycarbonate, silicon, and stainless steel (4, 10, 11).

The US Food and Drug Administration has used a lengthy procedure to check for the presence of E. sakazakii in dry infant formula milk (IFM). This procedure includes rehydrating IFM in sterile distilled water, followed by selective enrichment in Enterobacteriaceae enrichment broth, then by streaking on violet red bile glucose agar (VRBG) and on tryptic soy agar (TSA), respectively. Finally, presumptive yellow pigmented colonies on TSA are confirmed using the API 20E biochemical identification system (12). Several other media have recently been developed to detect for E. sakazakii in IFM such as OK medium (13), a chromogenic medium (11), and PCR based assay such as the BAX® System (Qualicon, Wilmington, DE, USA). But these methods are laborious and time-consuming because a selective enrichment step is indispensable before plating on a selective-differential medium. In addition, further biochemical tests for confirmation may take days to obtain results.

Fourier transform infrared (FT-IR) spectroscopy is a technique that has the potential to both differentiate and characterize bacterial strains since biochemical information about cellular components including water, proteins and peptides, carbohydrates, murein, nucleic acids, and phospholipids are detectable (14, 15). With the latest developments in multivariate statistical analyses, FT-IR spectroscopy has become increasingly more applicable and useful for the classification and identification of bacteria with minimal sample preparation (16). FT-IR has recently been widely used to detect, discriminate, and classify various microorganisms, such as Campylobacter species (17), yeast (18), Escherichia coli O157:H7 (15, 19), cyanobacteria (20), lactic acid bacteria (21), Bacillus spp. (14), and to monitor lymphocyte cell line interacted with metal compounds (22) and lung cell responses to health hazards (23). Use of an attenuated total reflectance (ATR) cell with FT-IR spectroscopy provides a technique for measuring a wide variety of samples since little or no sample preparation is required.
The objective of this study was to establish a novel and accurate method to detect and discriminate *E. sakazakii* strains from other selected bacterial strains of the family *Enterobacteriaceae* using FT-IR spectroscopy in combination with multivariate statistical tools including principal component analysis (PCA) and soft independent modeling of class analogy (SIMCA).

**MATERIALS AND METHODS**

Preparation of Bacterial Cultures. The bacterial strains used in this study were obtained from the culture collection of the Department of Food Science and Human Nutrition at Washington State University, including eight *Enterobacter sakazakii* strains (ATCC 29004, ATCC 12868, FSM 292, FSM 287, 3, 55, 2.39, and 2.68), two *Enterobacter cloacae* strains (Rev 1210 and Rev 1343), three *Escherichia coli* strains (ATCC 25922, K12 2B and B E4a), and two *Klebsiella pneumoniae* strains (Revco 41 and K1a).

The bacterial strains were activated by being streaked onto VRBG (Becton Dickinson, Sparks, MD, U.S.A.) and grown at 37°C for 48 h. A representative colony was then picked and inoculated into brain heart infusion (BHI) broth (Becton Dickinson) and incubated at 37°C for 24 h. Bacterial cells were enumerated using a standard spread plating method on TSA (Becton Dickinson). The plates were incubated at 37°C for 48 h on TSA agar. Typical colonies appeared purple surrounded by a purple halo of precipitated bile acids after 48-72 h incubation at 37°C on VRBG agar.

**FT-IR Spectroscopy.** Bacterial cells were harvested by filtering 15 ml bacterial/broth suspension through an aluminum oxide membrane filter (Anodisc, 0.2 µm pore size and 25 mm OD, Whatman Inc., Clifton, NJ, U.S.A.). A Whatman® vacuum glass membrane filter holder was used. Also, as a control, 15 ml of broth without bacteria was filtered through an Anodisc membrane. The filters were then air dried under laminar flow at room temperature for 10 min to obtain a dry, homogeneous film of bacterial cells. Three membrane samples (N=3) were prepared for each bacterial strain. FT-IR spectra were collected using a ThermoNicolet Avatar 380 FT-IR spectrometer (Thermo Scientific Inc., Madison, WI, U.S.A.). During measurement, the membrane filters coated with bacterial cells were placed in direct contact with an ATR diamond crystal. The resolution was set at 4 cm⁻¹ with each spectrum composed of an average of 64 separate scans. In this study, ten spectra were collected from different locations on each membrane and thirty spectra were acquired at room temperature for each culture.

Study of Reproducibility. To study the reproducibility of this FT-IR spectroscopic method, two *E. sakazakii* isolates (e.g. 3 and 2.39) were selected for an additional experiment that was conducted at a different time. Bacterial cultures and Anodisc membranes were prepared as mentioned above. Ten more spectra were acquired from each sample following the experimental design described above.

**Data Analysis.** Data analysis was conducted using OMNIC (Thermo Electron Inc.) and Delight version 3.2.1 (DSquared Development Inc., LaGrande, OR, USA) software. First, original spectra were undergone an “ATR correction”. In ATR analysis, the smaller the wavenumber (lower frequency), the deeper the energy penetrates into the sample, leading to stronger band intensities. This problem can be solved by applying “ATR correction”. Following this step, an “automatic baseline correction” and “normalization” were applied to ATR-corrected spectra to adjust any tilted baselines and
compensate for the pathlength effect. This is important because during FT-IR ATR spectral measurement, sample thickness can vary, for example, a thick bacterial biofilm can absorb more infrared energy than a thin biofilm, resulting in greater peak heights. Normalizing the sample spectra compensates for this pathlength effect and makes peak heights comparable. Other data pre-processing algorithms were also employed with the spectra smoothed using a Gaussian function over 6 cm\(^{-1}\). This was followed by a second derivative transformation with a gap value of 10 cm\(^{-1}\). After data pre-processing, principal component analysis (PCA) and soft independent modeling of class analogy (SIMCA) were employed (24).

RESULTS AND DISCUSSION

In this study, mid-infrared spectroscopy could be employed to detect and discriminate \textit{E. sakazakii} using an FT-IR ATR technique to obtain spectral features of bacterial samples. Figure 1 shows typical FT-IR ATR spectra (after data pre-processing) of four selected representative bacterial strains (\textit{E. sakazakii} ATCC 12868, \textit{E. cloacae} Rev 1210, \textit{E. coli} B E4a, and \textit{K. pneumoniae} K1a). FT-IR spectroscopy reflects important absorption and combination bands from various biochemical functional groups in the FT-IR region between 4000 and 600 cm\(^{-1}\). These bands arise from major bacterial cellular components such as water, lipids, proteins, carbohydrates, and nucleic acids. For example, the prominent absorption peaks around 3400 cm\(^{-1}\) are mainly from O-H stretching of water. The absorption peaks at 2960, 2929, and 1740 cm\(^{-1}\) are believed to be from fatty acids and lipids in the bacterial cell membrane.

Figure 2 shows typical second derivative transformation (10-point gap) of FT-IR ATR absorbance spectra of those four bacterial strains (\textit{E. sakazakii} ATCC 12868, \textit{E. cloacae} Rev 1210, \textit{E. coli} B E4a, and \textit{K. pneumoniae} K1a) in the region between 1800 and 800 cm\(^{-1}\). The FT-IR spectral region between 1500 and 800 cm\(^{-1}\) conveys important bacterial strain-dependent information that can be linked to a particular functional group or molecular bond (25). This “fingerprint region” contains deformation-, bending- and ring-vibrations from various biochemical functional groups, making it possible to identify and characterize microorganisms. However, absorption bands in the FT-IR region commonly overlap. A major challenge is how to separate overlapped absorption bands and single out useful spectral information. In this regard, data pre-processing algorithms, such as second derivative transformation were used to address this problem and enhance spectral differences between bacterial strains.

Figure 3 shows second derivative transformation (10-point gap) of FT-IR ATR spectra of two \textit{E. sakazakii} strains (FSM 287 and 3) and \textit{E. cloacae} Rev 1210 at wavenumbers between 1200 and 900 cm\(^{-1}\). This region reflects a large number of important absorption peaks from stretching vibrations of the carbohydrates, phosphate, and nucleic acids, presenting clear strain-dependent differences between the bacterial strains. Prominent peaks were observed around 1175, 1125, 1080, 1052, 970, 920 cm\(^{-1}\). The peak around 1080 cm\(^{-1}\) is from ring vibrations of glucose; while the peak around 1052 cm\(^{-1}\) can be linked to C-O, C-O-C functional groups of polysaccharides and disaccharides such as sucrose and trehalose (23, 26). \textit{E. sakazakii} bacterial cells contain some important carbohydrate compounds, such as lipopolysaccharide (LPS) on the surface of bacterial outer membrane; trehalose which is a non-reducing disaccharide of glucose and plays a critical role in protecting bacteria against drying by stabilizing...
phospholipid membranes and proteins (27). In Figure 3, there were some clear differences between these three strains, especially in absorption bands associated with carbohydrates.

A mean centered PCA analysis was conducted on the second derivative transformed spectra of bacterial strains over the entire working range (4000-600 cm\(^{-1}\)) as well as some specific regions. Figure 4 shows the 3-dimensional PCA clustering results for four bacterial species (E. sakazakii, E. cloacae, E. coli, and K. pneumoniae). Clear segregation with distinct sample clusters were observed between these four bacterial species, demonstrating that PCA could differentiate microorganisms at the species level based upon differences in FT-IR spectral features. Similarly, clear segregations were also observed in 2-dimensional PCA results (Figure 5) with distinct spectral samples clustering between nine bacterial strains.

Figure 6 shows the PCA clustering results for eight E. sakazakii strains (ATCC 29004, ATCC 12868, FSM 292, FSM 287, 3, 55, 2.39, and 2.68). These results were based on the PCA analysis of spectral data in the region between 1200-900 cm\(^{-1}\). However, satisfactory segregations for eight E. sakazakii strains were not obtained when analyzing spectral data over the whole working range (4000-600 cm\(^{-1}\)), indicating that most significant differences between E. sakazakii strains may lie in the carbohydrate profiles of their cell membrane, especially, exopolysaccharide (biofilm) forming capability.

SIMCA is a commonly used statistical model that is based upon the generation of a PCA model from the training set. Each sample in the test set is then validated against the model and assigned to a class according to its analogy to the training samples (28). In this study, each of four categories (e.g. E. sakazakii, E. cloacae, E. coli, and K. pneumoniae) was chosen as the training set to construct a model, and then this model was used to predict the remaining three categories. Average values for prediction of each category were obtained based on SIMCA: 90% of E. sakazakii spectral samples were correctly classified, while 88% of E. cloacae, 91% of E. coli, and 91% of K. pneumoniae samples were correctly classified, demonstrating that FT-IR coupled with SIMCA may provide an accurate tool to detect and classify Enterobacteriaceae strains.

Reproducibility and lower limit of detection are two major concerns about detecting and classifying microorganisms with spectroscopic methods. Factors affecting FT-IR spectral reproducibility include growth medium, growth conditions (e.g. incubation time and temperature, etc), sample preparation, and FT-IR spectral acquisition. Consistent and reproducible spectra can be obtained through a careful standardization of experimental procedures including medium selection, culture time and temperature, sample preparation, spectral measurements, and data pretreatment and analysis. Schmalreck et al. (29) reported that some strains of yeast isolated from patients were correctly identified repetitively over a 3-year period using FT-IR spectroscopy. In this study, an additional experiment was conducted with two selected E. sakazakii isolates (3 and 2.39) to test the reproducibility of FT-IR spectroscopy. SIMCA classification results for two E. sakazakii isolates (3 and 2.39) were obtained (Figure 7). Interestingly, 87% of E. sakazakii 3 spectra collected were correctly classified, while 90% of E. sakazakii 2.39 spectra were correctly classified. These results demonstrate the capability of FT-IR spectroscopy to classify Enterobacter strains with an appreciable degree of
reproducibility. On the other hand, lower level of detection for FT-IR could reach $10^3$ CFU/ml in apple juice (30) and $10^4$ CFU/mm² on Anodisc membrane (15).

Despite the fact that the incubation time of 24 h was chosen in this study to obtain enough stationary phase bacterial cells, FT-IR holds promise as a simple and sensitive alternative to traditional methods by replacing time consuming biochemical testing for confirmation, which is one of three essential steps (e.g. enrichment, plating, and confirmation) in the traditional culture based methods. Nevertheless, further studies are needed to investigate spectral features of *E. sakazakii* under various growth conditions such as exponential growth phase as well as to study the applicability of using FT-IR to detect and discriminate this foodborne pathogen under different circumstances.

**CONCLUSION**

FT-IR analysis coupled with multivariate statistical analyses is capable of detecting and classifying strains of *E. sakazakii* from other bacteria at the species and strain levels. Subtle differences in the cell membrane composition, especially carbohydrates, of different strains are reflected in spectral features, providing the basis for detection and classification of *E. sakazakii* strains. However, additional efforts are needed to investigate the relationship between spectral features in FT-IR region with biochemical characteristics such as biofilm forming abilities that are associated with the pathogenicity of these bacterial strains and their control within a processing environment.

**ACKNOWLEDGEMENTS**

Special thanks are given to Dr. Dong-Hyun Kang and Mr. Peter Gray at Washington State University for kindly providing bacterial strains.
LITERATURE CITED


(16) Naumann, D. FT-infrared and FT-Raman spectroscopy in biomedical research. In: Gremlich, H.-U.; Yan, B. (Eds.), Infrared and Raman spectroscopy


Figure 1. Typical FT-IR ATR absorbance spectra (after data pre-processing) of four strains (E. sakazakii ATCC 12868, E. cloacae Rev 1210, E. coli B E4a, and K. pneumoniae K1a).

Figure 2. Second derivative transformation (10-point gap) of FT-IR ATR absorbance spectra of four bacterial strains (E. sakazakii ATCC 12868, E. cloacae Rev 1210, E. coli B E4a, and K. pneumoniae K1a) in the region between 1800 and 800 cm\(^{-1}\).

Figure 3. Second derivative transformation (10-point gap) of FT-IR ATR spectra of three Enterobacter strains (E. sakazakii 3, E. sakazakii FSM 287, and E. cloacae Rev 1210) in the region between 1200 and 900 cm\(^{-1}\).

Figure 4. Three-dimensional PCA clustering results for four bacterial species (E. sakazakii, E. cloacae, E. coli, and K. pneumoniae). Parameters: smoothing: 6 cm\(^{-1}\), 2\(^{nd}\) derivative transformation: 10 cm\(^{-1}\), 5 latent variables, wavenumber range: 4000-600 cm\(^{-1}\).

Figure 5. PCA clustering results for bacterial strains: K. pneumoniae Revco 41 (■); E. coli ATCC 25922 (P); E. coli B E4a (O); E. coli K12-25 (●); E. sakazakii ATCC 12868 (X); E. sakazakii ATCC 29004 (#); E. cloacae Rev 1343 (√); E. cloacae Rev 1210 (*); Parameters: smoothing: 6 cm\(^{-1}\), 2\(^{nd}\) derivative transformation: 10 cm\(^{-1}\), 5 latent variables, wavenumber range: 4000-600 cm\(^{-1}\).

Figure 6. PCA clustering results for eight E. sakazakii strains (ATCC 29004 (x), ATCC 12868 (▲), FSM 292 (√), FSM 287 (*), 3 (+), 55 (■), 2.39 (●), and 2.68 (♦)). Parameters: smoothing: 6 cm\(^{-1}\), 2\(^{nd}\) derivative transformation: 10 cm\(^{-1}\), 5 latent variables, wavenumber range: 1200-900 cm\(^{-1}\).

Figure 7. SIMCA classification of E. sakazakii 3 (circle) and E. sakazakii 2.39 (triangle). Parameters: smoothing: 6 cm\(^{-1}\), 2\(^{nd}\) derivative transformation: 10 cm\(^{-1}\), 5 latent variables.