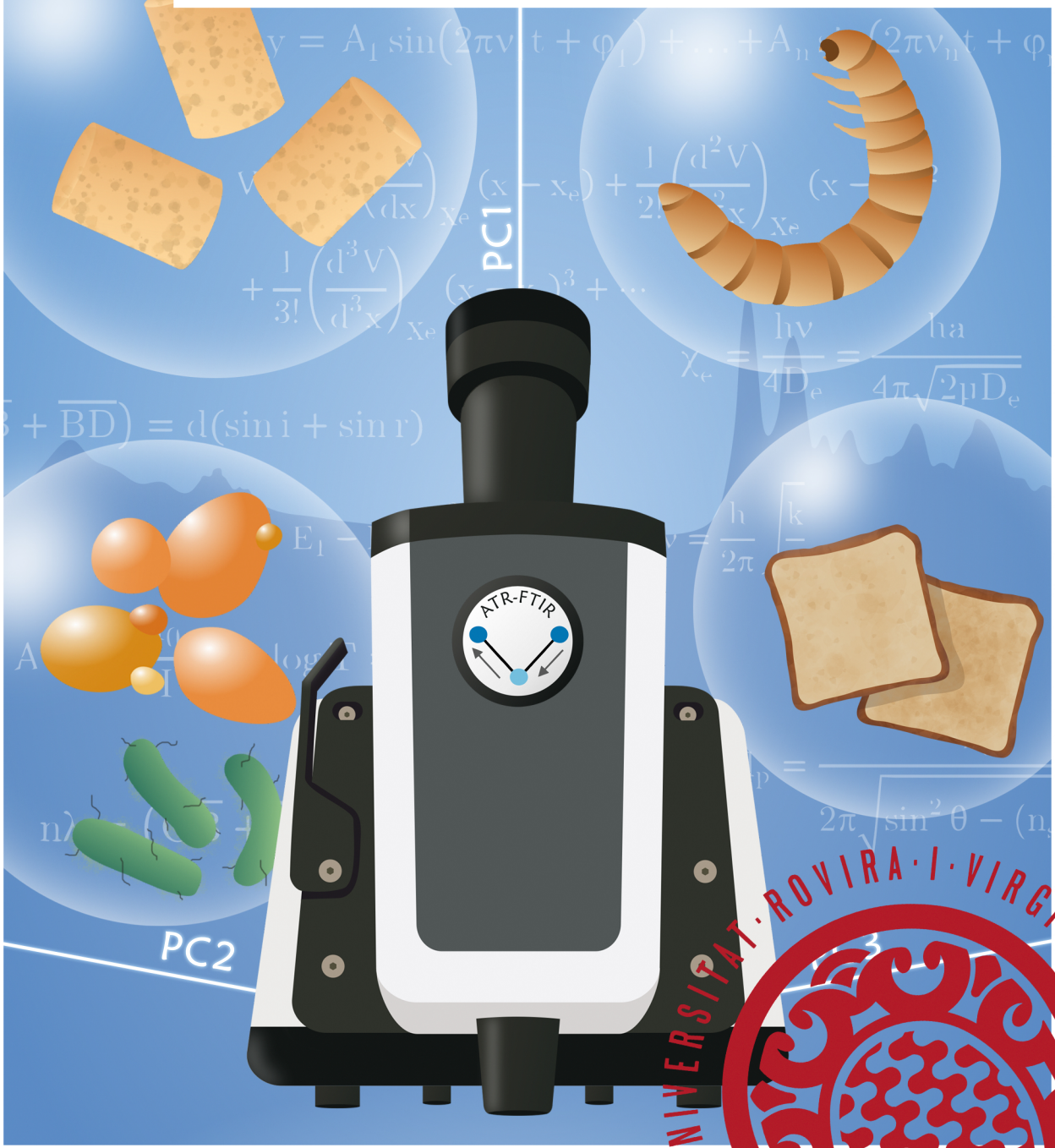
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS

Jorge Mellado Carretero

Evaluation of portable vibrational spectroscopy equipment combined with multivariate analysis for food applications

JORGE MELLADO CARRETERO

DOCTORAL THESIS | UNIVERSITAT ROVIRA I VIRGILI



UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

Jorge Mellado Carretero

Evaluation of portable vibrational spectroscopy equipment combined with multivariate analysis for food applications

Doctoral Thesis

Supervised by
Dr. Sílvia de Lamo Castellví

Co-supervised by
Dr. Diego García Gonzalo

UNIVERSITAT ROVIRA i VIRGILI

Department of Chemical Engineering
Food Innovation and Engineering (FoodIE)

Tarragona 2022

UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

UNIVERSITAT ROVIRA i VIRGILI

Departament d'Enginyeria Química,
ETSEQ Campus Sescelades
Avinguda dels Països Catalans 26, 43007, Tarragona, Spain

I STATE that the present study, entitled "Evaluation of portable vibrational spectroscopy equipment combined with multivariate analysis for food applications", presented by Jorge Mellado Carretero for the award of the degree of Doctor, has been carried out under our supervision at the *Departament d'Enginyeria Química* of Universitat Rovira i Virgili and that it fulfils all the requirements to obtain the doctoral degree.

Tarragona, April 1, 2022

Doctoral Thesis Supervisors

Firmado por SILVIA DE LAMO CASTELLVI - DNI 39895162Y el día 04/04/2022 con un certificado emitido por EC-Ciutadania

GARCIA GONZALO DIEGO - DNI 17750507G

Firmado digitalmente por GARCIA GONZALO DIEGO - DNI 17750507G
Nombre de reconocimiento (DN): c=ES, o=UNIVERSIDAD DE ZARAGOZA, ou=CERTIFICADO ELECTRONICO DE EMPLEADO PUBLICO, serialNumber=IDCES-17750507G, sn=GARCIA GONZALO, givenName=DIEGO, cn=GARCIA GONZALO DIEGO - DNI 17750507G
Fecha: 2022.03.31 23:12:47 +02'00'

Dr. Sílvia de Lamo Castellví

Dr. Diego García Gonzalo

UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

AGRADECIMIENTOS

Me gustaría agradecer enormemente toda la ayuda y supervisión que mi directora de tesis, la Dra. Sílvia de Lamo, me ha prestado durante el transcurso de este tortuoso, pero bonito viaje, que ha sido la obtención del título de doctorado. Gracias por hacerme ver cuál es mi vocación y a qué quiero dedicar el resto de mi vida.

También querría agradecer al Dr. Diego García, mi codirector de tesis, por compartir sus conocimientos de forma desinteresada y por haber estado ahí siempre cuando lo necesitaba.

Muchísimas gracias a todos los compañeros y amigos que han compartido laboratorio conmigo. Al Dr. Wael Kaade, con el que he tenido muchos desencuentros, pero aprecio sobremanera. A la Dra. Junjing Wang, con la que he compartido innumerables abrazos y risas. A los dos, espero visitaros pronto. Gracias a Aurélie, Carmen y Jitesh por toda la ayuda y comprensión que me habéis brindado, soy consciente que siempre que os he necesitado me habéis dado fuerzas para llevar a cabo mis objetivos. También quiero dar las gracias a Halima, Morane, Tamara, Yasmine y Madushika por la amabilidad y cariño que me han demostrado durante mi estancia en el grupo FoodIE. Agradecer a la Dras. Carme Güell y Montse Ferrando esos pequeños momentos que hemos compartido y el compañerismo que han tenido en momentos en que lo he necesitado.

Por otro lado, quiero dar las gracias al Dr. Daniel Berdejo, a Elisa Pagán y a Natalia Merino por la cálida bienvenida que me dieron en su laboratorio y amenizar mi estancia en la Universidad de Zaragoza.

Gracias a Josep Maria, Ana y Ángel por todo el soporte técnico que me habéis dado durante el transcurso de la tesis.

Agradecer a Claudia, Shreya, Xiaoyan y Siyu todos esos momentos que hemos vivido juntos en el grupo de investigación del Dr. Rodriguez-Saona, al que agradezco también el haberme dado la oportunidad de incorporarme en su laboratorio. Un “gracias muy grande” para Travis, Michael y Edmon, mis amigos de Estados Unidos, y que me apoyaron muchísimo en momentos muy difíciles. ¡Espero volver pronto!

Quiero agradecer a mi familia, especialmente a mis padres, el estar siempre ahí dándome amor y compañía a pesar de mi temperamento y mis bromas pesadas.

Finalmente, querría agradecer a una persona que ha dado tanto por mí sin pedir nunca nada a cambio. Nerea, me alegro de haberte conocido y guardo en mi corazón todos esos momentos en que me has apoyado y ayudado. Estemos lejos o cerca, siempre querré ser tu amigo.

CONTENTS

LIST OF ABBREVIATIONS

ATR	Attenuated total reflectance
CVF	Circular variable filter
DTGS	Deuterated triglycine sulphate
EO	Essential oil
FIR	Far-infrared
FTIR	Fourier transform mid-infrared
HCA	Hierarchical cluster analysis
ICD	Interclass distance
IR	Infrared
LVF	Linear variable filter
MCT	Mercury cadmium telluride
MEMS	Microelectromechanical systems
MIR	Mid-infrared
MLR	Multivariate linear regression
MRP	Maillard reaction product
MSC	Multiplicative scatter correction
NIR	Near-infrared
PC	Principal component
PCA	Principal component analysis
PCR	Principal component regression
PLS-DA	Partial least squares – discriminant analysis
PLSR	Partial least squares regression
PRESS	Predicted residual error sum of squares
RMSECV	Root mean square error of cross-validation
RMSEP	Root mean square error of prediction
SECV	Standard error of cross-validation
SEP	Standard error of prediction
SIMCA	Soft independent modelling of class analogy
SVD	Singular value decomposition
SVM	Support vector machine
UV	Ultraviolet

ABSTRACT

During these last decades vibrational spectroscopy has gathered attention in the food science field since it provides ways of analysing and characterizing a wide range of food commodities in a fast and simple manner. Moreover, its cost-effectiveness, reproducibility, and sensitivity have also been proved. Recently, vibrational spectroscopy has experimented a breakthrough regarding its miniaturisation, achieving in certain equipment sizes as small as the palm of a hand. Miniaturized spectroscopic devices open new possibilities towards the applications of vibrational spectroscopy for the in situ (e.g., in the field, industry, etc.) analysis of several food quality control parameters without altering the sample.

The objective of this doctoral thesis was to evaluate the feasibility of portable and handheld Raman, mid-infrared (MIR), and near-infrared (NIR) spectrometers combined with powerful chemometric algorithms for food applications including the study of chemical and physical parameters and species typing.

Firstly, the capability of a portable attenuated total reflectance Fourier transform mid-infrared (ATR-FTIR) spectrometer was tested for monitoring the Maillard reaction in protein-polysaccharide (sodium caseinate and gum Arabic or sodium carboxymethyl cellulose) model systems as an alternative to conventional methods that rely on the use of toxic and harmful reagents. Protein-polysaccharide mixtures obtained by spray drying were subjected to different reaction times to study the process of conjugate formation. The outcomes obtained from the chemometric models built up based on soft independent modelling of class analogy (SIMCA) were in accordance with the data collected from the standard methods of reference, such as the measurement of the conjugate stability effect through turbidity and potential measurements. Moreover, ATR-FTIR spectroscopy provided information about the components formed during the different reaction times tested. Therefore, the results of this study showed the potential of a portable ATR-FTIR spectrometer combined with multivariate analysis as a novel approach to monitor the Maillard reaction.

Edible insects have been proposed as an alternative source of animal and human protein since they present nutritional, economic, and environmental benefits. However, the consumer's rejection towards eating

insects, mainly occurring in Western countries, has led to their further processing into more palatable forms (e.g., powders). Because of that, authentication methods are required for assessing insect's origin and species to prevent frauds and adulterations. In this second study the same portable ATR-FTIR spectrometer was used to discriminate different edible insect powders by species (*Tenebrio molitor*, *Alphitobius diaperinus*, *Gryllodes sigillatus*, *Acheta domesticus*, and *Locusta migratoria*) and origin. A SIMCA model built up with IR data showed an outstanding classification performance and provided chemical information about the components that aided for species identification, which were linked to lipids and chitin.

The feasibility of other miniaturized vibrational spectroscopy devices was further investigated for determining the presence of surface treatments applied to different varieties of still and sparkling wine cork stoppers. NIR spectroscopy seems promising for the food sector due to the high degree of miniaturization achieved during these last years, facilitating its implementation for online process monitoring. Therefore, this research mainly focused on the use of a handheld NIR spectrometer for surface treatment authentication of cork closures. It was proven the potential of our approach for discriminating the treated corks from those that were not, with high sensitivity and specificity values achieved by the SIMCA model created ($>$ 90%). The NIR bands responsible of the discrimination between treated and untreated stoppers were attributed to the components of the surface treatments applied (i.e., silicone). Moreover, a partial least squares regression (PLSR) algorithm was developed for predicting the extraction force values in cork stoppers made for still wines, with high predicting ability. A handheld Raman spectrometer was also tested for one variety of stoppers, which confirmed the presence of the same components detected by NIR spectroscopy.

Portable vibrational spectroscopy was also further evaluated to wider the range of applications regarding the food microbiology field. Our portable ATR-FTIR spectrometer was used in two different microbiological applications. In the first application, a supervised chemometric model based on principal component analysis (PCA) was developed to differentiate *Escherichia coli* mutants resistant to essential oil constituents (carvacrol, citral, and limonene oxide); and to evaluate their biochemical differences as compared to their parental strain in order to get new insights about their mechanism of resistance. PCA models were also developed for wine-making yeasts strains for discriminating them; and SIMCA classification models were

created to provide a way for identifying *Saccharomyces cerevisiae* and non-*Saccharomyces* species. PCA models successfully discriminated all strains (*E. coli* and yeasts), while the SIMCA model accurately classified non-*Saccharomyces* and *S. cerevisiae* yeasts.

The results of the present doctoral thesis show the potential of portable vibrational spectroscopy combined with powerful chemometric techniques as a rapid, reproducible, and sensitive tool for its application in the food sector.

RESUMEN

Durante estas últimas décadas, la espectroscopia vibracional ha captado la atención en el campo de la ciencia de los alimentos ya que proporciona maneras sencillas y rápidas de analizar y caracterizar una amplia gama de productos alimentarios. Además, su rentabilidad, reproducibilidad y sensibilidad han sido probadas. Recientemente, la espectroscopia vibracional ha experimentado un gran desarrollo respecto a su miniaturización, alcanzando en algunos equipos tamaños tan pequeños como el de la palma de una mano. Los dispositivos espectroscópicos miniaturizados abren nuevas posibilidades respecto a las aplicaciones in situ de la técnica para el control de la calidad de los alimentos sin alteración de la muestra.

El propósito de esta tesis doctoral ha sido evaluar la viabilidad de espectrómetros portables Raman, de infrarrojo medio e infrarrojo cercano combinados con potentes algoritmos quimiométricos para aplicaciones alimentarias que engloban el estudio de parámetros químicos y físicos y la tipificación de especies.

En primer lugar, se probó la capacidad de un espectrómetro de infrarrojo con transformada de Fourier (ATR-FTIR) portable para realizar el seguimiento de la reacción de Maillard en sistemas modelo de proteína-polisacárido (caseinato de sodio y goma arábiga o carboximetilcelulosa sódica) como alternativa a los métodos convencionales basados en el uso de reactivos tóxicos y dañinos. Las mezclas de proteína-polisacárido obtenidas mediante secado por pulverización se sometieron a diferentes tiempos de reacción para estudiar el proceso de formación de conjugados. Los resultados conseguidos a partir de los modelos construidos basados en soft independent modelling of class analogy (SIMCA) concordaron con los datos obtenidos a partir de métodos de referencia, como el efecto en la estabilidad de las mezclas analizados a través de medidas de turbidez y potencial zeta. Además, la espectroscopia ATR-FTIR proporcionó información sobre los compuestos formados durante la reacción. Por lo tanto, los resultados de este estudio mostraron el potencial que tiene un espectrómetro ATR-FTIR portable combinado con análisis multivariante para el seguimiento de la reacción de Maillard.

Los insectos comestibles han sido propuestos como una fuente alternativa de proteína animal y humana que ofrece beneficios nutricionales, económicos y ambientales. Sin embargo, el rechazo que presentan los consumidores hacia los insectos, y que ocurre principalmente en los países occidentales, ha dado lugar a su procesamiento en formas más apetecibles (por ejemplo, polvos). Por este motivo, se requieren métodos de autentificación para evaluar el origen y la especie utilizadas para evitar fraudes y adulteraciones. En este segundo estudio se utilizó el mismo espectrómetro ATR-FTIR portable para discriminar por especie (*Tenebrio molitor, Alphitobius diaperino, Gryllodes sigillatus, Acheta domesticus* y *Locusta migratoria*) y origen diferentes polvos de insectos. Un modelo SIMCA construido con los datos de IR mostró una clasificación sobresaliente y proporcionó información química sobre los componentes que facilitaron la identificación de las diferentes especies, que se vincularon a lípidos y quitina.

Se investigó también la viabilidad de otros dispositivos de espectroscopia vibracional miniaturizados para determinar la presencia de tratamientos superficiales aplicados a diferentes variedades de tapones de corcho utilizados en vinos tranquilos y espumosos. La espectroscopia de infrarrojo cercano (NIR) se considera prometedora para el sector alimentario debido al alto grado de miniaturización que ha alcanzado durante estos últimos años, lo que facilita su implementación para el análisis del control de calidad en procesos en línea. Por lo tanto, se investigó el uso de un espectrómetro portable NIR para autentificar los tratamientos superficiales aplicados. Se comprobó el potencial que tenía nuestra idea para discriminar los tapones tratados de los que no lo estaban, alcanzando valores altos de sensibilidad y especificidad ($> 90\%$). Las bandas NIR responsables de dicha diferenciación se atribuyeron a los componentes químicos de los tratamientos superficiales aplicados (es decir, silicona). Además, se desarrolló un algoritmo de regresión de mínimos cuadrados parciales para predecir los valores de fuerza de extracción en los tapones de corcho elaborados para vinos tranquilos, que mostró una alta capacidad de predicción. También se utilizó un espectrómetro Raman portable para una de las variedades de tapón analizadas, que confirmó la presencia de los mismos componentes detectados por espectroscopia NIR.

La espectroscopia vibracional portable también se evaluó en el campo de la microbiología de los alimentos. Nuestro espectrómetro ATR-FTIR portable se utilizó en dos aplicaciones microbiológicas distintas. En la primera, se desarrolló un modelo quimiométrico supervisado basado en el análisis de

componentes principales (PCA) para diferenciar mutantes de *Escherichia coli* a compuestos de aceites esenciales (carvacrol, citral y óxido de limoneno); y evaluar sus diferencias bioquímicas en comparación con la cepa parental para obtener nuevos conocimientos acerca de su mecanismo de resistencia. También se desarrollaron modelos de PCA para cepas de levaduras de interés enológico con el objetivo de discriminarlas; y se crearon modelos de clasificación basados en SIMCA para identificar especies *Saccharomyces cerevisiae* y no-*Saccharomyces*. Los modelos PCA discriminaron con éxito todas las cepas (*E. coli* y levaduras), mientras que el modelo SIMCA clasificó con precisión las levaduras no-*Saccharomyces* y *S. cerevisiae*.

Los resultados de la presente tesis doctoral muestran el potencial de la espectroscopia vibracional portátil combinada con potentes técnicas quimiométricas como herramienta rápida, reproducible y sensible para su aplicación en el sector alimentario.

Introduction and objectives

UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

Food is one of the most important aspects in our daily routine. Besides being the energy source that sustains our lives, it constitutes the basis of many social interactions, having a great economic and environmental impact. Therefore, one of the biggest concerns among consumers and food industries is food safety, primarily caused by the emergence of food-related incidents that pose a public health threat.

1.1 A brief tour through food safety history

Since ancient times, humanity has been recognizing and avoiding those foods that were toxic or not ready to eat. Our ancestors learned new ways to preserve foods by trial and error (e.g., sun drying, salting, and fermentation), which made them safer to consume. It was not until the industrial revolution that the first food laws such as the Act of 1860 for "Preventing the Adulteration of Food and Drink" or *Codex Alimentarius Austriacus* arose, mainly because of the rapid urbanization (i.e., the need for a massive food production) and the poor hygiene conditions. Nonetheless, both laws were not enforceable and thus did not work. The first major consumer's protection law to be enforced was the Food, Drug and Cosmetic Act in 1938 by the United States (US) Food and Drug Administration (FDA)[1].

The regulatory framework for food safety based on the hazard analysis and critical control points (HACCP) and the good practices (GPs) was developed and established in the US during the 1960s and 1970s as a voluntary program. In 1971, a scandal splattered the Pillsbury Company regarding the contamination of their wheat flour with shattered glass. Dr. Howard Bauman, the Pillsbury's Director of Research, decided then to apply to all Pillsbury's production a system created during the 1960s for the safety assurance of food supplied to NASA astronauts, which was based on the assessment of critical control points (CCPs) and good manufacturing practices (GMPs)[2].

After the occurrence in US of several outbreaks of *Clostridium botulinum* in canned foods, Dr. Baumann gave a training program to the FDA, whose material included the first references to the use of the HACCP system. Two decades after the HACCP principles were turned into a harmonized and globalized system when the *Codex Alimentarius* Commission (CAC) published its final report on HACCP in 1997[2,3]. That report included the description of the seven current HACCP principles[1,4]:

i. Conduct a hazard analysis

ii. Identify the CCPs and hazards
iii. Establish critical limits for each preventive measure
iv. Establish a system to monitor the CCPs
v. Establish the corrective action to be taken when monitoring indicates a deviation in CCPs
vi. Establish procedures for verification to confirm the validity of the HACCP system
vii. Establish recordkeeping procedures of documentation concerning these principles and their application

The CAC, which was created in 1963, is the coordination organism of the Food Standards Programme from the joint World Health Organization (WHO) and the United Nations Food and Agricultural Organization (FAO). Besides CAC, other entities and regulatory frameworks coexist and are different from region to region. In Europe, the regulation is defined by the European Commission (EC) Directives and set by the European Food Safety Authority (EFSA). On the other hand, US legislation is based on "Acts" produced by the FDA. Both entities recommended the application of HACCP, GPs and International Standards Organization (ISO) standards[5].

ISO was created in 1946 as an independent and non-governmental organization in order to coordinate and unify the industrial standards, acting as a bridge between the public and private sectors. ISO has published more than 19500 standards, of which ISO 22000 family includes standards dealing with the management of food safety[1]. ISO 22000 standards were published in 2005 and combine the HACCP principles with GPs. Local national HACCP standards are intended to be replaced with ISO 22000. GPs [also known as prerequisite programs (PrPs)] are in turn procedures that describe conditions that help to run operations more smoothly for keeping a good food safety assurance program[4]. Good hygiene practices (GHP), good agricultural practices, and good manufacturing practices are examples of GPs and the most implemented in the food industry[1].

1.2 Food safety, quality, and fraud

All the regulations previously discussed lead us to the definition of some important concepts used in this field. According to the CAC guidelines:

- Food safety: "The assurance that food will not cause harm to the consumer when prepared and/or eaten according to its intended use."
- Food quality: "Quality includes all the attributes that influence a product value to the consumer. This includes negative attributes such as spoilage, contamination with filth, discoloration, and off-odours and positive attributes such as the origin, colour, flavour, texture, and processing method of the food."
- Quality Assurance: "All those planned and systematic actions necessary to provide adequate confidence that a product or service will satisfy given requirements for quality."
- Food Fraud: "The deliberate and intentional substitution, addition, tampering, or misrepresentation of food, food ingredients, or food packaging; or false or misleading statements made about a product, for economic gain that could impact consumer health." TABLE 1.1 includes the different types of food fraud[6]:

TABLE 1.1 Food fraud types. Adapted from Spink (2019)[6].

Fraud type	Definition
Adulteration	One of the ingredients is fraudulent
Tamper	Legitimate product and packaging are used in a fraudulent way
Overrun	The food product is made in excess of production agreements
Theft	Legitimate product is stolen and passed off as legitimately procured
Diversion	The sale or distribution of legitimate products outside of intended markets
Simulation	Illegitimate product is designed to look like but not exactly copy the legitimate product
Counterfeit	All aspects of the fraudulent product and packaging are fully replicated

It can be seen from HACCP principles and the previous definitions that a traceability system is needed in the safety assurance of foods to prove their identity and quality. Such system should focus on the following key aspects: quality and safety (food standards, adulterants, and labelling), geographical indication (species and processing conditions) and manufacturing systems (organic, conventional, or traditional procedures)[5,7]. Food safety and quality monitoring is mandatory in the European Union as well as in other countries. Moreover, the increase in the complexity of adulteration is demanding an improvement in the authentication and analytical protocols to ensure the correct assessment of the fraudulent components[8].

1.3 Food quality assessment

1.3.1 Quality factors

Food quality assessment is based on the acquisition of quantitative information from internal and external quality factors of the product, such as size, shape, colour, presence of defects, flavour, composition, and texture[9,10]. Therefore, the appropriate technology should be selected to measure the factor of interest. Technologies for the determination of the quality of foods can be based on mechanical, electromagnetic, and electrical energy, among others. TABLE 1.2 summarizes some of these techniques and the corresponding parameter measured[9]. As it will be described in the next section, the chemical composition of food is one of the most studied quality factors for the evaluation of food quality and safety.

TABLE 1.2 Measuring parameters regarding energy sources for food quality evaluation. Adapted from Cho (2011)

Energy Source	Method	Measuring Parameter
Mechanical energy	Vibration	Firmness, viscoelasticity
	Impact	Firmness, internal cavity
	Ultrasonic	Elasticity, internal texture
	Creep/stress relaxation	Viscoelasticity
Electromagnetic energy	X-ray	Internal texture
	UV spectroscopy	Chemical composition
	Visible spectroscopy	Chemical composition, colour
	Near-infrared (NIR) spectroscopy	Chemical composition, internal texture
	Visible imaging technology	Colour, shape dimension, surface texture
	NIR imaging technology	Temperature, surface texture
	Laser scattering	Particle size, texture
	Nuclear magnetic resonance	Chemical composition
Electrical energy	Impedance	Moisture, internal texture
	Potentiometric	Freshness
Others	Gas sensor	Volatile component
	Biosensor	Biomolecule, cell
	Chromatography	Chemical composition
	Microbiological techniques	Contamination

1.3.2 Fingerprinting: study of the chemical composition of food

The chemical composition of food is a good indicator of its authenticity, quality, and origin. There are targeted and untargeted approaches for the analysis of food product's authenticity and safety. Targeted methods focus on the analysis of a single or a family of compounds that constitute specific

markers. On the other hand, the untargeted strategy aims to explore the chemical composition of foods to identify unique patterns that might change due to environmental alterations, different storage conditions, or adulteration practices[5,8,11,12]. The latter approach is named fingerprinting since the uniqueness of the signals (i.e., spectrum, chromatogram, or image) reminds the one from a human's fingerprint.

Several fingerprinting methods have been widely used in food science. Among them, gas chromatography (GC), high-performance liquid chromatography (LC), and mass spectrometry (MS) are considered the gold standards or official methods for assessing food quality and safety. However, such techniques demand complex protocols for sample preparation and hence are time-consuming. Besides that, they are also expensive and sometimes involve the use of toxic reagents[13–15].

For this reason, it has been considered the use of non-destructive techniques in the food sector since they allow measuring repeated times different parameters on the same sample and thus, reduce waste, analysis time, and economic losses. In the recent years spectroscopic methods have gathered attention for food fingerprint analysis because of their robustness, speed, low-cost, and lack of sample preparation[16,17]. Throughout this introduction, the theory and principles of vibrational spectroscopy will be discussed and the advantages of each of the techniques explained will be commented in food safety and quality assessment.

1.4 Vibrational spectroscopy: Basic theory and principles

Spectroscopy can be divided into two main groups: photonic spectroscopy, which deals with the interaction of light with matter, and particle spectroscopy (i.e., mass spectrometry and derived methods)[18,19]. Regarding photonic spectroscopy, it can be divided depending on the interacting species (atomic or molecular), the type of light-matter interaction measured (e.g., absorption, emission or, scattering), and the region of the electromagnetic spectrum used[18].

1.4.1 Light properties

Two ways to explain how light interacts with matter were in competition until the twentieth century, the corpuscular and undulatory theories. The development of quantum physics implied that light should be

thought of particles of a certain energy that travel through space with wavelike properties: in other words, light has a dual nature[18,19]. Some light properties can be successfully explained using the undulatory theory, such as refraction, diffraction, and interference phenomena, while emission and absorption are best understood in terms of its particulate nature[18].

1.4.1.1 Wavelike properties

The wavelike properties of light are explained by considering it as a plane polarized wave of perpendicularly oriented magnetic and electric oscillating fields that propagate at a certain direction (Figure 1.1)[18]. Considering only the electrical component of light, this can be represented as a sinusoidal function:

$$y = A\sin(\omega t + \varphi) \tag{1.1}$$

where A is the amplitude or maximum value for y (N/C) at a certain time t (s), ω is the angular velocity (rad/s), and φ is the phase (rad). It is possible to relate the angular velocity with the frequency ν, which is the number of oscillations per second (s^{-1}):

$$\omega = 2\pi\nu \tag{1.2}$$

Combining Equations (1.1) and (1.2) gives:

$$y = A\sin(2\pi\nu t + \varphi) \tag{1.3}$$

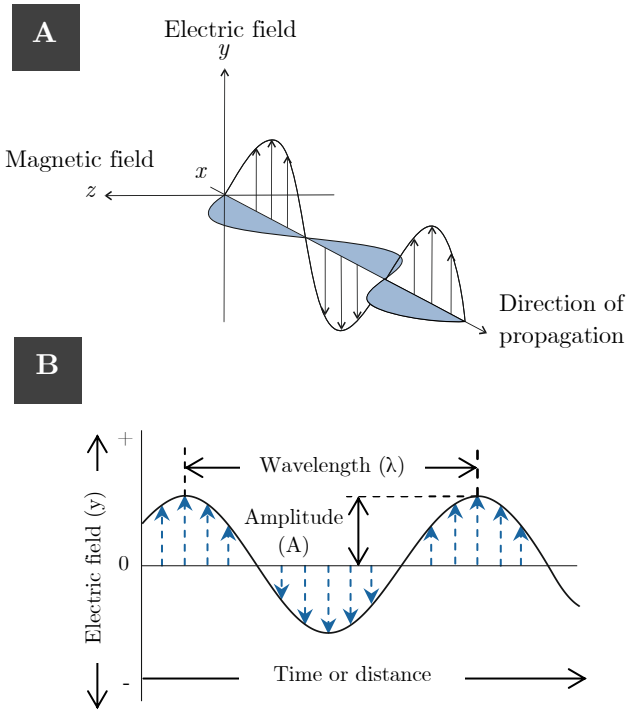


FIGURE 1.1 Plane polarized wave oscillating along the x axis (A) and its electric field component (B). Adapted from Skoog, Holler and Crouch (2018)[20].

A wave can also be described by its wavelength λ, which is the distance covered by light in a cycle. Wavelength can be then expressed in terms of its frequency:

$$\lambda = \frac{c}{\nu} \tag{1.4}$$

where c is the wave velocity of propagation in meters per second (m/s). Wavelengths are expressed in different units in the spectroscopy field depending on the spectral region analysed (e.g., cm, mm, µm, nm, and Å are usually used). Another unit widely used in mid-infrared (MIR) spectroscopy is the wavenumber $\bar{\nu}$, and is the reciprocal of the wavelength in cm^{-1}[19,20]. Wavelength and frequency can be used to define the different regions of the electromagnetic spectrum (FIGURE 1.2).

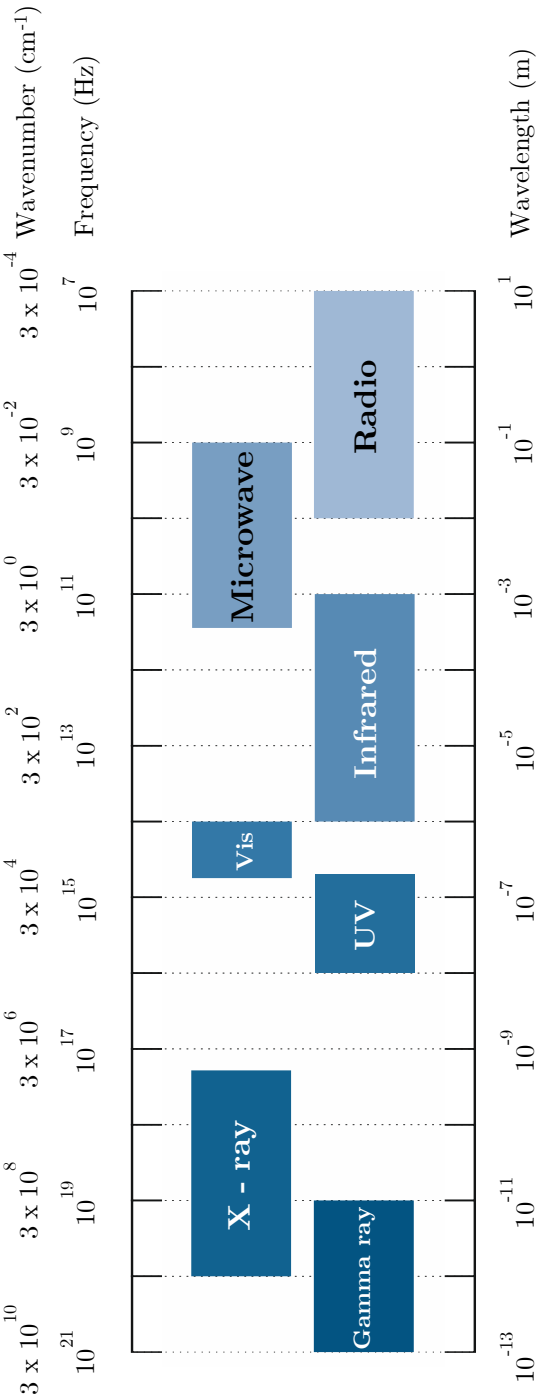


FIGURE 1.2 Regions of the electromagnetic spectrum.

When two or more waves encounter one another, a disturbance occurs and a resultant wave is produced, whose amplitude is the sum of the amplitudes of all the individual waves at the point of intersection. The law that explains this behaviour is named the principle of superposition. The following equation applies for n electromagnetic waves:

$$y = A_1 \sin\left(2\pi\nu_1 t + \varphi_1\right) + \ldots + A_n \sin\left(2\pi\nu_n t + \varphi_n\right) \tag{1.5}$$

Maximum constructive interference takes place when two waves are completely in phase (i.e., $\varphi_2 - \varphi_1 = n \cdot 360$), while maximum destructive interference occurs when they are out of phase by 180° (i.e., $\varphi_2 - \varphi_1 = 180 + n \cdot 360$; Figure 1.3). Vibrational spectroscopy takes advantage of interference phenomena to provide measurements in a fast and reliable fashion[18,20].

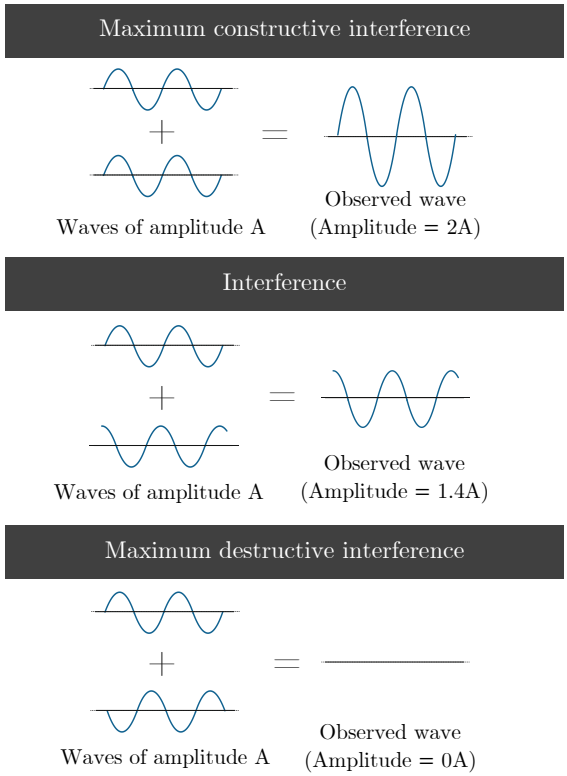


Figure 1.3 Interference patterns obtained with two identical waves. Adapted from Penner (2017)[18].

Other light phenomena can be explained by means of the wavelike nature of the light. Diffraction occurs when a parallel beam of radiation bents as it crosses a barrier with a small slit. When the slit width is shorter than the wavelength of the beam, a new source of circular waves is generated. If a screen is placed in front of the slit, a pattern can be observed, which is the result of the interference events generated by such waves. Thomas Young studied the diffraction phenomenon (in the so-called double slit experiment in 1800) and concluded that the difference in the light pathlength from both slits corresponds to the wavelength (or the wavelength multiplied by an integer) at the maximum constructive interference points:

$$n\lambda = \overline{CF} = \overline{BC}\sin\theta = \frac{\overline{BC}\ \overline{DE}}{\overline{OD}} = \frac{\overline{BC}\ \overline{DE}}{\overline{OE}} \tag{1.6}$$

Equation (1.6) implies that the linear displacement $\overline{DE}$ of the diffracted beam along the screen depends on the distance $\overline{OE}$ from the slits and the spacing $\overline{BC}$ between the slits (Figure 1.4)[20].

1.4.1.2 Quantum properties

Some properties have been discussed based on the undulatory properties of the light (e.g., diffraction and interference). However, two of the most important properties involved in spectroscopy are the absorption and emission of light. After the observation of the photoelectric effect (i.e., generation of a measurable current as a consequence of irradiation of a surface with light), Albert Einstein concluded that it is not possible to explain those results with the classical wave model, and a quantum model is needed instead. Then, Einstein proposed a relationship between the energy (E) of a photon and its frequency (ν):

$$E = h\nu \tag{1.7}$$

where h is the Planck's constant ($6.6254 \cdot 10^{-34}$ joule per second). Equation (1.7) makes necessary to assume that the energy is not uniformly distributed but rather concentrated in bundles or "quanta" of energy[18,20,21].

Since the energy of matter is quantized, atoms and molecules may exist in a series of discrete energy states. Most of the molecules and atoms are in the ground or lowest energy state under normal conditions and they can be promoted to a higher energy state once they gain a certain amount of energy, which is determined by the set of available energy states. The latter together with the fact that energy transitions are characteristic of a species, makes absorption and emission phenomena a good candidate as a fingerprinting methodology[18].

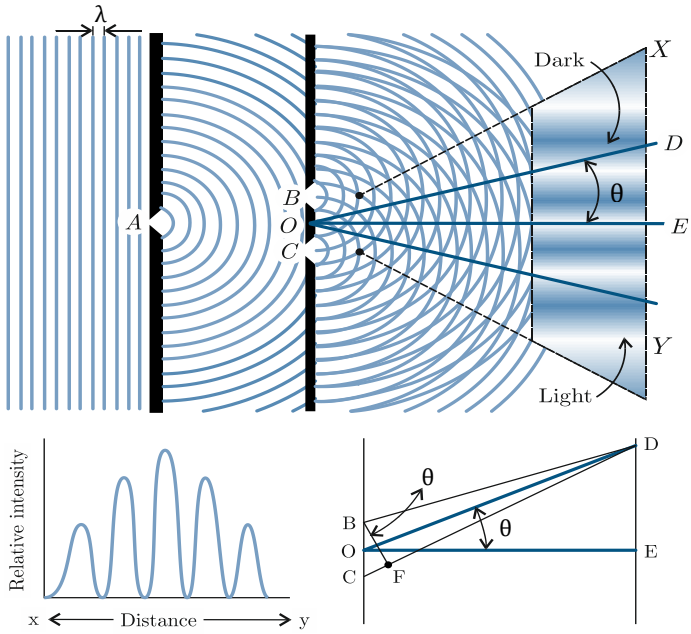


FIGURE 1.4 Diffraction pattern by a monochromatic light going through slits. Adapted from Skoog, Holler, and Crouch (2018)[22].

1.4.2 Energy states of matter and energy transitions

Molecules exhibit other types of transitions since the atoms they are made of can vibrate in many ways. All these vibrations are quantized as well. Energy differences between different vibrational states are called vibrational transitions. Moreover, the potential energy of a molecule can also be described in terms of the energy associated with the rotational motions. Therefore, the energy of an atom is dependent on its electronic energy levels, while the energy of a molecule is associated with its electronic, vibrational, and rotational states

(Figure 1.5). Since the energy spacings between states are different in electronic, vibrational, and rotational transitions, different regions of the electromagnetic spectrum can be used to induce them, then different types of spectroscopies exist, which are depicted in Table 1.3[18].

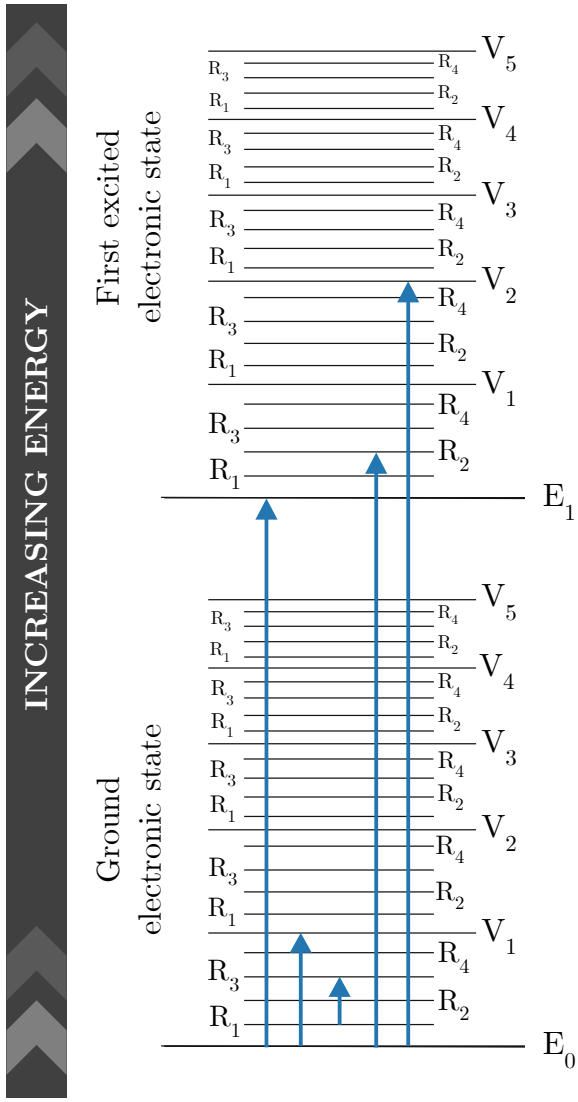


Figure 1.5 Molecular energy-level diagram comprising electronic, vibrational, and rotational transitions. Adapted from Penner (2017)[18].

TABLE 1.3 Spectroscopic methods and associated transitions. Adapted from Penner (2017)[18].

Wavelength region	Wavelength limits	Type of spectroscopy	Usual wavelength range	Types of transitions in chemical systems with similar energies
Gamma rays	0.01-1 Å	Emission	< 0.1 Å	Nuclear proton/neutron arrangements
X-rays	0.1-10 nm	Absorption, emission, fluorescence, and diffraction	0.1-100 Å	Inner-shell electrons
Ultraviolet (UV)	10-380 nm	Absorption, emission, and fluorescence	180-380 nm	Outer-shell electrons in atoms, bonding electrons in molecules
Visible (Vis)	380-750 nm	Absorption, emission, and fluorescence	380-750 nm	Same as ultraviolet
Infrared (IR)	0.075-1000 µm	Absorption	0.78-300 µm	Vibrational position of atoms in molecular bonds
Microwave	0.1-100 cm	Absorption	0.75-3.75 mm	Rotational position in molecules
		Electron spin resonance	3 cm	Orientation of unpaired electrons in an applied magnetic field
Radio wave	1-1000 m	Nuclear magnetic resonance	0.6-10 m	Orientation of nuclei in an applied magnetic field

1.4.3 Vibration of a diatomic molecule

1.4.3.1 Harmonic oscillator

If we consider a diatomic molecule as two masses connected by a spring that vibrate as a harmonic oscillator, the quantum treatment (i.e., solutions for the eigenvalues of the Schrodinger's equation; $H\psi = E\psi$) of such vibrations gives the following equation for the potential energy (E):

$$\mathrm{E} = \left(v + \frac{1}{2}\right)\frac{\mathrm{h}}{2\pi}\sqrt{\frac{\mathrm{k}}{\mu}} = \left(v + \frac{1}{2}\right)\mathrm{h}\nu \tag{1.8}$$

where k is the force constant of the chemical bond (N/m), µ is the reduced mass of each of the atoms of the molecule (kg), h is the Planck's constant, v is the vibrational quantum number (which can only take positive integer values, including zero), and ν is the vibrational frequency of the bond, respectively. Note that the frequency of the bond is related to the bond strength and the reduced mass of the atoms involved in the vibration. The energy difference between the ground state E_0 and the first excited state E_1 is:

$$\Delta E = E_1 - E_0 = \frac{3}{2}h\nu - \frac{1}{2}h\nu = h\nu = \frac{h}{2\pi}\sqrt{\frac{k}{\mu}} \quad (1.9)$$

Therefore, the frequency of the radiation being able to induce the promotion of a molecule from the ground to the excited state must be the same as the vibrational frequency of the bond. This kind of transition is called the fundamental.

1.4.3.2 Anharmonicity

Several deviations from the harmonic behaviour are observed due to the coulombic repulsions between the nuclei of the two atoms when approaching one another and their dissociation at a high interatomic distance. One of the deviations is the existence of transitions from the ground state to states higher than the first (e.g., $0 \rightarrow 2$). Such transitions are called overtones[18,19,21–23], which are the basis of near-infrared (NIR) spectroscopy.

Instead of modelling bond vibrations as a harmonic oscillator, another approximation was suggested by Morse in 1929, which considered both coulombic repulsions and bond dissociation's energy (Morse and harmonic oscillator energy potential functions are depicted in Figure 1.6):

$$V(x) = D_e[1 - \exp\{-a(x - x_e\}]^2 \quad (1.10)$$

where x is de internuclear distance, D_e is the dissociation energy of the bond, and a is a constant. Morse function can be expanded by a Taylor series into the following polynomial:

$$V(x) = V(x_e) + \left(\frac{dV}{dx}\right)_{Xe}(x - x_e) + \frac{1}{2!}\left(\frac{d^2V}{d^2x}\right)_{Xe}(x - x_e)^2 + \frac{1}{3!}\left(\frac{d^3V}{d^3x}\right)_{Xe}(x - x_e)^3 + \cdots \quad (1.11)$$

Considering $\left(\frac{d^2V}{d^2x}\right)_{Xe} = k$ and ignoring the terms with orders higher than two, the following equation holds:

$$V(x) = \frac{1}{2}k(x - x_e)^2 \tag{1.12}$$

Therefore, Morse's equation represents a function close to the harmonic oscillator approximation in the region close to the equilibrium nuclear distance x_e. The high order terms in Equation (1.11), $(x - x_e)^3$, $(x - x_e)^4$ and so on, are called the anharmonic terms. When the Schrödinger equation is solved using the Taylor series expansion considering up to the cubic term we get:

$$E = \left(v + \frac{1}{2}\right)h\nu - \left(v + \frac{1}{2}\right)^2 h\nu\chi_e \tag{1.13}$$

where χ_e is the anharmonicity constant, which holds the following relationship:

$$\chi_e = \frac{h\nu}{4D_e} = \frac{ha}{4\pi\sqrt{2\mu D_e}} \tag{1.14}$$

Molecules with lighter atoms (μ) and lower dissociation energies (D_e) have high anharmonicity constants and, consequently, exhibit overtones at lower energies[23]. This is one of the reasons why NIR spectroscopy exhibits only overtones of vibrations in which there is a hydrogen atom involved. Moreover, spacings at higher energy levels become narrower, thus the overtones do not exactly match with the fundamental frequency multiplied by an integer.

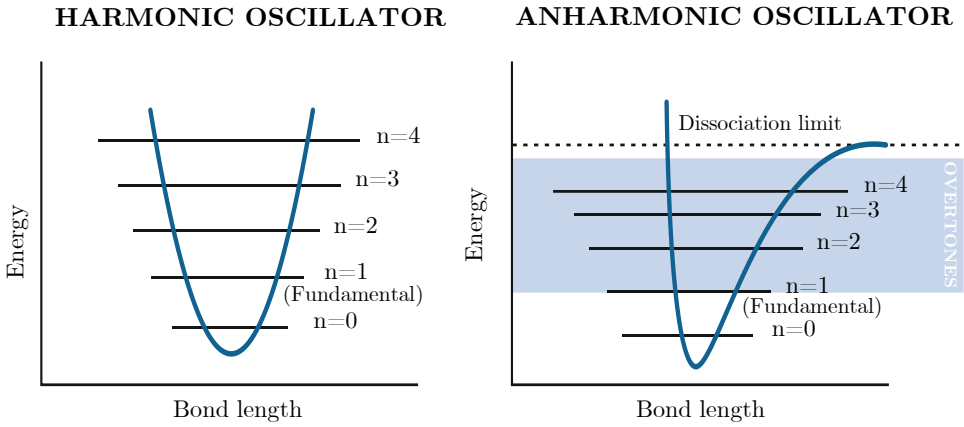


Figure 1.6 Potential energy curves of harmonic (A) and anharmonic (B) oscillator models. Adapted from Rodriguez-Saona, Ayvaz, and Wehling (2017)[21].

1.4.4 Vibrational modes of molecules

Considering a molecule with N atoms, three coordinates are necessary to locate each of the atoms the molecule is made of, then it has 3N degrees of freedom. Such degrees of freedom can be divided into translational, vibrational, and rotational motions. Three degrees of freedom (i.e., three coordinates) are required to describe the translational motion of the molecule's centre of gravity. Another three are taken to explain the rotational motions of a molecule. Hence, 3N-6 vibrations or normal modes are possible for a polyatomic molecule, except linear ones that present 3N-5 vibrations since they cannot rotate around the bond axis[19,22–24]. It is worth mentioning that, for nonlinear molecules, the symmetry plays an important role in the types of vibrations observed. Thereby, a water molecule, which has an axis of symmetry, presents symmetric and antisymmetric vibrations[19]. Vibrations can be separated into two categories: stretching and bending. Stretching describes the change of interatomic distance along the bond axis, while bending involves a change in the angle between two bonds. All types of vibrations are displayed in Figure 1.7[22,25].

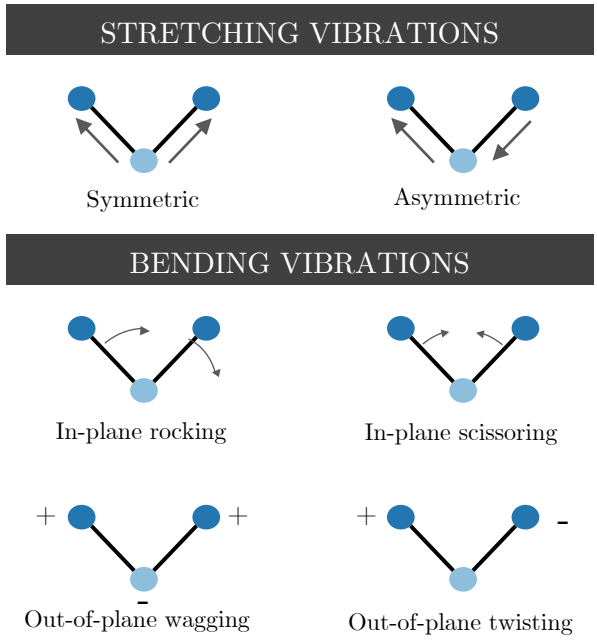


Figure 1.7 Types of molecular vibrations. Note that + and - indicate motions towards the reader and away from the reader, respectively. Adapted from Skoog Holler, and Crouch (2018)[22].

1.4.5 Light interaction with matter

An incident beam of light of intensity I_0 can undergo several phenomena upon interacting with matter (e.g., food), such as specular reflection (I_R), diffusive reflection or scattering (I_S), absorption (I_A), and transmission (I_T). When the light interacts with food, about 4% of the incident light is reflected by its surface, while the rest is transmitted through the matrix. Some of the transmitted light can be either absorbed or scattered at different angles. By computing an energy balance of the processes illustrated in Figure 1.8, the following equation is obtained:

$$I_0 = I_R + I_S + I_A + I_T \tag{1.15}$$

I_A is the only parameter that gives information about the chemical composition of a sample. Even though this value cannot be measured directly, it can be assessed by measuring the rest of parameters. Most commercially available spectrometers use only one detector to measure either I_0 and one of the other parameters (usually I_S or I_T), then I_A can be calculated by bringing the rest to zero or close to zero through sample preparation[10,16].

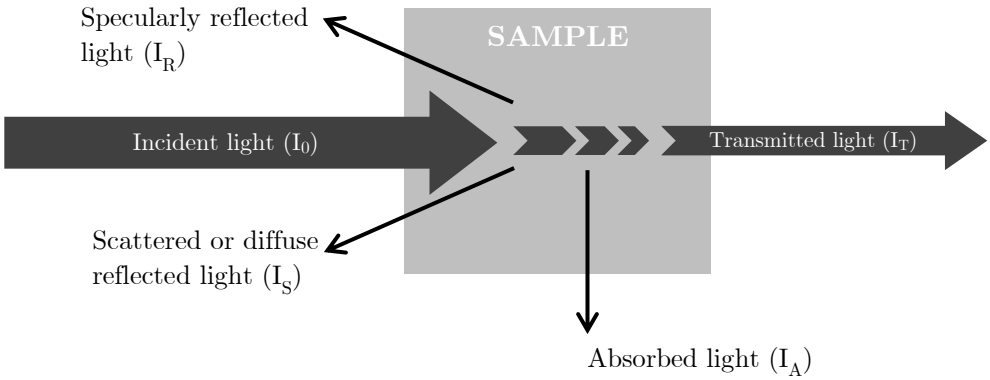


Figure 1.8 Energy balance of the light interacting with a sample. Adapted from Craig, Franca, and Irudayaraj (2015)[16].

1.4.5.1 Law of Beer

In the nineteenth century, Beer demonstrated that chemical compounds could absorb light at certain wavelengths, and found that the

amount of absorbed light (i.e., absorbance) is directly proportional to the concentration of the absorbing species:

$$A = \log \frac{I_0}{I} = -\log T = \varepsilon bc \tag{1.16}$$

where I_0 and I are the intensities of the incident light and the one reaching the detector (I_T), respectively, ε is the absorptivity ($L \cdot mol^{-1} \cdot cm^{-1}$), b is the pathlength (cm), and c is the concentration of the analyte (mol/L)[18,22,26]. Therefore, through the study of the absorption of the light it is possible to gather information about the chemical composition of a sample as well as determine the quantity of its compounds.

1.5 Infrared (IR) spectroscopy

The infrared region covers the radiation with wavelengths ranging from 0.78 to 1000 µm and is divided into three regions according to the differences in the instrumentation needed for data collection, the near-infrared (NIR), the mid-infrared (MIR) and the far-infrared (FIR) regions[22].

1.5.1 History

In 1800, Herschel described the existence of an "invisible light" that goes beyond the visible red and heated a thermometer that he placed in one of his experiments. One year after, Johan Ritter came up with the same conclusion for light beyond the violet. Maxwell (1864) then predicted that ultraviolet, visible, and infrared light are all electromagnetic waves, which was proved afterwards by Hertz in 1888[27].

Interest in IR spectroscopy was based mainly in the MIR region during the first half of the twentieth century due to the difficulties in analysing NIR spectra. The first MIR instrument was commercially available in 1942 by Beckman Instruments, and the first double-beam IR instrument was sold in 1947 by PerkinElmer. The latter popularized the use of MIR spectroscopy in many branches of analytical chemistry including food science.

In the 1950s, NIR spectroscopy was recognized as a useful tool in quantitative analysis being firstly employed as an accessory in many UV/Vis (Cary 14 from Cary Instruments, 1954) and MIR spectrophotometers. The breakthrough for NIR spectroscopy was led by Karl Norris of the US

Department of Agriculture (USDA) when he was studying a way to measure the moisture content of cereal products. Some difficulties were faced because of the interference obtained from other compounds such as protein and starch. Norris solved that problem by using multivariate statistical methods to correlate the NIR spectral features with the reference values. For this reason, Norris is recognized as the "father of NIR spectroscopy". The use of statistical methods in NIR spectroscopy by Norris supposed the starting point for the development of chemometrics. After Norris's work, many applications came out in the food science and agricultural field[27,28].

Another big progress was made when interferometry and an algorithm called Fourier transform (FT) were implemented into MIR spectrometers during the 1960s thanks to the advances achieved in computer science. FTIR instruments have several advantages such as the cost, sensitivity, high-throughput, and precision. Until the 1980s, most MIR equipment were of the dispersive type which used diffraction gratings to separate the light in its constituting frequencies[22,28].

During the last years, the technological advances in hardware (i.e., detectors, light sources, optics) and software has allowed developing faster, smaller, and more accurate instruments. Nowadays is possible to obtain a spectrum within seconds while in the 1990s it could take several hours[28].

1.5.2 Selection rules

Only under certain conditions a molecule absorbs IR light. Satisfying Equation (1.9) does not always imply the appearance of an IR band related to a certain fundamental mode. Through a quantum mechanical treatment of the harmonic oscillator approximation two selection rules are obtained[16,20,23]:

- Only vibrational transitions in which the quantum number changes by unity ($\Delta v = \pm 1$) are allowed
- IR absorption can only take place when the vibration mode changes the dipole moment of the vibrating group

The first selection rule it is not rigorously followed because of the anharmonic behaviour of the chemical bond, as it has been discussed previously (see Section 1.4.3.2). Let us discuss the second selection rule by looking into the fundamental modes of carbon disulphide. Carbon disulphide is a linear molecule that has 4 normal modes ($3N - 5 = 4$), two stretching modes (symmetric and asymmetric), and two bending modes (scissoring; in-plane and

out of plane). As it can be seen from Figure 1.9, the symmetric stretching does not cause any change in dipole moment since both oxygen atoms move away from the central carbon atom at the same time (the distance vectors cancel giving a net dipole moment of zero). This is not the case for the asymmetric stretching, which exhibits a change in the charge distribution and thus in dipole moment. Moreover, both scissoring vibrations produce only a single band in the IR spectrum since they are identical in energy (note that both normal modes are the same but rotated 90°). Quantum states having the same energy are called degenerate[22].

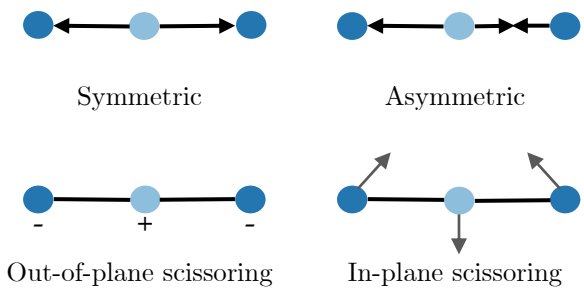


Figure 1.9 Fundamental modes of carbon disulphide.

1.5.3 Mid-infrared (MIR) spectroscopy

MIR spectroscopy measures the absorption of radiation in the range of 4000 to 400 cm^{-1} (from 2.5 to 25 µm in frequency). Most of the transitions observed in this region correspond to vibrational modes since rotational transitions are observed at wavenumbers below 300 cm^{-1}. For this reason, MIR spectroscopy is also called vibrational spectroscopy[29]. MIR region can be divided into two different regions: the functional group (4000-1500 cm^{-1}) and the fingerprint regions (1500-400 cm^{-1}). The functional group region can be further divided into three distinct regions: the single-bond region (4000-2500 cm^{-1}; O-H, N-H and C-H stretching modes), the region of triple bonds (2500-2000 cm^{-1}; C≡C and C≡N), and the double-bond region (2000-1500 cm^{-1}; C=C, C=N and C=O). The functional group region allows the structural determination of molecules[17,30,31]. On the other hand, the interpretation of the fingerprint region is seldom possible since the complexity of its many overlapping bands. However, such complexity and richness in spectral features can be used for identification purposes in food science, such as bacteria typing or food authentication[17,32].

1.5.3.1 Instrumentation

Two major types of MIR equipment are available, dispersive and Fourier transform (FT) instruments[16]. Though, this section will primarily describe the driving mechanism and components of FT spectrometers since they are the most commonly used in MIR analysis.

The most important component of FT spectrometers is the interferometer. The invention of the interferometer is attributed to Albert Abraham Michelson, who constructed the first interferometer in 1891. One year later, Lord Rayleigh recognized that it was possible to retrieve a spectrum from an interference pattern by means of a mathematical treatment, the FT algorithm. However, the implementation of FT spectrometers was hindered by the lack of instrumentation and computing technology[33]. Since the Michelson interferometer is the most used in FT spectrometers, its design will be discussed here.

A Michelson interferometer consists of a source of light, a beam splitter and, fixed and moving mirrors (Figure 1.10). The IR light generated from the source is divided into two beams at the beam splitter. Each of the beams is directed to one of the mirrors, either the fixed or the moving one. The two beams are reflected then onto the beam splitter, where they recombine. The moving mirror generates a difference in the distance travelled by of the beams along the time (i.e., the position of the moving mirror), thus an interference pattern can be obtained. Such distance is called the optical path difference (OPD). The recombined beam is directed to the sample and then the signal captured by the detector[21,33].

When the distance between the mirrors is the same, the beams are in phase (i.e., OPD = nλ, where n is an integer) and constructive interference takes place. In other words, they are at zero path difference (ZPD). The signal intensity at this point is the highest. On the other hand, when the two beams are out of phase they interfere destructively [complete interference shows at OPD = $(n + \frac{1}{2})\lambda$]. At other distances destructive and constructive interference phenomena occur (see Figure 1.5).

The plot of the electrical signal (in volts, V; Figure 1.11) versus optical path difference is called the interferogram, which represents the forward and backward motions of the moving mirror from the initial point to the ZPD point and vice versa, respectively. The point of highest intensity occurs at ZPD and is called the centerburst[33]. Such point is taken as a reference as the

sampling start point when averaging many measurements[22]. The regions on both sides of the centerburst are the wings, in which constructive and destructive interferences take place.

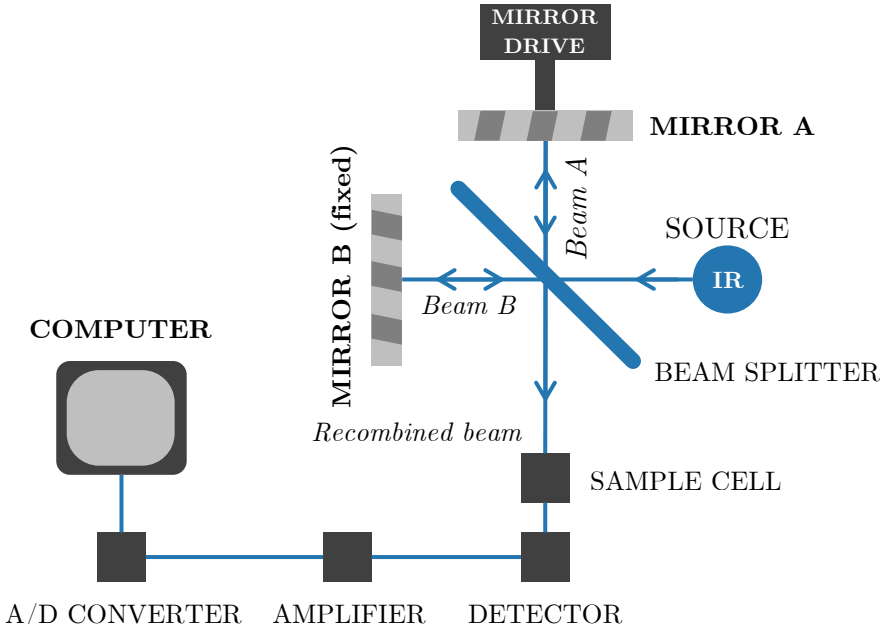


FIGURE 1.10 Schematic diagram of a FTIR instrument. Adapted from Rodriguez-Saona, Ayvaz, and Wehling (2017)[21].

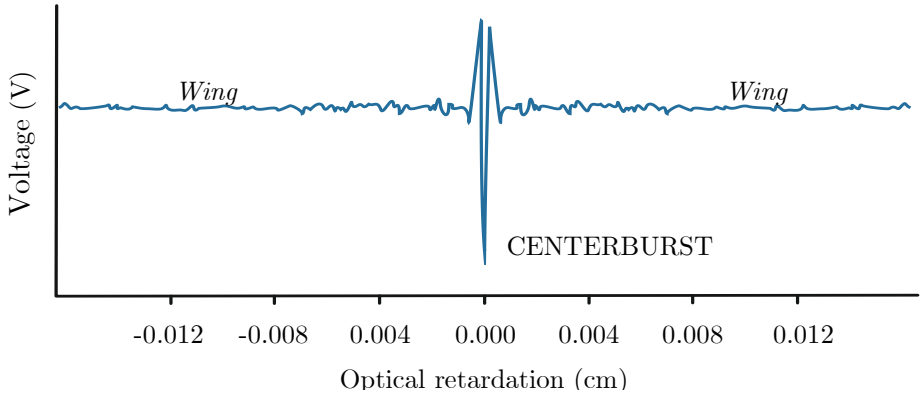


FIGURE 1.11 Typical interferogram obtained with a FTIR instrument. The wings contain information about sample's composition. Adapted from Subramanian and Rodriguez-Saona (2009)[33].

Some mathematical transformations are applied to the interferograms obtained for their better interpretation (Figure 1.12). First, apodization is performed, which consists in multiplying the interferogram by a certain function to avoid the appearance of side lobes or oscillations around spectral features once the Fourier transform is performed. However, the apodization treatment lowers the spectral resolution. The most common apodization functions are the "Boxcar" (no apodization), triangular, Norton-Beer, and Happ-Genzel. Phase correction algorithms are also performed on interferograms to remove their asymmetry (ideally an interferogram should be symmetric on either side of the centerburst). Thus, by applying a phase correction it is possible to record only one of the sides of the interferogram saving time and costs. After such mathematical treatments, the interferogram must be converted into a spectrum by means of the FT. FT basically decomposes a signal into sine waves for each of the wavelengths in the light, then they are arranged to form an interpretable spectrum. After that, the spectrum obtained is truncated to the MIR or NIR region and the reference signal is subtracted (i.e., measured with no sample) from the former to obtain the final spectrum[33].

The sources commonly used in MIR analysis are the Nernst Glower (rare earth oxide), the Globar (silicon carbide), and the Nichrome wire. All sources consist of inert solids that emit continuum IR radiation when heated at temperatures between 1500 to 2200 K. The Globar source has replaced the Nernst Glower since the former can be heated up by passing an electrical current while the latter needs external heat. Refrigeration of the sources can be achieved by means of air or water cooling. On the other hand, detectors can be classified into two categories: thermal or pyroelectric detectors and photoconducting or photosensitive detectors. Pyroelectric detectors are made of dielectric materials (i.e., insulator), being triglycine sulphate $(NH_2CH_2COOH)_3 \cdot H_2SO_4$ (usually deuterated, DTGS) the most commonly used. When IR radiation hits a pyroelectric detector, it changes its temperature and consequently alters its charge distribution, which is detected as a current by an external circuit connected to it. In turn, photoconducting detectors are based on conductivity changes caused by the absorption of radiation in a semiconductor [e.g., lead sulphide, mercury cadmium telluride (MCT), or indium antimonide]. Even though photoconducting detectors exhibit faster detection times than the pyroelectric ones, they typically require cryogenic cooling[21,22,28].

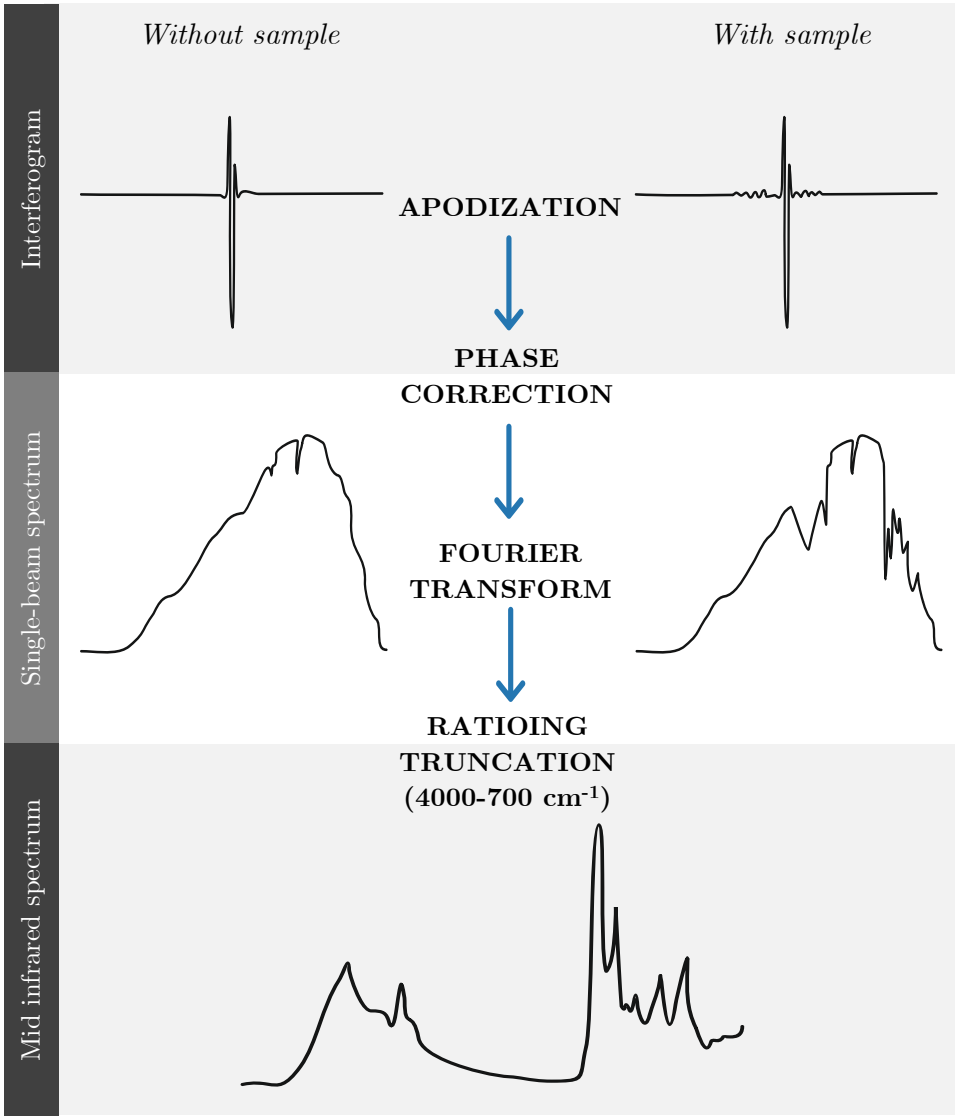


FIGURE 1.12 Processing process before obtaining the final MIR spectrum from a collected interferogram. Adapted from Subramanian and Rodriguez-Saona (2009)[33].

1.5.3.2 Sampling techniques

Sample measuring techniques in IR spectroscopy (i.e., MIR and NIR regions) can be divided into three main types: transmittance, reflectance (diffusive or specular) and a combination of both. The measurement modes mostly used in MIR analysis are transmittance and attenuated total reflectance (ATR).

As discussed in Section 1.4.5, light can undergo different processes as it interacts with matter. In transmittance measurements, the light travels through the sample and is absorbed, and later the remaining light exiting the sample is captured by the detector. If scattering and reflection phenomena are minimized, the absorption of light can be calculated based on the amount of transmitted light measured (see Figure 1.8). Gaseous samples can be measured using cylindrical cells with suitable windows, with pathlengths ranging from few centimetres to several metres, the latter being used in trace analysis. Nevertheless, for liquid and solid samples, the pathlength is restricted to the formation of thin films of less than one centimetre. Liquid samples are measured in cells made of a non-absorbing IR material, such as potassium bromide (KBr) or sodium chloride. Solid samples demand a complex preparation process. They are usually grounded to fine powders to prevent light scattering. Then the powder is mixed with KBr and pressed to form pellet or mixed with a heavy hydrocarbon to produce a mull. A low sample amount is needed in transmittance measurements of solid materials, but this can be a disadvantage when analysing heterogeneous matrixes such as food[16,21,25,32,33].

ATR spectroscopy has become popular recently due to the little or no sample preparation and its high reproducibility[21]. ATR is based on the following principle: the sample is placed onto a crystal of a high refractive index material. When the beam is at the interface between the crystal and a sample of low refractive index such as food, reflection phenomena occur. However, before reflection takes place, the light penetrates the sample in the form of an evanescent wave (i.e., the light electric field decays exponentially across the distance of penetration). The sample absorbs the entering light and once is reflected, the remaining is captured by the detector (Figure 1.13)[16,21,29,32,33]. ATR spectra look quite different from transmission spectra since the depth of penetration is function of the wavelength:

$$d_p = \frac{\lambda_c}{2\pi\sqrt{\sin^2\theta - (n_s/n_c)^2}} \tag{1.17}$$

where λ_c is the wavelength of the light impinging the crystal, θ is the angle of incidence, and n_s and n_c are the refractive indexes of the sample and the crystal, respectively. Thus, the depth of penetration can be properly altered by changing the angle of the incident light or the refractive index of the crystal[32]. Table 1.4 shows the different types of crystals available for ATR measurements. The use of one crystal or another depends on the sample properties and the application.

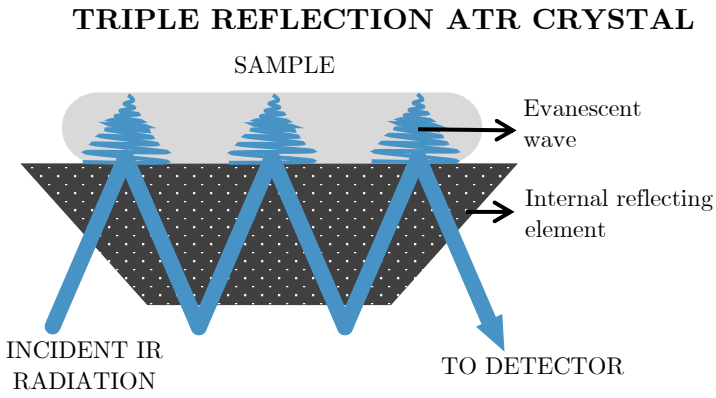


Figure 1.13 Reflection phenomena in an ATR accessory. Adapted from Rodriguez-Saona, Ayvaz, and Wehling (2017)[21].

Table 1.4 Properties of common ATR accessories. Adapted from Rodriguez-Saona, Ayvaz, and Wehling (2017)[21].

Material	Spectral range (cm^{-1})	Hardness (Kg/mm)	pH range	Refractive Index
Diamond	50000-2500[a]	9000	1-14	2.4
Germanium	5000-550	780	1-14	4.0
Silicon	8333-33	1150	1-12	3.4
KRS-5	17900-250	40	5-8	2.4
ZnSe	20000-500	130	5-9	2.4
AMTIR	11000-725	170	1-9	2.5

[a]Also covers the 2200-400 cm^{-1} range.

ATR technique can be used in both solid and liquid samples. However, a pressure clamp system is needed for the solid ones to ensure an intimate contact between the crystal and the sample analysed. Besides high reproducibility and lack of sample preparation, another advantage of ATR technique is that the signal is not influenced by sample's thickness since the depth of penetration is limited to few micrometres (i.e., the amount of sample placed onto the crystal does not need to be standardized in many cases)[21,25]. However, such property becomes a disadvantage when dealing with non-homogeneous samples[33]. Moreover, ATR measurements exhibit lower signal to noise ratios than transmission ones in most of the cases[17].

1.5.4 Near-infrared (NIR) spectroscopy

The region in NIR spectroscopy includes the upper end of the visible region at around 800 nm (12500 cm^{-1}) to 2500 nm (4000 cm^{-1}). Broad bands with overlapping wavelengths are observed in NIR region, which are mainly arising from overtones and combinations of normal modes of vibration that take place in the MIR region (from 3000 cm^{-1} to 1700 cm^{-1}), though some electronic transitions can also be seen.

The NIR region can be divided into two main regions: the short wave near-infrared (SW-NIR), encompassing the region of 800-1200 nm (12500-8500 cm^{-1}), and the long wave near-infrared (LW-NIR), which covers the range of 1200-2500 nm (8500-4000 cm^{-1}). The SW-NIR region, has been also termed "the Herschel region" in honour to William Herschel. The SW-NIR region provides information about the electronic transitions and higher-order overtones. It is expected then that the intensities recorded in this region are low and therefore this region shows high transparency to many materials. This property becomes an advantage for biomedical applications, such us blood oxygen level monitoring. The LW-NIR region can be further divided into two other regions. The region from 1200 to 1800 nm (8500-5500 cm^{-1}) contains the first and second overtones of carbon-hydrogen, oxygen-hydrogen, and nitrogen-hydrogen chemical bonds (stretching vibration and combination modes). The region of 1800-2500 nm (5500-4000 cm^{-1}) mainly exhibits combination bands and shows poor permeability.

1.5.4.1 Instrumentation

NIR instruments consist, as in MIR equipment, of a source, a wavelength selector or interferometer, and a detector, all of them being interfaced by different optical materials. NIR instruments can be classified into two types: Fourier transform and dispersive-like[34]. Even though FT instruments are popular in MIR devices (i.e., based on the use of an interferometer), dispersive-like devices are predominant in most portable NIR measurements (benchtop spectrometers almost entirely adopted the FT approach). The setup of dispersive NIR equipment can be divided into two different types. Many instruments use a wavelength selector to disperse the white light by wavelength, to later irradiate the sample and record the intensities sequentially. On the other hand, in some devices the white light hits the sample, which is then separated into its constituting wavelengths. Then, a multichannel array detector measures all the wavelengths at once[21].

The mechanism of wavelength separation plays a major role in IR instrumental design and thus MIR and NIR instruments can be also classified according to the wavelength selector used. Based on that, there are monochromator-based instruments and filter-based instruments[28].

A monochromator is a device used to disperse and convert light with wide range of wavelengths to a monochromatic light. Formerly, prisms were commonly used as dispersion devices for both NIR and MIR instruments. However, most materials used for prism construction had a lower transmittance and a limited useful wavelength range and thus diffraction gratings replaced them[28,35]. A diffraction grating is a piece of substrate with fine parallel groves equally spaced and coated with a reflecting material such as aluminium or gold (Figure 1.4). When the light strikes the groves of the grating, each of its constituting wavelengths is dispersed at a specific angle. Each of the grooves acts as a new source and interferences among the reflected beams are produced. As stated in Section 1.4.1.1, for constructive interference to occur, the difference in the paths between two different light beams must be equal to nλ. Therefore:

$$n\lambda = \left(\overline{CB} + \overline{BD}\right) = d(\sin i + \sin r) \tag{1.18}$$

where n is the diffraction order and can be zero or any integer, d is the spacing between grooves, and i and r are the angles of the incident and diffracted light,

respectively. Equation (1.18) implies that several orders of diffraction might be found at a same angle r, which is named as multiple order of diffraction. This can be minimized by using filters that remove the higher-order lines, such as glass, which absorbs and eliminates those wavelengths that are in the visible region[28,35,36].

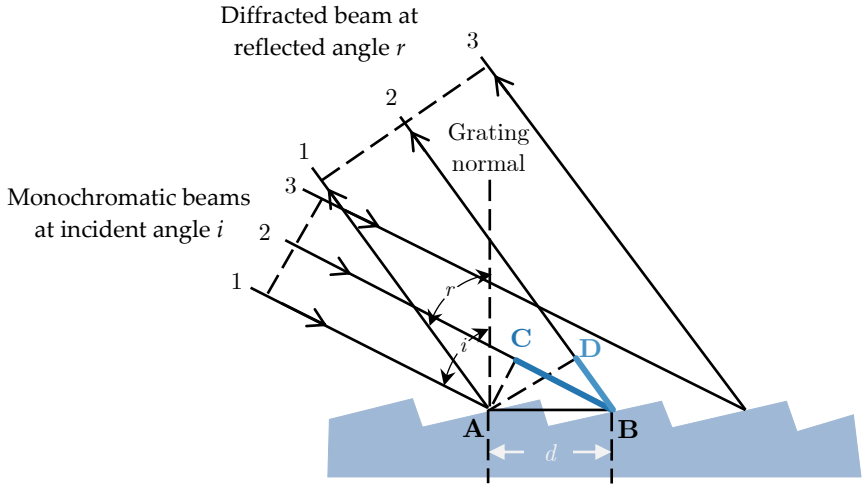


Figure 1.14 Scheme illustrating the diffraction phenomena produced in a grating selector. Adapted from Skoog, Holler, and Crouch (2018)[36].

On the other hand, filters can also be used as wavelength selectors and can be divided into interference filters and electrically tuneable filters[28]. Only interference filters will be introduced in this section since they are the most commonly used. Interference filters consist of multiple layers of dielectric films placed between semi-transparent films made of a highly reflective material. Such array of films is sandwiched between two transparent plates of glass. The light undergoes several reflections inside the dielectric films and, depending on its thickness, some wavelengths are enhanced due to constructive interference while others are attenuated by destructive interference. Then, the wavelength λ transmitted through the filter is given by:

$$m\lambda = 2nd\cos\theta \tag{1.19}$$

where m is the order of interference, n the refractive index of the dielectric, d the thickness of the dielectric layer, and θ the angle of the incident beam. Hence, by changing the thickness of the dielectric layer throughout the filter it is possible to adjust the wavelengths obtained from a polychromatic source[35,36]. Circular variable filters (CVF) are based on this approach. CVFs are rotated along their central axis and then the light transmitted through a certain point of the filter is used for the spectral analysis. Linear variable filters (LVF) follow the same principle but in turn are rectangular and move along the linear direction[35].

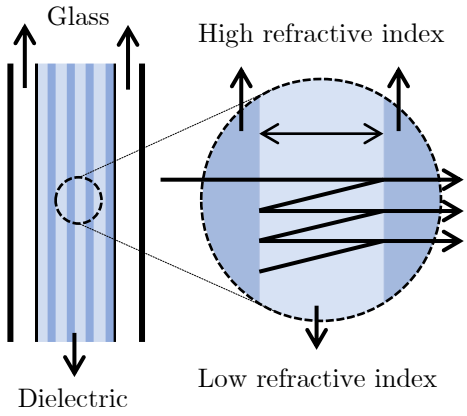


FIGURE 1.15 Schematic diagram of an interference filter. Adapted from Ikehata (2021)[35].

Sources used in MIR equipment like the Globar or the nichrome wire are also used in NIR instruments. However, the most popular light source in the NIR region is the tungsten-halogen lamp with a quartz bulb. As the thermal sources discussed in SECTION 1.5.3.1, the tungsten-halogen lamp needs to be heated at high temperatures to produce light emission[35]. The halogen gas presents some cleaning action by depositing back the evaporated tungsten placed in the walls of the bulb. Thus, such sources exhibit more stable performances and longer life when compared to standard tungsten filament lamps. On the other hand, for applications where the full spectrum acquisition is not necessary, lasers, LEDs, and laser diodes are used. Ruggedness and small sizes can be obtained when using the later sources[28].

As for detectors, they range from photoconductors to image sensors. Examples of photoconductors, which were described in SECTION 1.5.3.1, are the

indium gallium arsenide (InGaAs) and the lead sulfide (PbS) detectors. An image sensor is a detector that consists of multiple sensors (i.e., semiconductors) arranged in a row (linear image sensor) or in two dimensions (area image sensor). An example of such detectors is the charge couple device (CCD) sensor. Image sensors are usually employed in dispersive instruments, where multiple wavelengths have to be captured simultaneously[21,28,35].

1.5.4.2 Sampling techniques

Different sampling modes are available in NIR analysis depending on the application, being transmittance (discussed in Section 1.5.3.2) and diffusive reflectance measurements the most commonly performed. Interactance and transflectance have been also applied, but less frequently.

When a beam of light strikes the surface of a fine powder, some of the light is directly reflected back (specular reflection) from the surface without entering the sample, but the rest is partially transmitted through. The light inside the sample is scattered and absorbed many times, with part of it escaping the sample. The latter is called diffusive reflectance light and gathers information about the chemical composition of a sample like the one provided in transmission measurements. Reflectance is calculated then:

$$R = \frac{I}{I_0} \tag{1.20}$$

where I and I_0 are the intensities measured at a given wavelength from the sample and the reference (i.e., non-absorbing material). Then, the analogous absorbance expression for the reflectance can be calculated as $\log(1/R)$. Reflectance measurements involve a faster and easier preparation of the sample compared to transmittance ones. However, the signal to noise ratio is usually lower. Food products are ideal for this sampling mode because of their physical properties (i.e., rough shape and surfaces)[17,21,33,35].

Transflectance can be considered as a combination of transmission and reflection modes. The sample is placed in a non-absorbing container (e.g., quartz) and the light transmitted is reflected by means of a high-reflective material (e.g., gold) that is placed behind the sample. Transflectance mode is suitable for transmittance measurements of clear liquids and smooth surfaces which otherwise cannot be measured by diffusive reflectance technique since they either transmit or specularly reflect the light[17,35].

1.6 Raman spectroscopy

It is considered that Raman spectroscopy was born when C. V. Raman and K. S. Krishnan discovered in 1928 that a small fraction of the light scattered by a molecule undergoes a wavelength shift and this change in the wavelength depends on the chemical structure of the scatterer. Raman was awarded with the Nobel Prize in physics in 1931 for this discovery. Raman spectroscopy was not widely applied until the invention of the laser sources in the 1960s[24,37].

1.6.1 Inelastic scattering

When a photon of energy hν strikes a certain molecule, the former can be scattered in different ways. During the collision, the molecule reaches an unstable vibrational energy state called virtual state that lasts for a very short time. Most of the excited molecules relax to the ground state and reemit the photon having the same energy as the incident one. This process is known as Rayleigh scattering. However, few molecules relax and emit radiation with shifted frequency (i.e., energy). This is called inelastic scattering or Raman scattering. In all these processes the photon does not need to match the energy between the different vibrational energy states. It is noteworthy to mention that the shift in energy corresponds to the energy gap between the ground and excited vibrational energy states. Therefore, energy shifts observed in Raman analysis should be the same as the energy transitions obtained in IR absorption measurements. Two different scattering events are observed in Raman phenomena. After reaching a certain virtual energy state, molecules can either exhibit energy transitions from the ground state to an excited vibrational state or vice versa. Such processes are called Stokes and anti-Stokes shifts, respectively[21,24,37]. Figure 1.15 illustrates all these processes and IR absorption.

1.6.2 Selection rules

When the electric field component (E) of radiation [with a certain frequency (ν_{ex})] interacts with the electron cloud around the bond of a certain molecule, a dipole moment μ is induced:

$$\mu = \alpha E = \alpha E_0 \cos(2\pi\nu_{ex} t) \tag{1.21}$$

where α is a constant called the polarizability of the bond and measures the deformability of the electron cloud around the bond of a molecule. In turn, the polarizability changes as the distance between nuclei varies:

$$\alpha = \alpha_0 + (x - x_e)\left(\frac{d\alpha}{dx}\right) \tag{1.22}$$

where α_0 is the polarizability at the equilibrium internuclear distance x_e. The distance between nuclei also changes depending on the frequency of vibration ν_v:

$$x - x_e = x_0 \cos(2\pi\nu_v t) \tag{1.23}$$

Therefore, by substituting Equations (1.22) and (1.23) into Equation (1.21) and doing some rearrangements:

$$\begin{aligned}\mu = \alpha_0 E_0 \cos(2\pi\nu_{ex} t) + \frac{E_0}{2} r_m \left(\frac{d\alpha}{dr}\right) \cos[2\pi(\nu_{ex} - \nu_v)t] \\ + \frac{E_0}{2} r_m \left(\frac{d\alpha}{dr}\right) \cos[2\pi(\nu_{ex} + \nu_v)t]\end{aligned} \tag{1.24}$$

The first, second and third terms in this equation correspond to the Rayleigh scattering (ν_{ex}), and the Stokes ($\nu_{ex} - \nu_v$) and anti-Stokes ($\nu_{ex} + \nu_v$) shifts in Raman scattering, respectively. It is noteworthy to mention that it can be concluded from Equation 1.24 that for Raman scattering to occur, a molecule must undergo a change in the polarizability during the vibration ($d\alpha/dr \neq 0$)[37,38]. Reconsidering once more the normal modes of vibration of the molecule of carbon disulphide (already discussed in Section 1.5.2), the symmetric stretching of this molecule is not active in the IR since it does not exhibit any change in its dipole moment. However, when the carbon disulphide vibrates symmetrically, the electron cloud changes and the total polarizability is not zero. Then, such fundamental mode is Raman active. Figure 1.16 displays a summary of the IR and Raman active modes of vibration based on their selection rules[24].

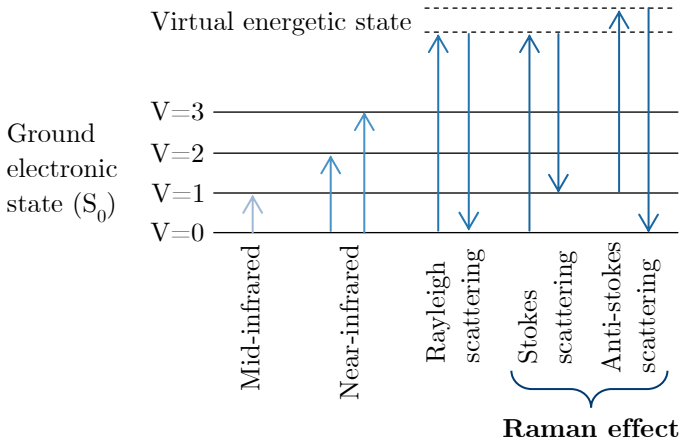


FIGURE 1.15 Rayleigh scattering and Raman and IR transitions.

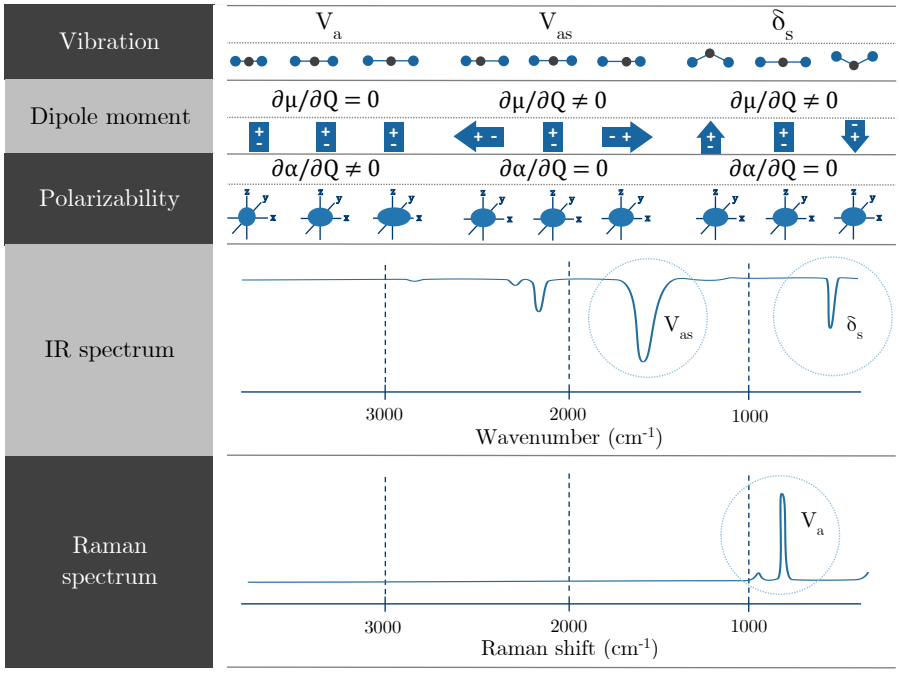


FIGURE 1.16 IR and Raman active fundamental modes of carbon disulphide with their resultant spectra. Adapted from Smith and Dent (2019)[24].

1.6.3 Instrumentation

The optical components in Raman spectrometers, either dispersive or Fourier transform, are similar to the ones discussed in prior sections. Dispersive-like Raman spectrometers use interference filters to reject the Rayleigh bands (i.e., having the same wavelength as the source) emitted from the sample and a diffractor grating to disperse the rest into its different wavelengths, whose intensities are detected in a multichannel detector, such as a CCD array detector. On the other hand, Fourier transform Raman (FT-Raman) instruments use a Michelson interferometer that produces and interference pattern from the scattered light coming from the sample. After that, the interference pattern is collected by a transducer such as InGaaAs and nitrogen cooled Ge detectors, and a spectrum is obtained by means of a Fourier transform algorithm[21].

Lasers are used as light sources since they are monochromatic, and their high intensity allows to easily observe the weak Raman bands. There are several lasers of different wavelengths available. In dispersive instrumentation the typical wavelengths used are 785, 633, 532, and 473 nm. However, fluorescence phenomena occur at shorter wavelengths because they possess higher energy to match the electronic energy transitions. As the wavelength becomes higher, fluorescence is minimized. FT-Raman spectrometers commonly use near-IR sources (1064 nm) that eliminate the fluorescence by sacrificing the intensity of the Raman scattering, since the latter is proportional to $1/\lambda^4$[21,37,39].

1.7 Multivariate analysis

Over the last decades food authentication has been based on the non-targeted analysis of food matrices, also known as fingerprinting. Nowadays, modern analytical instruments focused on fingerprinting analysis provide significant amounts of data which involve several factors at a time. There is the need for analysis tools that can extract meaningful information from such big amounts of data[40]. Multivariate analysis (also referred as chemometrics) was born with the use of statistical multivariate methods by Karl Norris with the aim of evaluating big sets of highly correlated NIR data for predicting a certain variable of interest (see Section 1.5.1). Furthermore, chemometrics also refers to the use of mathematical techniques to reduce noise and interferences

originated during data collection (i.e., data pre-processing)[41]. Chemometric analysis has also applications on the observation of natural groupings or trends in sample's distribution that are consistent with its chemical or physical properties (i.e., discrimination of authentic and non-authentic food products)[42].

Chemometric methods can be divided in exploratory/unsupervised, classification/supervised and prediction methods[14,40,42,43]. In this section, one exploratory, one classification, and one prediction algorithm will be discussed in depth, which have been the approaches taken to analyse the spectroscopic data in the present thesis.

Moreover, when building up a chemometric model, a common sequence is followed. First, the data is pre-processed to remove noise and other interferences. Then, a set of data called training set is used to build up the models, a process known as calibration step. Finally, the model performance needs to be tested for future samples, so a validation set of data (either internal or external) is used for that purpose[40,41].

1.7.1 Exploratory analysis: principal component analysis (PCA)

Exploratory analyses are usually the first approach to take when analysing multivariate data. They are performed to evaluate the repeatability of the measurements, for outlier detection, and as a variable reduction tool (i.e., simplification of the dataset)[43]. Furthermore, these methods are also called unsupervised since they are able to identify clusters without any prior knowledge about the structure of the data (i.e., sample groupings are not stated before running the algorithm). PCA and hierarchical cluster analysis (HCA) are the most common exploratory analysis used in food fingerprinting[14,40].

PCA is a dimension reduction algorithm that computes new variables that are linear combinations of the originals (i.e., wavelengths or wavenumbers). This eases the visualization and interpretation of highly collinear and complex data[43]. Considering a matrix of spectroscopic data X, which contains the absorbance values x at n different wavelengths (columns) for m different samples (rows):

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1n} \\ x_{21} & x_{22} & \cdots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mn} \end{bmatrix} \tag{1.25}$$

The covariance of the first and second variables (v_1 and v_2, respectively) can be calculated as:

$$\mathrm{cov}(v_1, v_2) = \frac{1}{n-1}\sum_{n=1}^{m}(x_{1n} - \bar{x}_1)(x_{2n} - \bar{x}_2) \tag{1.26}$$

where $\bar{x}_1$ and $\bar{x}_2$ are the averages of the first and second variables, respectively. The covariance is a measure of the joint variability between two different variables. To compute the covariance matrix of X, a new matrix (mean-centred) must be defined:

$$U = \begin{bmatrix} x_{11}\text{-}\bar{x}_1 & x_{12}\text{-}\bar{x}_2 & \cdots & x_{1n}\text{-}\bar{x}_n \\ x_{21}\text{-}\bar{x}_1 & x_{22}\text{-}\bar{x}_2 & \cdots & x_{2n}\text{-}\bar{x}_n \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1}\text{-}\bar{x}_1 & x_{m2}\text{-}\bar{x}_2 & \cdots & x_{mn}\text{-}\bar{x}_n \end{bmatrix} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{m1} & u_{m2} & \cdots & u_{mn} \end{bmatrix} \tag{1.27}$$

Then, the covariance matrix of X can be obtained by multiplying U^T by U and dividing by n−1:

$$\begin{aligned} \frac{1}{n-1}(U^T U) &= \frac{1}{n-1}\begin{bmatrix} u_{11} & u_{21} & \cdots & u_{m1} \\ u_{12} & u_{22} & \cdots & u_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ u_{1n} & u_{2n} & \cdots & u_{mn} \end{bmatrix}\begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{m1} & u_{m2} & \cdots & u_{mn} \end{bmatrix} = \\ &= \frac{1}{n\text{-}1}\begin{bmatrix} \sum u_{m1}^2 & \sum u_{m1}u_{m2} & \cdots & \sum u_{m1}u_{mn} \\ \sum u_{m1}u_{m2} & \sum u_{m2}^2 & \cdots & \sum u_{m2}u_{mn} \\ \vdots & \vdots & \ddots & \vdots \\ \sum u_{m1}u_{mn} & \sum u_{m2}u_{mn} & \cdots & \sum u_{mn}^2 \end{bmatrix} \\ &= \mathrm{cov}(X) \end{aligned} \tag{1.28}$$

The elements of the diagonal constitute the sample variance of each of the variables of the data set (e.g., $s_1^2 = \sum u_{m1}^2$), while the elements out of the diagonal are the covariances among variables [e.g., $\mathrm{cov}(v_1, v_2) = \frac{1}{n\text{-}1}\sum u_{m1}u_{m2}$].

It is worth mentioning that the matrix is symmetrical [i.e., $cov(v_1, v_2) = cov(v_2, v_1)$].

The main objective in PCA is to eliminate those variables that contain redundant information. To achieve that, a spectral or eigenvalue decomposition of the covariance matrix of X is performed:

$$cov(X) = L\Lambda L^{-1} \tag{1.29}$$

where L is the matrix of eigenvectors or loadings and Λ is a diagonal matrix of eigenvalues. The superscript -1 in matrix L indicates the inverse. Then, the mean-centred data U can be represented in terms of the loadings:

$$Y = UL \tag{1.30}$$

Y is the scores matrix, whose covariance is:

$$cov(Y) = \frac{1}{n-1}Y^TY = \frac{1}{n-1}(UL)^TUL = \frac{1}{n-1}L^TU^TUL \tag{1.31}$$

By combining Equations (1.28) and (1.31), substituting the output with Equation (1.29) and applying the orthogonal property $L^TL = I$ give:

$$cov(Y) = L^Tcov(X)L = L^TL\Lambda L^{-1}L = \Lambda \tag{1.32}$$

The covariance matrix of Y turns out to be diagonal, thus no correlation exists among the different variables (i.e., covariances take zero values) that build up the new principal component space (matrix L). The diagonal of matrix Λ give the variability explained (i.e., variance) by each principal component[44–46]. The first principal component gathers the highest variability, the second lies on the direction of the highest remaining variability and is orthogonal to the first, and so on. Figure 1.17 illustrates this. Basically, principal components are new variables that account for the sample's maximum variance (i.e., separation between data points), which facilitates the visualization of groupings in very complex datasets. In the banana example, the first component, which accounts

for most of the variance, represents the length of the banana. The second corresponds to the span of its curvature, and the third is the axis drawn at right angles in respect to PC1 and PC2. As it can be seen in the lower left side of the figure, first and second principal components provide the most information about the structure of the banana.

Other PCA outcomes to consider are the loadings and the scores. Loadings provide information about the variables that are most important for sample separation and are used for computing the scores, which in turn are a weighted sum of the former ones and are represented in the space of principal components (i.e., PC1, PC2, and so on). Examples of loading and score plots can be found in the next chapters.

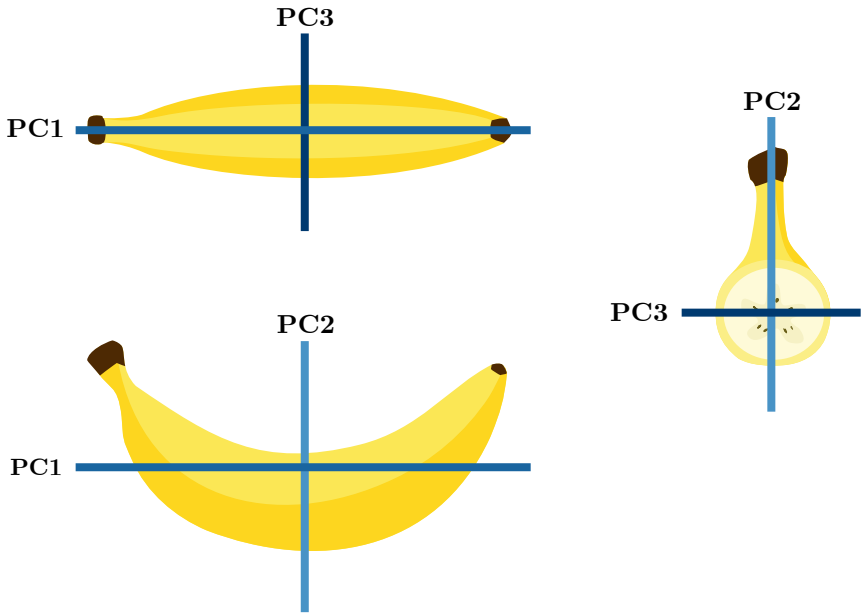


FIGURE 1.17 Banana views based on the three first principal components.

Finally, the spectral decomposition in PCA algorithms is no longer used because of its low efficiency in terms of time consumption. Instead, singular value decomposition (SVD) is applied. For understanding the PCA approach and its outputs only its computing through eigenvalue decomposition will be explained[45].

1.7.2 Classification methods: soft independent modelling of class analogy (SIMCA)

The aim of supervised methods, also named pattern recognition techniques, is to classify samples based on similarities (i.e., spectral features). Such methods require information about class membership prior building up the model and can be divided into discriminant analysis and class-modelling analysis. The first aims to divide, based on different delimiters (e.g., Euclidean distances), the data space in different regions that correspond to each of the samples to be analysed. In turn, discriminant analysis focuses on modelling each class independently so samples can be either assigned or rejected to that class[14,40,42,43]. The classification methods mostly used are SIMCA and partial least squares-discriminant analysis (PLS-DA).

SIMCA is a class-modelling technique in which each class is defined through its own model of principal components (see Section 1.7.1), which is being described in the matrix of loadings L. The mean-centred data U_1 belonging to the first class (e.g., authentic samples) is projected on the principal component space of the second class (e.g., fraudulent samples), and then projected back to the space of original variables:

$$\overline{U}_1 = U_1 L L^T \tag{1.33}$$

Then, the matrix of residuals can be obtained:

$$E = \overline{U}_1 - U_1 \tag{1.34}$$

Every row vector (e_i) in E corresponds to each of the samples in the first class. A residual between first and second classes can be computed:

$$s_{12} = \sqrt{\left(\frac{1}{(m - k_2)n_1} \sum_{i}^{n_1} e_i e_i^T\right)} \tag{1.35}$$

where m is the number variables (e.g., wavelengths), k_2 the number of principal components of the second-class model, and n_1 the number of samples in the

first class. The residual gives information about the degree of separation among classes which can be represented in the so-called Coomans plot (examples can be found on Chapter 2 and 5). Another way to measure class separation is the Interclass Distance (ICD):

$$ICD_{12} = \sqrt{\left(\frac{s_{12}^2 + s_{21}^2}{s_{11}^2 + s_{22}^2}\right)} - 1 \tag{1.36}$$

As it can be seen in this equation, the interclass distance is a ratio of interclass to intraclass distance (based on residuals) among two different classes. This formula is repeated for each pair of classes; thus, a symmetrical square matrix is obtained since $ICD_{12} = ICD_{21}$ (the diagonal elements are zero). High ICD values indicate good separation between two different classes, which are considered statistically different when they reach an ICD value higher than 3.

Furthermore, similar to the information provided by the loadings in PCA, the discriminating power is used to assess which variables are most important for separating the classes (examples of its use can be found in the next chapters). Discriminating power plots can take values from 0 to infinite with higher values meaning more power for discriminating among classes. Finally, other parameters such as the Mahalonbis distance (i.e., variance weighted Euclidean distances) are used for outlier detection. If the Mahalanobis distance of a certain sample lies beyond a threshold defined by a probability distribution (e.g., F-distribution), such sample will be considered an outlier[47,48].

1.7.3 Regression methods: partial least squares regression (PLSR)

The aim of regression methods is to stablish a relationship between the concentrations of an analyte and the signal obtained. In spectroscopy is possible to do that because of the law of Beer:

$$\text{signal} = \text{concentration} \cdot \text{sensitivy} + \text{error} \tag{1.37}$$

However, in multivariate calibration an inverse of Equation (1.37) is needed in order to obtain the concentrations of interest from the multivariate signal (m variables = concentrations) collected in a set of samples (i = 1,2, ..., n):

$$y_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \ldots + b_m x_{im} + e_i \quad (1.38)$$

Or expressed in matrix notation (with mean-centring to remove the b_0):

$$y = X^T b + e \quad (1.39)$$

Where y is a column vector of analyte concentrations, X^T is the spectral signal matrix, b is the column matrix of regression coefficients (i.e., is a weighted sum of independent variables, either wavelengths or wavenumbers), and e is the vector of concentration residuals. The aim is to obtain the vector of regression coefficients by applying the least squares criterion (i.e., the error sum of squares is minimized):

$$e^T e = (y - Xb)^T (y - Xb) \quad (1.40)$$

Then:

$$y - Xb = 0 \quad (1.41)$$

The vector of regression coefficients can be obtained:

$$b = (X^T X)^{-1} X^T y \quad (1.42)$$

The prediction of the concentration of unknown samples is based on the vector of regression coefficients. However, due to the high collinearity of the spectroscopic data such approach, known as Multivariate Linear Regression (MLR), gives inaccurate and poor predictions. To solve this problem, PLSR and other algorithms such as principal component regression (PCR) were proposed and are based on extracting all the redundant information as in the PCA approach.

In PLSR, the redundancy from high collinear spectral data is extracted by means of an algorithm called non-iterative partial least squares (NIPALS). Unlike PCR, which is based on SVD (see Section 1.7.1), PLSR gives two

different loading matrixes, which are intended to explain the maximum variance among the spectral data (weight loadings) and the maximum correlation between the data X and the concentration y (loadings). Because of that, the PLSR algorithm usually yields more parsimonious (i.e., less factors) and reliable models[49–51].

1.7.4 Model validation: figures of merit

After model development a validation step is conducted to test its performance in future unknown samples. Validations can be either internal or external. The most performed internal validation is the cross-validation approach, and different strategies can be taken when selecting the samples from the training set to perform the prediction, such as the leave-one-out, leave-n-out, random subsets, venetian blinds, among others (other internal validation approaches will be discussed in the following chapters). On the other hand, an external validation uses an independent sample set (i.e., test set), which provides a better and more reliable estimate of the model prediction ability[41,52].

Depending on the type of model obtained, either for classification or prediction purposes, a set of different validation parameters can be assessed. The output obtained when a sample is classified in supervised methods is whether it belongs (or not) to a certain category. Different outcomes are obtained when comparing the output to the real membership of the sample: true positive (TP), true negative (TN), false positive (FP), false negative (FN), and not assigned or assigned to more than one category. From these outcomes, parameters such as sensitivity and specificity are studied. Sensitivity is the percentage of samples belonging to a certain category that are correctly identified by the model:

$$\text{sensitivity} = \frac{\text{TP}}{(\text{TP} + \text{FN})} \cdot 100 \tag{1.43}$$

On the other hand, the specificity accounts for the percentage of samples belonging to other classes that are correctly rejected by the model:

$$\text{specificity} = \frac{\text{TN}}{(\text{TN}+\text{FP})} \cdot 100 \tag{1.44}$$

In turn, prediction models can be assessed in terms of the errors obtained when predicting test samples. The coefficient of determination (R^2) is used to assess how well the measured responses fit with the ones computed by the model, with higher values of R^2 meaning a good linear fit. Besides the R^2, the root mean square error can be calculated for the cross-validation (RMSECV) and prediction steps (RMSEP):

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{1.45}$$

where n is the number of samples in the training or calibration set, $\hat{y}_i$ are the responses obtained during the cross-validation or prediction (i.e., external validation), and y_i are the measured responses. From RMSE derive the standard error of cross-validation (SECV) and prediction (SEP). Other parameters provided from the plot of predicted values vs measured ones are the slope and intercept[52].

1.8 Towards miniaturization: portable and hand-held technologies

1.8.1 A brief overview

The properties of IR and Raman spectroscopies make these technqiues suitable for the in-situ analysis of food products. They are fast, sensitive, robust, and a minimal sample preparation is required, which allow them to be used outside of the laboratories under challenging conditions. Nowadays, thanks to the advent of nano and materials sciences, benchtop spectrometers have been down-sized to portable and hand-held devices, some of them attaining palm-size dimensions.

All miniaturized spectrometers are based on the integration of microelectromechanical systems (MEMS), which substitute the big and heavy components of benchtop spectrometers, such as interferometers and diffraction gratings, and are fabricated on silicon substrates by means of microfabrication processes (e.g., chemical vapor deposition and photolithography). Besides MEMS, the emergence of new sources such as lasers and LEDs and high-

performance detector arrays has also flattened the path towards miniaturization (see Section 1.5.4.1)[21,34,53].

However, not all technologies have been miniaturized to the same extend. NIR spectroscopy has been the technology that exhibited most of the advances in miniaturization due to the resolution requirements (i.e., NIR has broader bands than MIR spectroscopy) and the availability of more efficient sources. For this reason, even though MIR spectroscopy has been scaled down, the size of most miniaturized MIR spectrometers has been limited to the one of a lunch box (e.g., Cary 360, 4500, and 5500 by Agilent Technologies) due to the need of moving parts (i.e., Michelson interferometer), higher resolutions, and constraints in detector cooling[53].

Most NIR miniaturized devices use single and array InGaAs detectors, and tungsten or tungsten-halogen lamps and LEDs as NIR sources. Most of the miniaturized spectrometers are dispersive-like, in which a wide range of wavelength selectors can be used. The microPHAZIR hand-held NIR spectrometer from Thermoscientific, which was the first miniaturized NIR device to be launched to the market, is based on the use of a Hadamard mask implemented through MEMS technology. The complex spectral signal obtained is resolved into its constituting wavelengths by means of an inverse Hadamard transform. Several spectrometers have implemented other wavelength selectors such as diffraction and holographic gratings and LVFs (see Section 1.5.4.1). Finally, one of the latest NIR sensors to be commercially available was the palm-sized FT-NIR instrument based on a nanoscale Michelson interferometer launched by Si-Ware Systems (Neospectra Micro)[34].

Raman spectrometers have also experienced a certain degree of miniaturization. Most of them operate with 785 nm excitation sources and CCD detectors (see Section 1.6.3). One example of such spectrometer is the hand-held MIRA analyser by Metrohm, that incorporates the orbital raster scan technology that tries to overcome the errors produced when statically focusing the beam on non-homogeneous samples. There are also FT-Raman approaches based on the use of a 1064 nm laser source, to minimize the fluorescence signal from biological and food samples[21].

1.8.2 Applications of miniaturized spectroscopic devices

Rapid and non-destructive methods are preferred because of their online and in-field capabilities. Thus, miniaturized spectrometers are good candidates for such purpose. However, the miniaturization of infrared and

Raman spectroscopies has been achieved by compromising to some extend their resolution and sensitivity. Therefore, with the emergence of such wide range of small spectroscopic sensors, there is the need to assess their suitability in food applications. A bibliographic search using the Web of Science (WoS) database was conducted using the following terms: "spectro*" (i.e., the * sign includes spectroscopy and spectrometer), "hand-held OR portable", "food" and "infrared OR Raman"[54]. From all the results retrieved by the database, only the ones from last 9 months were considered, which are displayed in Table 1.5.

Most of the articles published during the last months regarding spectroscopy and food are based on the use of NIR spectroscopy because of its unique advantages when compared to other vibrational spectroscopic techniques. NIR spectroscopy can evaluate higher samples volumes, which makes this technology suitable for quantitative analysis (i.e., higher reproducibility). Moreover, aqueous samples can be measured without any prior preparation because of the low absorptivity of water in the NIR region. On the other hand, this technique has limitations such as the low selectivity and complexity of the spectra, thus it becomes challenging to obtain insights in the chemical composition of the sample analysed[16,21,23,28,53].

NIR sensors have been implemented in different applications to prove their suitability outside of the laboratory. Kappacher et al. (2022) reported that the results obtained using a NIR sensor when performing predictions of the harvesting time of black truffle where similar to the ones obtained by means of a benchtop FT-NIR spectrometer (e.g., $R^2_{VAL} = 0.86$ and $R^2_{VAL} = 0.75$ for the benchtop spectrometer and the sensor, respectively), proving the capability of such small devices to be implemented for the in-situ analysis of food. The authors were also able to successfully classify different truffle species achieving sensitivities of 100% for all the instruments tested[55]. The performance of NIR sensors have been also assessed for the discrimination of bitter and sweet almonds. In the studied conducted by Torres et al. (2021), two different sensors were evaluated, and high sensitivities and specificities were achieved (see Table 1.5)[56]. Other authors have applied NIR sensors for food safety purposes. Cebrián et al. (2021) presented a novel use of a NIR sensor for the discrimination of Ochratoxin A (OTA) producing moulds and non-producing ones[57]. However, NIR spectroscopy has been mostly tested in the meat science field. Dashti et al (2022) reported the use of two hand-held spectrometers for the speciation of meat for halal certification purposes, obtaining high accuracies [e.g., 94.7% with support vector machine (SVM)

models]⁵⁸. In another study, the composition and quality characteristics of meat were studied in the abattoir by means of portable and hand-held sensors. However, Patel et al. (2021) reported poor prediction abilities for parameters such as the protein and lipid content, with coefficient of determination values of 0.42 and 0.62, respectively⁵⁹.

Publications regarding the use of MIR portable spectrometers are less common due to the restrictions of this technology for their miniaturization. However, they have been used in applications where the sensitivity and the chemical information are important factors to consider. A hand-held (TruDefender) and a portable ATR-FTIR (Agilent 5500) spectrometers were used to determine the acrylamide content in French fries, obtaining high determination coefficients (R^2_{VAL} = 0.99 and R^2_{VAL} = 1.00)⁶⁰. In the study of Ma et al. (2021), the same hand-held spectrometer was used for the quantification of nitrate in vegetable leaves with a standard error of prediction (SEP) of 799.7 mg/kg and a R^2_{VAL} of 0.93⁶¹.

Finally, the feasibility of miniaturized Raman spectrometers for food science applications has been also investigated. Raman spectroscopy also has some advantages when compared to the conventional FT-IR spectroscopy (i.e., MIR). The weak Raman scattering of water is one of the major advantages, thus aqueous samples can be analysed. Moreover, the confocality of the laser makes this technique suitable for the measurement of reagents through transparent packaging. The main commodity that has been lately studied by means of portable Raman sensors are vegetable oils. Varnasseri et al. (2021) proved the capability of a hand-held Raman spectrometer when performing measurements through bottles of olive oil to quantify the adulteration levels with cheaper ones, achieving a high prediction ability (R^2_{VAL} = 0.99)⁶². Furthermore, Martín-Gómez et al. (2021) classified dry-cured ham based on their feeding and quality, achieving a sensitivity of 83%⁶³.

TABLE 1.5 Recent applications of portable vibrational spectroscopy in the food science field.

Application	Technology /sampling	Instrument (Brand)	Wavelength selector	Source/detector	Chemometrics	Figures of merit	Range	Ref.
Trans fat quantification in butter and margarine	ATR-FTIR (1 bounce)	TruDefender (Thermo Scientific)	Michelson's interferometer	DTGS	SIMCA (C)	-	NA	64
					PLSR (P)	r_{VAL} = 0.99, SEP = 0.36%	0-5.3%	
	ATR-FTIR (5 bounces)	5500a (Agilent Technologies)	Michelson's interferometer	DTGS	SIMCA (C)	-	NA	
					PLSR (P)	r_{VAL} = 1.00, SEP = 0.12%	0-5.3%	
	FT-NIR	NeoSpectra Micro (Si-Ware Systems)	MEMS Michelson's interferometer	InGaAs	SIMCA (C)	-	NA	
					PLSR (P)	r_{VAL} = 1.00, SEP = 0.12%	0-5.3%	
Acrylamide quantification in frozen in frozen French fries	ATR-FTIR	4500a (Agilent Technologies)	Michelson's interferometer	DTGS	PLSR (P)	r_{VAL} = 0.94, SEP = 55.1 µg/kg	57.9-783.4 µg/kg	60
Classification and quality evaluation (harvesting time) of black truffle	NIR	MicroNIR 1700 (VIAVI)	LVF	Tungsten	LDA (C)	SENS = 100%	NA	55
				InGaAs array	PLSR (P)	R^2_{VAL}= 0.75, SEP = 3.46 days	0-26 days	
	NIR	Enterprise Sensor (Tellspec)	LVF	Tungsten -halogen	LDA (C)	SENS = 100%	NA	
				InGaAs	PLSR (P)	R^2_{VAL}= 0.72, SEP = 3.71 days	0-26 days	
	NIR	SCiO (Consumer Physics)	Bandpass filter (?)	LED	LDA (C)	SENS = 100%	NA	
				Diode array	PLSR (P)	R^2_{VAL}= 0.64, SEP = 4.48 days	0-26 days	
Control of coriander oil adulteration with canola oil	NIR	DLP NIRScan Nano (Texas Instruments)	Diffraction grating*	Tungsten -halogen	KNN (C)	SENS = 97%, SPEC = 86%	NA	65
				InGaAs	PLSR (P)	R^2_{VAL}= 0.94, SEP = 8.34%	0-90%	

TABLE 1.5 Recent applications of portable vibrational spectroscopy in the food science field (*continued*).

Application	Technology /sampling	Instrument (Brand)	Wavelength selector	Source/detector	Chemometrics	Figures of merit	Range	Ref.
Ripening monitoring (fat in dry matter) of Turkish white cheese	FT-NIR	Neospectra module (Si-Ware Systems	MEMS Michelson's interferometer	InGaAs	SIMCA (C)	SENS = 100%, SPEC = 100%	NA	66
					PLSR (P)	r_{VAL} = 0.90, SEP = 0.34%	19.4-21.8%	
	ATR-FTIR	4500a (Agilent Technologies)	Michelson's interferometer	DTGS	SIMCA (C)	SENS = 100%, SPEC = 100%	NA	
					PLSR (P)	r_{VAL} = 0.97, SEP = 0.19%	19.3-21.8%	
	Raman	Progeny (Rigaku)		Nd:YAG laser (1064 nm)	SIMCA (C)	SENS = 100%, SPEC = 100%	NA	
Estimation of quail egg freshness (quality index (%))	NIR	DLP NIRScan Nano (Texas Instruments)	Diffraction grating*	Tungsten -halogen	PLSR (P)	R^2_{VAL}= 0.72, SEP = 11.02%	10-110% (approx.)	67
					SVMR (P)	R^2_{VAL}= 0.82, SEP = 9.26%		
				InGaAs	PLS-DA (C)	SENS = 89.4%, SPEC = 80.7%	NA	
					SVMC (C)	SENS = 93.8%, SPEC = 71.4%	NA	
Classification of organic and conventional cocoa beans	NIR	Enterprise Sensor (Tellspec)		Tungsten -halogen	RF (C)	SENS = 96.7%, SPEC = 100%	NA	68
					KNN (C)	SENS = 95.0%, SPEC = 90.6%	NA	
				InGaAs	LDA (C)	SENS = 100%, SPEC = 96.2%	NA	
					PLS-DA (C)	SENS = 100%, SPEC = 100%	NA	
Discrimination of ochratoxin A producing and non-producing molds	NIR	MicroNIR OnSite (VIAVI)	LVF	Tungsten InGaAs array	SVM-DA (C)	SENS = 85%, SPEC = 86%	NA	57

TABLE 1.5 Recent applications of portable vibrational spectroscopy in the food science field (*continued*).

Application	Technology /sampling	Instrument (Brand)	Wavelength selector	Source/ detector	Chemometrics	Figures of merit	Range	Ref.
Evaluation of olive oil adulteration (with sunflower and soybean oils)	Raman	CBEx hand-held Raman (Snowy Range)		(785 nm)	PLSR (P)	R^2_{VAL}= 0.99, SEP = 4.23%	10-90%	62
		CBEx hand-held Raman (Snowy Range)		(1064 nm)	PLSR (P)	R^2_{VAL}= 0.99, SEP = 3.71%		
		Resolve (Agilent Technologies)		(830 nm)	PLSR (P)	R^2_{VAL}= 1.00, SEP = 1.98%		
		MIRA DS (Metrohm)		Laser diode (785 nm)	PLSR (P)	R^2_{VAL}= 0.99, SEP = 1.98%	0-100%	69
Assessment of the quality of essential oil (benzyl alcohol (%) in Rosemary oil)	Raman	Smart Raman XI (CloudMinds)			PLS-DA (C)	NA	NA	70
					RF (C)	SENS= 97%	NA	
Cuttlefish speciation (Northeastern Atlantic Ocean)	NIR	PoliSPECNIR (ITPhotonics)			SVM (C)	SENS = 96%, SPEC = 88%	NA	71

TABLE 1.5 Recent applications of portable vibrational spectroscopy in the food science field (*continued*).

Application	Technology/sampling	Instrument (Brand)	Wavelength selector	Source/detector	Chemometrics	Figures of merit	Range	Ref.
Study of beef composition [protein (%)] and quality characteristics in the abattoir	NIR	MicroNIR 1700 (VIAVI)	LVF	Tungsten Diode array	PLSR (P)	R^2_{VAL}= 0.42, SEP = 0.52%	20.5-24.2%	59
Quality grade classification of dry-cured ham	Raman	Raman BWS415 (B&W Tek)			LDA + KNN (C)	SENS= 83.3%	NA	63
Discrimination of bitter and sweet almonds	NIR	MicroNIR 1700 (VIAVI)	LVF	Tungsten InGaAs array	PLS-DA (C)	SENS = 87%, SPEC = 92% (CV)	NA	56
		Aurora (GraiNit)		Diode array	PLS-DA (C)	SENS= 97%, SPEC = 99%	NA	
Quantification of nitrate levels in vegetable leafs	ATR-FTIR (1 reflection)	TruDefender (Thermo Scientific)	Michelson's interferometer		ED-ELM	R^2_{VAL}= 0.93, SEP = 799.7 mg/kg	2000-10000 mg/kg (approx.)	61

As it can be seen in Table 1.5, no studies about microbiological applications have been conducted using portable spectroscopy devices during the last 9 months, and publications regarding this topic are scarcely found even before then. Therefore, there is a need of exploring the feasibility of such equipment for food microbiology applications (i.e., bacteria typing).

1.9 Objectives

The purpose of this dissertation was to evaluate the use of portable vibrational spectroscopy equipment combined with powerful chemometric algorithms for several food applications. Mid-infrared, near-infrared and Raman miniaturized devices were tested in monitoring a food chemical reaction, for the authentication of a novel food ingredient, in food packaging quality control, and for food microbiology applications. This main goal was accomplished based on the following nine specific objectives:

- To monitor Maillard reaction in protein-polysaccharide (sodium caseinate and gum Arabic or sodium carboxymethyl cellulose) mixtures using a portable ATR-FTIR spectrometer combined with SIMCA
- To develop a rapid method for studying the Maillard reaction of protein-polysaccharide mixtures without the need of using toxic and harmful reagents
- To build up chemometric models for edible insect powder authentication of *Tenebrio molitor*, *Alphitobius diaperinus, Gryllodes sigillatus, Acheta domesticus*, and *Locusta migratoria* species
- To evaluate how the rearing conditions affect to the chemical composition of edible insect powders
- To determine the presence of surface treatments applied to six still and sparkling wine cork stopper varieties (natural, agglomerated, micro-agglomerated and a mixture of agglomerated and micro-agglomerated) by means of hand-held NIR and Raman spectrometers using SIMCA
- To predict the extraction force for still wine cork stoppers using NIR data and PLSR
- To obtain new insights on the mechanism of resistance to different essential oil constituents (carvacrol, citral, and limonene oxide) in *E. coli* MG1655 evolved mutants using vibrational spectroscopy equipment (portable ATR-FTIR and Raman) and PCA

- To discriminate wine-making yeast strains (6 *Saccharomyces cerervisiae* strains and 4 non-*Saccharomyces*) by means of a portable ATR-FTIR spectrometer combined with PCA and evaluate the biochemical differences among them
- To create a pairwise SIMCA model for classifying wine-making *S. cerevisiae* and non-*Saccharomyces* yeasts

1.10 References

1. Raspor, P., Jevsnik Mokjca & Ambrozic, M. Food Safety Systems. in *Food Safety: Basic Concepts, Recent Issues, and Future Challenges* (eds. Selamat, J. & Zafar Iqbal, S.) 3–31 (Springer International Publishing AG Switzerland, 2016). doi:10.1007/978-3-319-39253-0.
2. Wallace, C. A., Sperber, W. H. & Mortimore, S. E. Origin and Evolution of the Modern System of Food Safety Management: HACCP and Prerequisite Programmes. in *Food Safety for the 21st Century: Managing HACCP and Food Safety Throughout the Global Supply Chain* (eds. Wallace, C. A., Sperber, W. H. & Mortimore, S. E.) 3–14 (John Wiley & Sons, Inc., 2018). doi:10.1002/9781119053569.ch1.
3. Stevens, S. K. Tracing the Food Safety Laws and Regulations Governing Traceability: A Brief History of Food Safety and Traceability Regulation. in *Food Traceability: From Binders to Blockchain* (eds. McEntire, J. & Kennedy, A. W.) 13–26 (Springer Nature Switzerland AG, 2019). doi:10.1007/978-3-030-10902-8.
4. Varzakas, T. ISO 22000, HACCP, and Other Management Tools for Implementation of Food Safety-Traceability Case Studies. in *Handbook of Food Processing: Food Safety, Quality, and Manufacturing Processes* (eds. Tzia, C. & Varzakas, T.) 105–140 (Taylor & Francis Group, LLC, 2016).
5. Medina, S., Perestrelo, R., Silva, P., Pereira, J. A. M. & Câmara, J. S. Current trends and recent advances on food authenticity technologies and chemometric approaches. *Trends in Food Science and Technology* 85, 163–176 (2019).
6. Spink, J. W. Introduction (Part 1 of 2): Food Fraud Definitions and Scope. in *Food Fraud Prevention: Introduction, Implementation, and Management* 1–43 (Springer Science+Business Media, LLC, 2019). doi:10.1007/978-1-4939-9621-6.
7. Reid, L. M., O'Donnell, C. P. & Downey, G. Recent technological advances for the determination of food authenticity. *Trends in Food Science and Technology* 17, 344–353 (2006).
8. Esteki, M., Shahsavari, Z. & Simal-Gandara, J. Gas Chromatographic Fingerprinting Coupled to Chemometrics for Food Authentication. *Food Reviews International* 36, 384–427 (2020).

9. Cho, Y.-J. Introduction. in *Emerging Technologies for Food Quality and Food Safety Evaluation* (ed. Cho, Y.-J.) 1–4 (Taylor & Francis Group, LLC, 2011). doi:10.1201/b10710.
10. El-Mesery, H. S., Mao, H. & Abomohra, A. E. F. Applications of non-destructive technologies for agricultural and food products quality inspection. *Sensors (Switzerland)* 19, (2019).
11. Medina, S., Pereira, J. A., Silva, P., Perestrelo, R. & Câmara, J. S. Food fingerprints – A valuable tool to monitor food authenticity and safety. *Food Chemistry* 278, 144–162 (2019).
12. Fiorino, G. M. *et al.* Overview on Untargeted Methods to Combat Food Frauds: A Focus on Fishery Products. *Journal of Food Quality* 2018, (2018).
13. Hussain, N., Sun, D. W. & Pu, H. Classical and emerging non-destructive technologies for safety and quality evaluation of cereals: A review of recent applications. *Trends in Food Science and Technology* 91, 598–608 (2019).
14. Medina, S., Perestrelo, R., Silva, P., Pereira, J. A. M. & Câmara, J. S. Current trends and recent advances on food authenticity technologies and chemometric approaches. *Trends in Food Science and Technology* 85, 163–176 (2019).
15. Su, W. H. & Sun, D. W. Fourier Transform Infrared and Raman and Hyperspectral Imaging Techniques for Quality Determinations of Powdery Foods: A Review. *Comprehensive Reviews in Food Science and Food Safety* 17, 104–122 (2018).
16. Craig, A. P., Franca, A. S. & Irudayaraj, J. Vibrational spectroscopy for food quality and safety screening. in *High Throughput Screening for Food Safety Assessment: Biosensor Technologies, Hyperspectral Imaging and Practical Applications* (eds. Bhunia, A. K., Kim, M. S. & Taitt, C. R.) 165–194 (Elsevier Ltd, 2015). doi:10.1016/B978-0-85709-801-6.00007-1.
17. Lohumi, S., Lee, S., Lee, H. & Cho, B. K. A review of vibrational spectroscopic techniques for the detection of food authenticity and adulteration. *Trends in Food Science and Technology* 46, 85–98 (2015).
18. Penner, M. H. Basic Principles of Spectroscopy. in *Food Analysis* (ed. Nielsen, S. S.) 79–88 (Springer International Publishing, 2017). doi:10.1007/978-3-319-45776-5_6.

19. Dufour, É. Principles of Infrared Spectroscopy. in *Infrared Spectroscopy for Food Quality Analysis and Control* (ed. Sun, D. W.) 3–27 (Academic Press/Elsevier, 2009). doi:10.1016/B978-0-12-374136-3.X0001-6.
20. Skoog, D. A., Holler, F. J. & Crouch, S. R. An Introduction to Spectrometric Methods. in *Principles of Instrumental Analysis* 120–147 (Coverage Learning, 2018).
21. Rodriguez-Saona, L., Ayvaz, H. & Wehling, R. L. Infrared and Raman Spectroscopy. in *Food Analysis* (ed. Nielsen, S. S.) 107–128 (Springer International Publishing, 2017). doi:10.1007/978-3-319-45776-5.
22. Skoog, D. A., Holler, F. J. & Crouch, S. R. An Introduction to Infrared Spectrometry. in *Principles of Instrumental Analysis* 389–411 (Cengage Learning, 2018).
23. Ozaki, Y. & Morisawa, Y. Principles and Characteristics of NIR Spectroscopy. in *Near-Infrared Spectroscopy: Theory, Spectral Analysis, Instrumentation, and Applications* (eds. Ozaki, Y., Huck, C., Tshuchikawa, S. & Balling Engelsen, S.) 11–36 (Springer Nature Singapore Pte Ltd., 2021). doi:10.1007/978-981-15-8648-4.
24. Smith, E. & Dent, G. Introduction, Basic Theory and Principles. in *Modern Raman Spectroscopy: A Practical Approach* 1–19 (John Wiley & Sons Ltd, 2019).
25. Candoğan, K., Evrim Altuntas, G. & İğci, N. Authentication and Quality Assessment of Meat Products by Fourier-Transform Infrared (FTIR) Spectroscopy. *Food Engineering Reviews* 13, 66–91 (2021).
26. Dahm, K. D. & Dahm, D. J. Theoretical Models of Light Scattering and Absorption. in *Near-Infrared Spectroscopy: Theory, Spectral Analysis, Instrumentation, and Applications* (eds. Ozaki, Y., Huck, C., Tsuchikawa, S. & Balling Engelsen, S.) 37–60 (Springer Nature Singapore Pte Ltd., 2021). doi:10.1007/978-981-15-8648-4.
27. Ozaki, Y. & Huck, C. Introduction. in *Near-Infrared Spectroscopy: Theory, Spectral Analysis, and Applications* (eds. Ozaki, Y., Huck, C., Tsuchikawa, S. & Balling Engelsen, S.) 3–10 (Springer Nature Singapore Pte Ltd., 2021). doi:10.1007/978-981-15-8648-4.
28. Lin, M., Rasco, B. A., Cavinato, A. G. & Al-Holy, M. Infrared (IR) Spectroscopy: Near-Infrared Spectroscopy and Mid-Infrared Spectroscopy. in *Infrared Spectroscopy for Food Quality Analysis and Control* (ed. Sun, D. W.) 119–143 (Academic Press/Elsevier, 2009).

29. Alvarez-Ordóñez, A. & Prieto, M. Technical and Methodological Aspects of Fourier Transform Infrared Spectroscopy in Food Microbiology Research. in *Fourier Transform Infrared spectroscopy in Food Microbiology* (eds. Hartel, R. W., Clark, J. P., Rodriguez-Lazaro, D. & Topping, D.) 1–18 (Springer, 2012). doi:10.1007/978-1-4614-3813-7.
30. Kamal, M. & Karoui, R. Analytical methods coupled with chemometric tools for determining the authenticity and detecting the adulteration of dairy products: A review. *Trends in Food Science and Technology* 46, 27–48 (2015).
31. Li, Q. *et al.* Application of Fourier transform infrared spectroscopy for the quality and safety analysis of fats and oils: A review. *Critical Reviews in Food Science and Nutrition* 59, 3597–3611 (2019).
32. Skoog, D. A., Holler, F. J. & Crouch, S. R. Applications of Infrared Spectrometry. in *Principles of Instrumental Analysis* 412–436 (Cengage Learning, 2018).
33. Subramanian, A. & Rodriguez-Saona, L. Fourier Transform Infrared (FTIR) Spectroscopy. in *Infrared Spectroscopy for Food Quality Analysis and Control* (ed. Sun, D. W.) 145–178 (Academic Press/Elsevier, 2009).
34. Huck, C. W. New Trend in Instrumentation of NIR Spectroscopy-Miniaturization. in *Near-Infrared Spectroscopy* (eds. Ozaki, Y., Huck, C., Tsuchikawa, S. & Balling Engelsen, S.) 193–210 (Springer Nature Singapore Pte Ltd., 2021). doi:10.1007/978-981-15-8648-4.
35. Ikehata, A. NIR Optics and Measurement Methods. in *Near-Infrared Spectroscopy* (eds. Ozaki, Y., Huck, C., Tsuchikawa, S. & Balling Engelsen, S.) 211–234 (Springer Nature Singapore Pte Ltd., 2021). doi:10.1007/978-981-15-8648-4.
36. Skoog, D. A., Holler, F. J. & Crouch, S. R. Components of Optical Instruments. in *Principles of Instrumental Analysis* 148–195 (Cengage learning, 2018).
37. Skoog, D. A., Holler, F. J. & Crouch, S. r. Raman Spectroscopy. in *Principles of Instrumental Analysis* 437–452 (Cengage Learning, 2018).
38. Li, Y. S. & Church, J. S. Raman spectroscopy in the analysis of food and pharmaceutical nanomaterials. *Journal of Food and Drug Analysis* 22, 29–48 (2014).

39. Craig, A. P., Franca, A. S. & Irudayaraj, J. Surface-enhanced Raman spectroscopy applied to food safety. *Annual Review of Food Science and Technology* 4, 369–380 (2013).
40. Esteki, M. *et al.* A review on the application of chromatographic methods, coupled to chemometrics, for food authentication. *Food Control* 93, 165–182 (2018).
41. Prieto, N., Pawluczyk, O., Dugan, M. E. R. & Aalhus, J. L. A Review of the Principles and Applications of Near-Infrared Spectroscopy to Characterize Meat, Fat, and Meat Products. *Applied Spectroscopy* 71, 1403–1426 (2017).
42. Callao, M. P. & Ruisánchez, I. An overview of multivariate qualitative methods for food fraud detection. *Food Control* 86, 283–293 (2018).
43. Borràs, E. *et al.* Data fusion methodologies for food and beverage authentication and quality assessment - A review. *Analytica Chimica Acta* 891, 1–14 (2015).
44. Li Vigni, M., Durante, C. & Cocchi, M. Exploratory Data Analysis. in *Chemometrics in Food Chemistry* (ed. Marini, F.) vol. 28 55–126 (Elsevier B.V., 2013).
45. Principal Components. in *Handbook of Chemometrics and Qualimetrics: Part A* (eds. Massart, D. L. et al.) vol. 20 519–556 (Elsevier Science B.V., 1998).
46. Analysis of Measurement Tables. in *Handbook of Chemometrics and Qualimetrics: Part B* (eds. Vandeginste, B. G. M. et al.) vol. 20 57–86 (Elsevier Science B.V., 1998).
47. Bevilacqua, M. *et al.* Classification and Class-Modelling. in *Chemometrics in Food Chemistry* (ed. Marini, F.) vol. 28 171–233 (Elsevier Science B.V., 2013).
48. Supervised Pattern Recognition. in *Handbook of Chemometrics and Qualimetrics: Part B* (eds. Vandeginste, B. G. M. et al.) vol. 20 207–241 (Elsevier Science B.V., 1998).
49. Westad, F., Bevilacqua, M. & Marini, F. Regression. in *Chemometrics in Food Chemistry* vol. 28 127–170 (Elsevier B.V., 2013).
50. Multivariate calibration. in *Handbook of Chemometrics and Qualimetrics: Part B* (eds. Vandeginste, B. G. M. et al.) vol. 20 349–381 (Elsevier Science B.V., 1998).

51. Olivieri, A. C. The Partial Least-Squares Model. in *Introduction to Multivariate Calibration* 103–121 (Springer Nature Switzerland AG, 2018). doi:10.1007/978-3-319-97097-4.
52. Riedl, J., Esslinger, S. & Fauhl-Hassek, C. Review of validation and reporting of non-targeted fingerprinting approaches for food authentication. *Analytica Chimica Acta* 885, 17–32 (2015).
53. Rodriguez-Saona, L., Aykas, D. P., Borba, K. R. & Urtubia, A. Miniaturization of optical sensors and their potential for high-throughput screening of foods. *Current Opinion in Food Science* 31, 136–150 (2020).
54. Clarivate Analytics. Web of Science. https://access.clarivate.com/login?app=wos&alternative=true&shibShireURL=https:%2F%2Fwww.webofknowledge.com%2F%3Fauth%3DShibboleth&shibReturnURL=https:%2F%2Fwww.webofknowledge.com%2F&roaming=true.
55. Kappacher, C. *et al.* Portable vs. Benchtop NIR-Sensor Technology for Classification and Quality Evaluation of Black Truffle. *Molecules* 27, (2022).
56. Torres, I., Sánchez, M. T., Vega-Castellote, M. & Pérez-Marín, D. Fraud detection in batches of sweet almonds by portable near-infrared spectral devices. *Foods* 10, (2021).
57. Cebrián, E., Núñez, F., Rodríguez, M., Grassi, S. & González-Mohino, A. Potential of near infrared spectroscopy as a rapid method to discriminate ota and non-ota-producing mould species in a dry-cured ham model system. *Toxins (Basel)* 13, (2021).
58. Dashti, A. *et al.* The feasibility of two hand-held spectrometers for meat speciation combined with chemometric methods and its application for halal certification. *Foods* 11, (2022).
59. Patel, N., Toledo-Alvarado, H. & Bittante, G. Performance of different portable and hand-held near-infrared spectrometers for predicting beef composition and quality characteristics in the abattoir without meat sampling. *Meat Science* 178, (2021).
60. Aykas, D. P. *et al.* Screening of Acrylamide of Par-Fried Frozen French Fries Using Portable FT-IR Spectroscopy. *Molecules* 27, (2022).
61. Ma, F., Du, C., Zheng, S. & Du, Y. In Situ Monitoring of Nitrate Content in Leafy Vegetables Using Attenuated Total Reflectance –

Fourier-Transform Mid-infrared Spectroscopy Coupled with Machine Learning Algorithm. *Food Analytical Methods* 14, 2237–2248 (2021).

62. Varnasseri, M. *et al.* Portable through bottle sors for the authentication of extra virgin olive oil. *Applied Sciences (Switzerland)* 11, (2021).
63. Martín-Gómez, A. *et al.* Portable raman spectrometer as a screening tool for characterization of iberian dry-cured ham. *Foods* 10, (2021).
64. Salas-Valerio, W. F. *et al.* In-field screening of trans-fat levels using mid- and near-infrared spectrometers for butters and margarines commercialized in the Peruvian market. *LWT* 157, (2022).
65. Kaufmann, K. C., Sampaio, K. A., García-Martín, J. F. & Barbin, D. F. Identification of coriander oil adulteration using a portable NIR spectrometer. *Food Control* 132, (2022).
66. Yaman, H., Aykas, D. P., Jiménez-Flores, R. & Rodriguez-Saona, L. E. Monitoring the ripening attributes of Turkish white cheese using miniaturized vibrational spectrometers. *Journal of Dairy Science* 105, 40–55 (2022).
67. Brasil, Y. L., Cruz-Tirado, J. P. & Barbin, D. F. Fast online estimation of quail eggs freshness using portable NIR spectrometer and machine learning. *Food Control* 131, (2022).
68. Anyidoho, E. K., Teye, E. & Agbemafle, R. Differentiation of Organic Cocoa Beans and Conventional Ones by Using Hand-held NIR Spectroscopy and Multivariate Classification Techniques. *International Journal of Food Science* 2021, (2021).
69. Barros, I. H. A. S. *et al.* Use of portable Raman spectroscopy in the quality control of extra virgin olive oil and adulterated compound oils. *Vib Spectrosc* 116, (2021).
70. Lebanov, L. & Paull, B. Smartphone-based hand-held Raman spectrometer and machine learning for essential oil quality evaluation. *Analytical Methods* 13, 4055–4062 (2021).
71. Currò, S. *et al.* Fast and green method to control frauds of geographical origin in traded cuttlefish using a portable infrared reflective instrument. *Foods* 10, (2021).

UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

ATR-FTIR spectroscopy combined with multivariate analysis, a novel approach to monitor Maillard reaction

This chapter is published as: Mellado-Carretero, J., Kaade, W., Ferrando, M., Güell, C. and de Lamo-Castellví, S., 2019. Attenuated total reflectance fourier transform midinfrared spectroscopy combined with multivariate analysis, a novel approach to monitor maillard reaction. Journal of Food Science, 84, p.2777-2784.

ABSTRACT

The aim of this work was to study the potential of using infrared spectroscopy and chemometrics to monitor Maillard reaction. Sodium caseinate (NaCAS) and gum Arabic (GA) or sodium carboxymethyl cellulose (CMC) powders were mixed at 1:1, spray-dried, and incubated at 60 °C and 76% of relative humidity from 0 to 72 h. Sample infrared spectra were collected, and browning degree, conjugation efficiency, and stabilization properties of the conjugates were analysed by spectrophotometry, fluorescence spectroscopy, turbidity, and zeta potential measurements. Pairwise soft independent modelling of class analogy (SIMCA) models showed significant chemical differences among NaCAS–GA mixtures incubated for 0 (Control) and 16 h, attributed to functional groups linked to different Maillard reaction products such as Schiff's base and pyridine compounds. Infrared spectroscopy combined with SIMCA is a powerful tool to monitor the formation of protein–polysaccharide conjugates by Maillard reaction.

2.1 Introduction

Proteins are important food ingredients because they can change the texture, taste, and appearance of foods. Therefore, proteins are frequently used in food industry, and due to their amphipathic behaviour, as emulsifying agents[1–3]. However, proteins show instability under certain conditions of temperature, pH, or the presence of organic solvents and high concentration of electrolytes. Thus, researchers have enhanced protein stability and solubility under these unfavourable conditions by using polysaccharides[1,3,4]. Polysaccharides are high molecular weight polymers with high stability but low interfacial activity[5]. Then, by combining properties from both proteins and polysaccharides though conjugation, it is possible to obtain novel emulsifiers with higher stability[1,3,4].

There are physical, chemical, or enzymatic treatments to produce protein-polysaccharide conjugates. Nonetheless, food manufacturers try to avoid using conjugates produced by chemical agents that may be potentially toxic and harmful for consumers. Therefore, a suitable approach to obtain these conjugates is by Maillard reaction[1,5,6].

The use of protein-polysaccharide conjugates seems promising, but it is necessary to monitor their production during Maillard reaction. The degree of protein glycation is commonly determined by measuring the formation of different Maillard reaction products (MRPs). Several methods such as high-performance liquid chromatography or mass spectrometry are suitable for analysing MRPs. However, these techniques are time consuming and not useful for quantification due to the lack of pure standards[7,8]. Moreover, sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE) or dinitrofluorobenzene (DNFB) and o-phtaldialdehyde colorimetric methods involve the use of toxic and harmful reactants.

Some researchers have used Fourier transform mid-infrared (FTIR) spectroscopy to evaluate the structural changes and interactions of polysaccharides and proteins during Maillard reaction[9–11]. Temenouga et al. (2016) found changes in the absorbance values of IR bands related to secondary structure of proteins and carbohydrates[12]. Moreover, Yang et al. (2015) confirmed the presence of different MRPs such as Schiff's bases and β-diketones. Compared with other classical methods of reference (e.g., chemical methods), FTIR spectroscopy is an environmentally friendly technique.

Nevertheless, the high complexity of the systems studied restricts the use of FTIR to qualitative analysis.

One of the major advances in IR spectroscopy is the use of multivariate analysis [e.g., principal component analysis (PCA) or soft independent modelling of class analogy (SIMCA)], which allow researchers to obtain valuable information from big sets of data[13]. The present work aims to determine if a portable FTIR spectrometer in combination with SIMCA can be used to rapidly and safely monitor sodium caseinate-gum Arabic (NaCAS-GA) and sodium caseinate-carboxymethyl cellulose (NaCAS-CMC) conjugates produced by Maillard reaction. Classical methods of reference such as browning degree, conjugation efficiency, and stability assessments of conjugates were performed by spectrophotometry, fluorescence spectroscopy, and turbidity and ζ-potential measurements, respectively.

2.2 Materials and methods

2.2.1 Materials

Sodium caseinate (NaCAS) from bovine milk, gum Arabic (GA) from acacia tree branched polysaccharide, sodium carboxymethyl cellulose (CMC) with a molecular weight of ≈250 kDa and a degree of substitution of 0.7, quinine hemisulphate salt monohydrate (90%), and fluorescamine (98%) were acquired from Sigma-Aldrich (Spain). Sodium phosphate dibasic heptahydrate (≥98%), sodium phosphate monobasic monohydrate, potassium bromide (≥99%), sodium tetraborate (98%) anhydrous, Nα-BOC-L-Lysine, and hydrochloric acid (0.1 N) were obtained from Fischer Scientific (Spain). Milli-Q water was acquired from a Millipore Simplicity 185 water system (Merck, Spain). All other reagents were analytical grade unless otherwise stated.

TABLE 2.1 Spray-drying operational parameters

Operational parameters	NaCAS	NaCAS-CMC	NaCAS-GA
Pump rate (%)	35	25	30
Airflow (m^3/h)	35	35	35
Aspirator rate (%)	100	100	100
Inlet temperature (°C)	150	150	150
Outlet temperature (°C)	72	80	69
Nozzle clean mode	1	1	1

2.2.2 Preparation of Maillard conjugates

NaCAS 5 wt%, CMC 3.75 wt%, and GA 5 wt% were prepared in Milli-Q water. Solutions were stirred at 400 rpm for 2 h and kept at 4 °C overnight. NaCAS solutions were mixed with CMC (1:1) and GA (1:1), and stirred at 400 rpm for 30 min at room temperature. The pH of each NaCAS-polysaccharide mixtures was adjusted to 7.0 using HCl 0.1 N. Then, a Mini Spray Dryer B-290 (Büchi, Spain) equipped with a two-fluid nozzle of 0.7 mm (operational parameters are shown in Table 2.1) was required for spray drying the samples. Maillard reaction was carried out as previously explained by Davidov-Pardo et al. (2015)[14]. Spray-dried samples were allowed to react at 0, 8, 16, 24, 36, 40, 48, 56, 64, and 72 h at 60 °C in a desiccator filled with a supersaturated solution of potassium bromide to achieve a constant relative humidity of 76 to 79%. After the reaction, samples were allowed to cool down to room temperature, crushed using a mortar and a pestle, and then kept in a desiccator under nitrogen atmosphere until their analysis. Conjugates were prepared in triplicate.

2.2.3 Turbidity and ζ-potential measurements

NaCAS and NaCAS-GA samples were diluted at 1:99 wt and NaCAS-CMC samples at 1:114 wt in 15 mM phosphate buffer (pH 6.8 to 7.0). Turbidity was measured by static multiple light scattering using a Turbiscan Lab Expert (Formulaction, Spain). Backscattering (BS) profiles were obtained to assess stability changes between samples at different reaction times. Laser Doppler-microelectrophoresis was also used to measure the ζ-potential of protein-polysaccharide mixtures by means of Zetasizer Nano ZS, equipped with a 633 nm He-Ne laser with a power of 4 mW (Malvern Instruments, Spain)[14]. All reported ζ-potential values, calculated using the Smoluchowski approximation, were an average of three measurements.

2.2.4 Browning degree

Browning degree produced by Maillard reaction of NaCAS, NaCAS-GA (diluted 1:99 wt in 15 mM phosphate buffer, pH 6.8 to 7.0), and NaCAS-CMC (diluted 1:114 wt) samples was determined by measuring their absorbance in a UV/Vis spectrophotometer (DR5000, Hach Lange, Spain) at 420 nm[14].

2.2.5 Conjugation efficiency

Reduction of free amino groups was evaluated though the fluorescamine assay based on the method of Yaylayan et al. (1992)[15]. NaCAS, NaCAS-GA, and NaCAS-CMC mixtures were diluted in 50 mM borate buffer, pH 8.5 (1 mg of protein/mL). An aliquot of 0.2 mL of diluted mixtures was transferred to vials and brought to 5.0 mL with 50 mM borate buffer. A volume of 1.0 mL of fluorescamine solution in acetone (15 mg/100 mL) was added to the vials while they were strongly shaken on a vortex type mixer. Fluorescence data were obtained at room temperature on a Cary Eclipse fluorescence spectrophotometer (Varian, Spain) equipped with a Xenon flash lamp and a thermostable cell holder. The excitation and emission wavelengths were 390 and 475 nm, respectively, with excitation and emission slit widths of 5 nm. The equipment was calibrated using a quinine sulphate solution in 0.05 M sulfuric acid (5 mg/250 mL). All measurements were done in triplicate. Calibration curves were built using Nα-BOC-L-Lysine and regressed with a power function ($y = ax^b$). Yaylayan et al. (1992) previously chose Nα-BOC-L-Lysine as the most precise external standard for estimating the amount of free amino groups in a commercial glycated BSA protein. Conjugation efficiency was determined according to the equation given by Davidov-Pardo et al. (2015)[14]:

$$\text{conjugation efficiency } (\%) = \left(1 - \frac{\text{no. amino groups after conjugation}}{\text{no. amino groups prior conjugation}}\right) \quad (1.1)$$

2.2.6 ATR-FTIR spectroscopy

Dried samples were dissolved in Milli-Q water (1:20 wt) and stored overnight at 4 °C. Aliquots of 0.6 µL of completely hydrated samples were placed onto the sample stage of a Cary 630 portable spectrometer (Agilent Technologies, Spain) interfaced with a single reflection attenuated total reflectance (ATR) diamond crystal accessory. Water contribution was reduced by vacuum drying prior to IR analysis. A background scan was deducted from every sample scan to prevent the effect of environmental changes on sample's spectra. Spectra were collected from 4000 to 800 cm^{-1} with 8 cm^{-1} resolution, from the average of 128 scans. Data acquisition was totally controlled using MicroLab PC software (Agilent Technologies).

2.2.7 Multivariate analysis: SIMCA

Multivariate analysis and data pre-processing were performed using chemometric software (Pirouette version 4.0, Infometrix, US). Data were mean-centred, processed to their second derivative using a second-order polynomial Savitzky-Golay filter of 25 points, and vector-length normalized. SIMCA was used to build pair-wise models based on PCA models for each class (0 h vs. 8, 16, 24, 32, 40, 48, 56, 64, and 72 h) to describe the variations occurred during the reaction. Sample residuals and Mahalanobis distances were used for outlier determination. SIMCA models were interpreted according to class distances, misclassifications, and discriminating power plots.

2.2.8 Statistical analysis

Analysis of variance (ANOVA) was performed using SPS statistics version 25.0 (IBM, US). Tukey post hoc tests were conducted to evaluate significant differences among sample means in turbidity, ζ-potential, absorbance, and conjugation efficiency data. Significance level was set at $P < 0.05$.

2.3 Results and discussion

2.3.1 Stability study of Maillard conjugates

BS data of protein-polysaccharide mixtures are shown in Figure 2.1 A. For NaCAS (control), no significant differences were detected throughout the reaction time applied (up to 72 h), showing a BS value close to 5%. On the other hand, NaCAS-GA mixtures exhibited a significant decrease of BS, being more pronounced between the 32 and 48 h to finally stabilized at 56 h of reaction (from 7.5% to 4%). According to the Mie scattering model of light for particles smaller than the incident wavelength ($\lambda = 880$ nm), a decrease of backscattered light flux entails a decrease of particle mean diameter[16]. When Maillard reaction arises, sugars covalently linked to proteins make proteins impossible to interact by steric hindrance and electrostatic interactions, forcing them to separate. Figure 2.1 A illustrates this behaviour for NaCAS-GA mixtures, with decreasing BS values and thus particle mean diameter during the reaction.

Electrophoretic light scattering is commonly used to assess colloid stability in solution based on its ζ-potential value. The ζ-potential is the

electric potential measured at the boundary surrounding the layer in which the less-attached ions form a stable entity (i.e., slipping plane, the first layer with strongly attached ions is called the Stern layer). Particles with ζ-potentials greater than 30 mV in absolute value are usually considered stable[17]. Figure 2.1 B summarizes the ζ-potential values obtained for NaCAS, NaCAS-GA, and NaCAS-CMC mixtures at different reaction times. An overall significant decrease in ζ-potential was observed for all samples, with NaCAS-GA mixtures showing the highest decrease (from −9 mV at 0 h to −25 mV at 56 h) and NaCAS the lowest decrease (from −9 mV at 0 h to −13 mV at 72 h). During Maillard reaction, the covalent bond formed between carbonyl groups of sugars and amino groups of proteins cause a shift in their isoelectric point towards a lower pH value. Thus, the ζ-potential should decrease with higher reaction times, leading to a higher stability of proteins[6,18,19]. Both BS and ζ-potential results confirmed the existence of Maillard reaction in NaCAS-GA mixtures.

2.3.2 Study of the formation of advanced MRPs

A rapid method to monitor Maillard reaction is measuring the absorbed light of the brown-coloured products formed. Final stage products are active at 420 nm[5,20]. Absorbance results are plotted in Figure 2.1 C. NaCAS and NaCAS-CMC mixtures did not show changes in absorbance (0.8 and 0.5, respectively), suggesting that there was not production of brown-coloured compounds and, therefore, Maillard reaction. On the other hand, NaCAS-GA mixtures experienced a linear increase of absorbance until 64 h (i.e., from 0.5 at 0 h to 1.1 at 64 h), to then reach a plateau. Colour development during Maillard reaction can be divided into different stages. In the first stage, commonly designated as induction stage, no changes are detected in A_{420} since melanoidins are end products of Maillard reaction. After that, a linear increase in A_{420} is observed until its stabilization, maybe due to the saturation of brown colour or the depletion of the reactants (i.e., sugars and proteins)[21,22]. Even though no induction was identified, our findings confirmed the existence of Maillard reaction in NaCAS-GA mixtures.

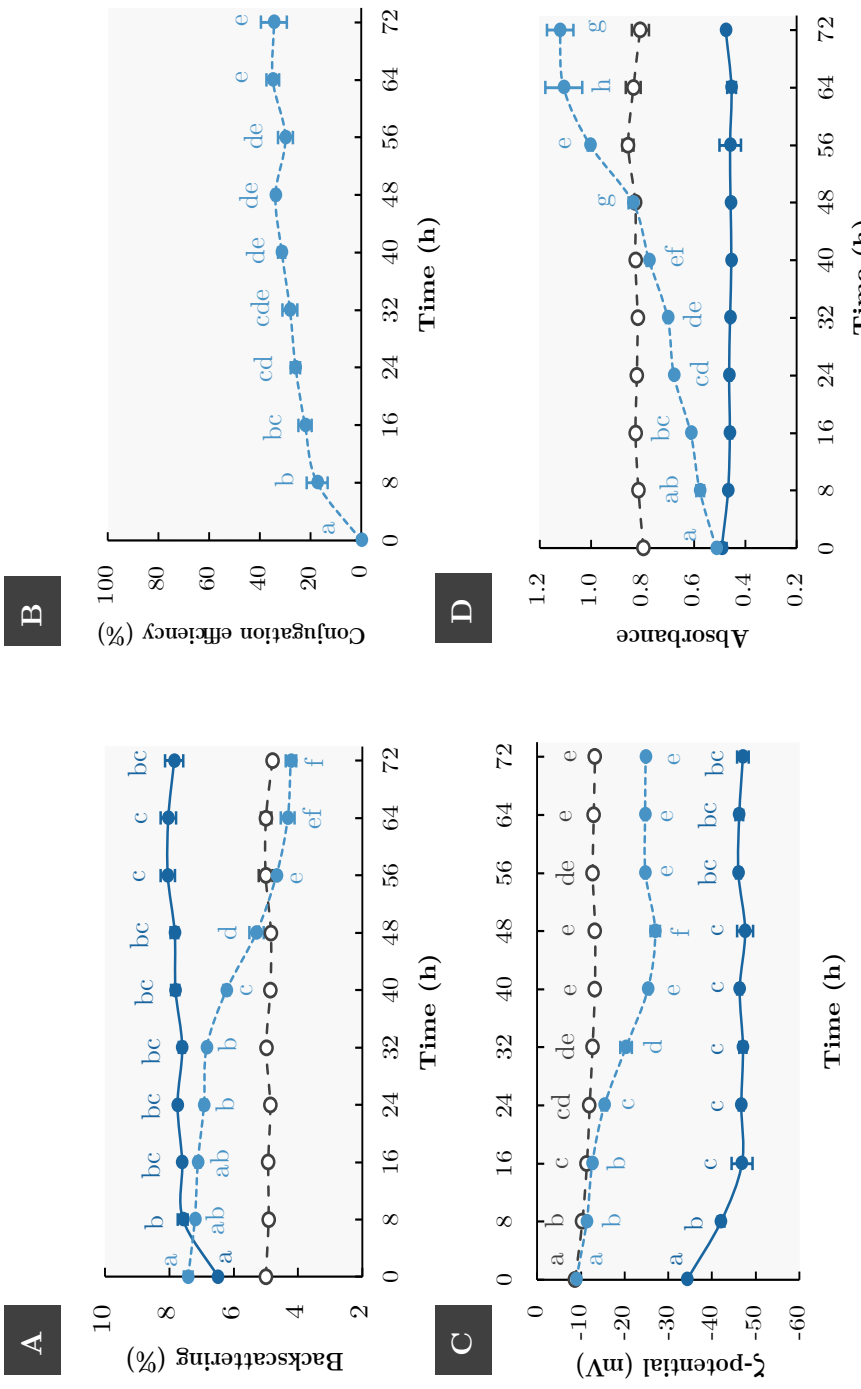


FIGURE 2.1 Backscattering (A), ζ-potential (B), conjugation efficiency (B), and absorbance at 420 nm (D) of NaCAS (white), NaCAS-CMC (dark blue), and NaCAS-GA (blue) mixtures at different reaction times (pH 6.8-7). Errors bars in the plot represent the standard deviation for three independent replicates. Means (n = 3) with different letters are significantly different (P > 0.95).

2.3.3 Conjugation efficiency

Fluorescence obtained from the reaction of fluorescamine with proteins is an indicator of their number of free amino groups[15]. Thus, with the appropriate standards, conjugation efficiency can be assessed during Maillard reaction by applying Equation 1[14]. Calibration curves obtained with Nα-BOC-L-Lysine were adjusted using a power regression, which gave the best fit (Figure 2.2). After linearizing the fluorescence data, the number of free amino groups was interpolated for NaCAS, NaCAS-CMC, and NaCAS-GA mixtures.

Conjugation efficiency increased for NaCAS-GA mixtures until reaching a constant level after 48 h. Davidov-Pardo et al. (2015) showed the same trend when producing NaCAS-dextran conjugates. In their study, conjugation efficiency was obtained using the o-phtaldialdehyde (OPA) assay, which reacts with free primary amines as fluorescamine but could be potentially toxic for the operator. Decrease in conjugation efficiency to a constant level could be attributed to a reduction of the number of free amino groups available to react with the fluorophore or the steric hindrance reached by the attachment of sugars to the protein[14]. No changes in fluorescence values were detected in NaCAS and NaCAS-CMC mixtures (data not shown). Considering that the Maillard reaction occurs between reducing sugars and free amino groups of proteins[20,23,24], the carbonyl end group available from CMC (a nonreducing sugar) would not be able to react with enough protein free amino groups to observe significant changes under the experimental conditions tested.

2.3.4 Maillard reaction monitoring by ATR-FTIR

Raw spectra of NaCAS, NaCAS-CMC, and NaCAS-GA mixtures at 0 h are shown in Figure 2.3. The tentative assignations of the IR bands observed are collected in Table 2.2. NaCAS spectrum displayed IR bands at 1636 and 1529 cm^{-1}, which could be related to amide I and amide II protein characteristic regions, respectively. IR bands at 1443, 1398, and 1241 cm^{-1} might be assigned to the amide II region of proteins as well. NaCAS-CMC spectrum had a broad band located at 1059 cm^{-1}, which is probably related to C-O stretching vibrations of carbohydrates. NaCAS-CMC mixtures also exhibited the amide I and II bands at 1636 and 1544 cm^{-1}, respectively[11,25–27].

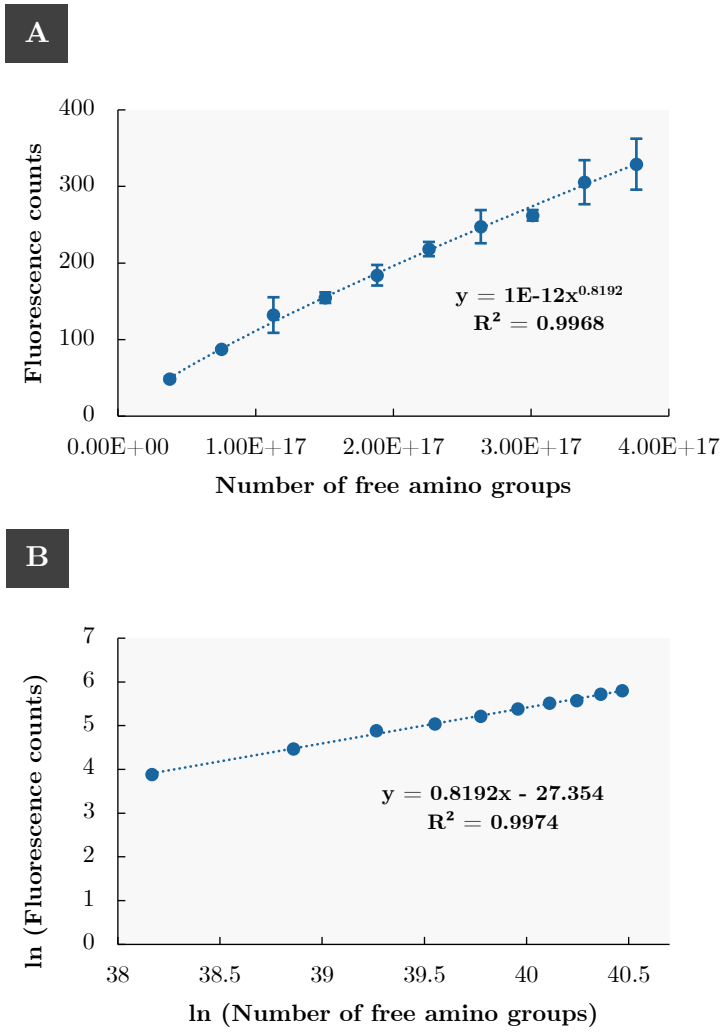


FIGURE 2.2 Calibration curve for Nα-BOC-L-Lysine standards expressed as means of n = 3 (A) and its linearisation (B).

Furthermore, the IR band at 1629 cm^{-1}, which was not present in the raw spectrum of NaCAS mixtures, could be linked to the electrostatic interaction between protein amino groups and carboxyl groups of CMC[28]. In NaCAS-GA spectrum, same bands related to sugars (at 1059 cm^{-1}) and proteins (at 1636 and 1544 cm^{-1}) were observed.

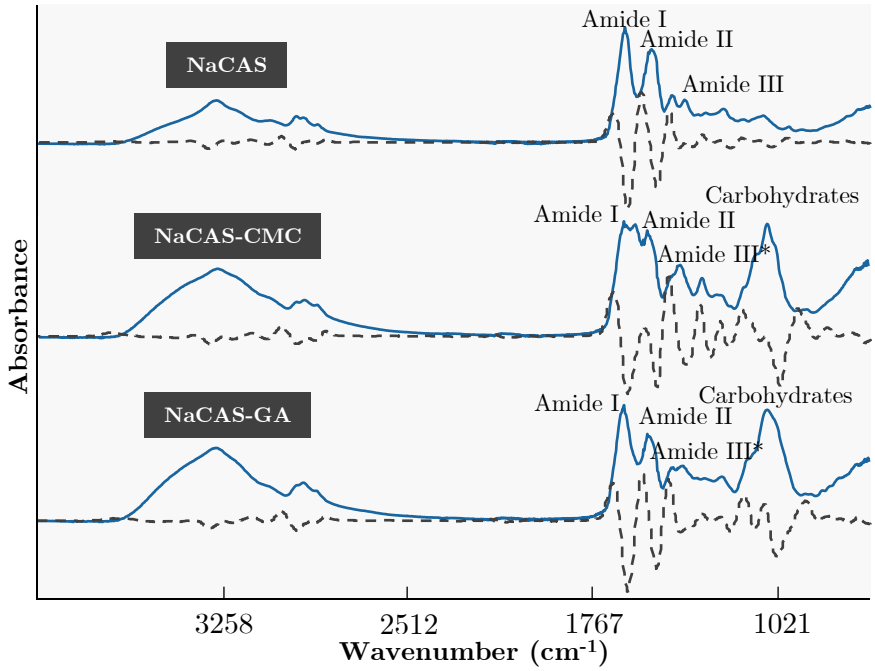


FIGURE 2.3 ATR-FTIR spectroscopy raw spectra of NaCAS, NaCAS-CMC, and NaCAS-GA mixtures (black), and their Savitzky-Golay second derivative (grey). *Some bands within the amide III region can be linked to carbohydrates (C-H and O-H bending vibrations).

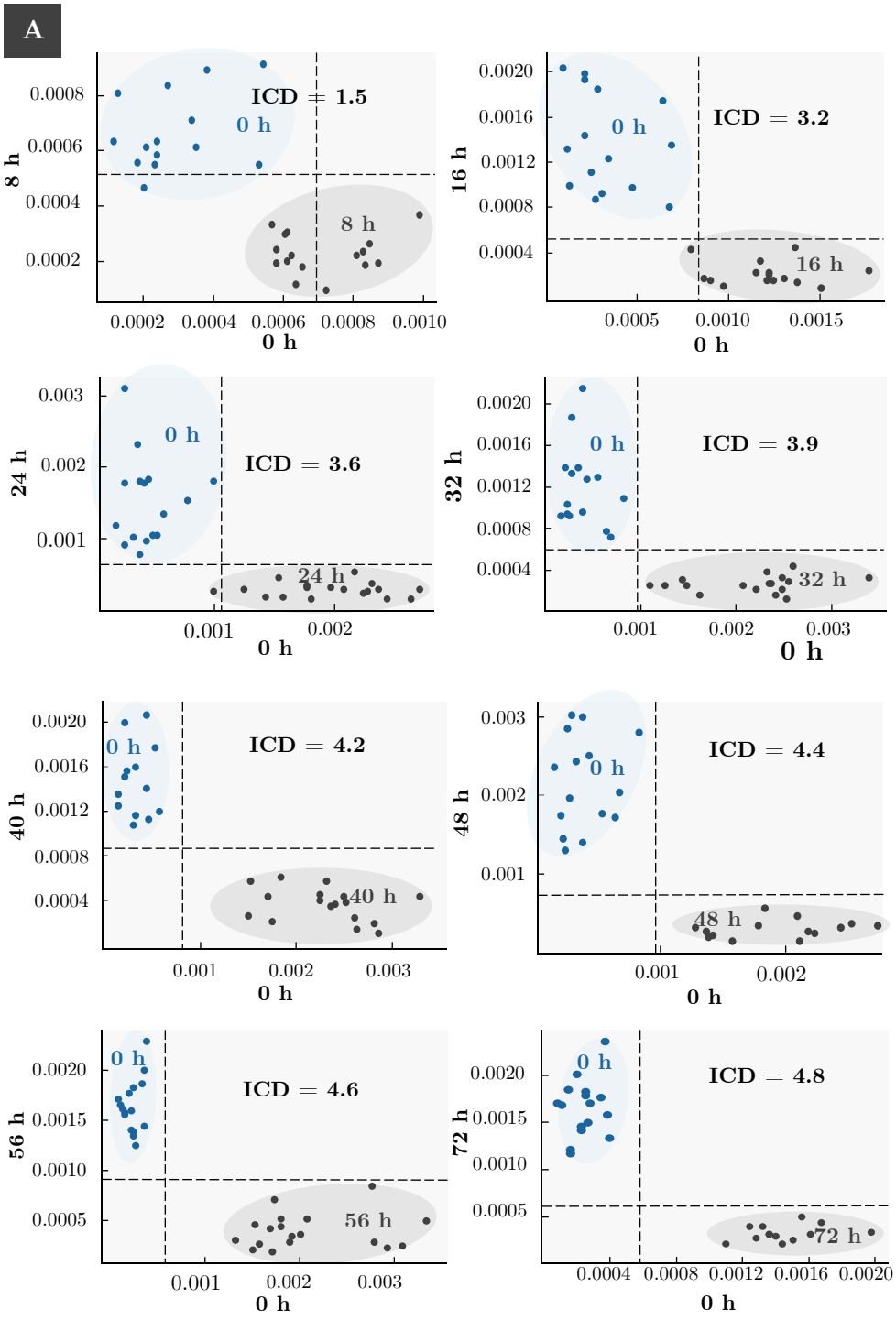
TABLE 2.2 Band assignments in raw ATR-FTIR spectra of NaCAS, NaCAS-GA, and NaCAS-CMC and SIMCA discriminating power plots of NaCAS-GA mixtures.

Wavenumber (cm^{-1})	Functional group assignment
1655	C=N stretching vibration of imine or enaminol groups of Schiff base
1636	Amide I region of proteins
1620	Amide I region of proteins after Maillard reaction (NaCAS-GA; downshifting)
1595 and 1565	C=O stretching vibration of carboxylate ion of MRPs (both), C=O stretching vibration of β-diketones, or ring stretching of pyridine derivatives (both)
1544 and 1529	Amide II region of proteins (NaCAS-CMC and NaCAS-GA, respectively)
1510	Amide II region of proteins after Maillard reaction (NaCAS-GA; downshifting
1455	O-H bending vibration of carbohydrates
1444, 1398, and 1241	Amide III region of proteins (NaCAS)
1413	-COO- stretching vibration of carbohydrates
1368	C-H bending vibration of carbohydrates
1059	C-O stretching vibration of carbohydrates

[a]Summarised from references: Estevez et al. (2019), Kosaraju et al. (2010), Pirestani et al. (2018), Socrates (2004), Su et al. (2010), and Yang et al. (2015)[21,25–29].

Two-class SIMCA models (0 h vs. 8, 16, 24, 32, 40, 48, 56, 64, and 72 h) were built up using pre-treated spectra (830 to 2000 cm^{-1}) of NaCAS, NaCAS-CMC, and NaCAS-GA mixtures. PCA factors and number of spectra rejected per class in NaCAS-GA, NaCAS-CMC, and NaCAS mixtures are displayed in Table 2.3, Table 2.4, and Table 2.5, respectively. One way to evaluate the distance among classes is by visualizing the Coomans plots. These plots are divided into four quadrants by two threshold lines, which represent the critical standard deviation of the residuals around the class models. Every data point (i.e., spectrum) falling under the threshold of a certain class is considered as a member of that class with a 95% confidence[30]. For the NaCAS-GA mixtures, SIMCA Coomans plots showed tight clustering and clear differentiation after 16 h of reaction (Figure 2.4 A). On the other hand, no differences were detected in NaCAS and NaCAS-CMC mixtures throughout the reaction under the conditions tested (Figure 2.4 B and Figure 2.4 C, respectively), considering that most of the spectra analysed belonged to both classes (bottom left quadrant). The interclass distance (ICD) value is a measure accounting for the degree of separation between classes. Classes with ICD values greater than 3.0 are considered statistically different[31]. In the NaCAS-GA mixtures, ICD values ranged between 1.5 and 4.8, overcoming the threshold value of 3.0 after 16 h of reaction (ICD = 3.2). However, ICD values of NaCAS and NaCAS-CMC mixtures ranged between 1.2 and 1.8 and 0.9 and 2.6, respectively, suggesting that no significant chemical changes had taken place during the experiments.

Wavenumbers in the spectral range analysed were represented against their power to discriminate and differentiate the classes. Discriminating power plots of NaCAS-GA mixtures at different reaction times are displayed in Figure 2.5. Different spectral wavenumbers explained the differences between NaCAS-GA samples at different reaction times compared to the 0 h ones (see Table 2.2). Several IR bands responsible for the discrimination among different samples have been previously linked to functional groups related to some MRPs. The IR band at 1655 cm^{-1} could be assigned to C=N stretching vibration linked to the formation of Schiff bases during the reaction[27,32]. IR bands at 1599 and 1565 cm^{-1} could be related to COO- stretching vibration, C=O stretching vibration of β-diketones, or ring stretching vibration of pyridine derivatives, respectively[27].


A
ICD = 1.5
0 h
8 h
8 h
0 h
ICD = 3.2
0 h
16 h
16 h
0 h
ICD = 3.6
0 h
24 h
24 h
0 h
ICD = 3.9
0 h
32 h
32 h
0 h
ICD = 4.2
0 h
40 h
40 h
0 h
ICD = 4.4
0 h
48 h
48 h
0 h
ICD = 4.6
0 h
56 h
56 h
0 h
ICD = 4.8
0 h
72 h
72 h
0 h

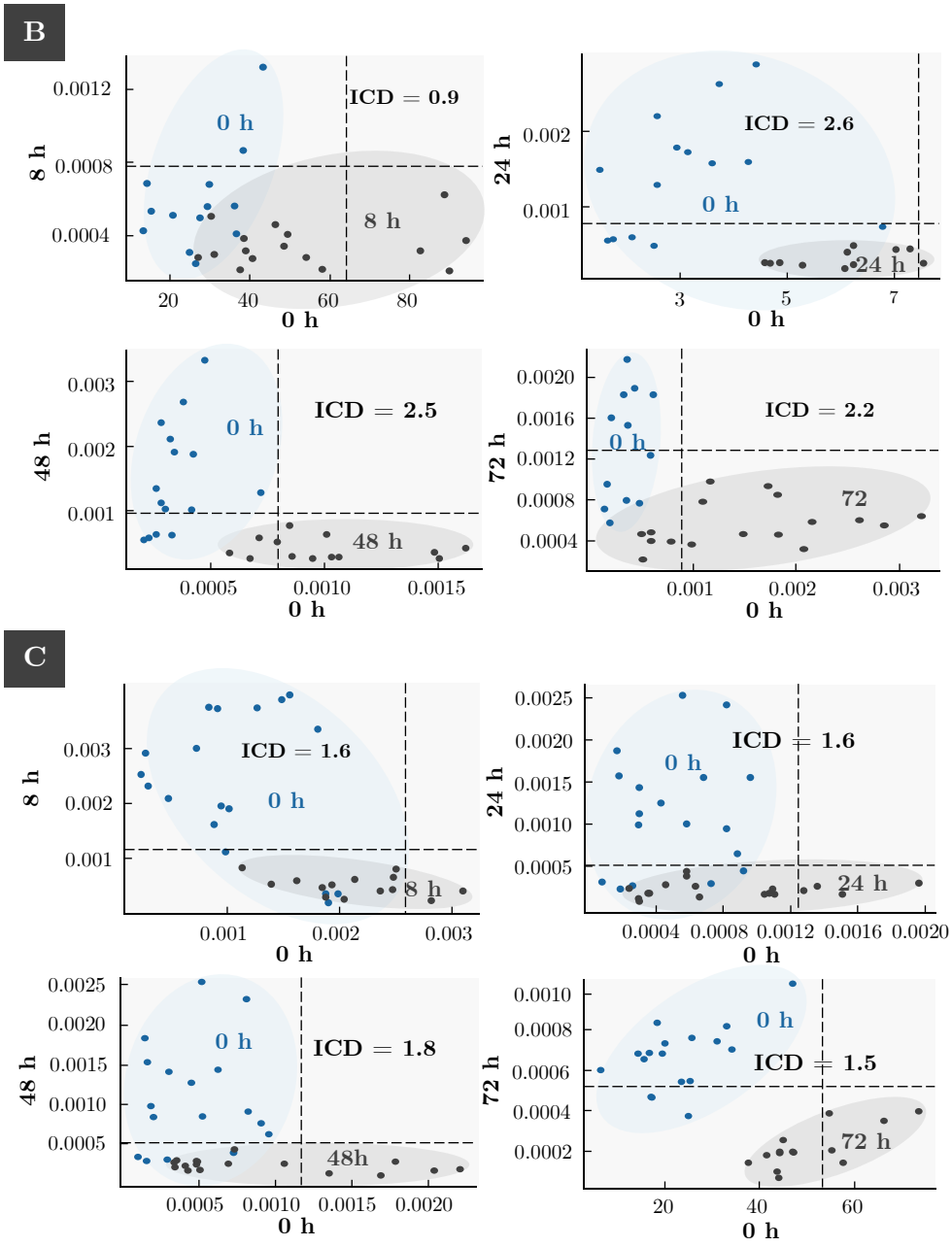


FIGURE 2.4 SIMCA Coomans plots of transformed (second derivative, 25-point window) ATR-FTIR spectra of NaCAS-GA (A), NaCAS-CMC (B), and NaCAS (C) at different reaction times. Samples in NW, SE, SW, and NE quadrants belong to 0 h, reaction time compared (8 to 72 h), both and none of them, respectively.

IR bands and at 1455 and 1368 cm^{-1} could be attributed to CH and OH bending vibrations of carbohydrates[26]. After 48 h of reaction, amide I and II downshifted by 20 cm^{-1} approximately. This trend has been reported by other authors. Kosaraju et al. (2010) found a downshift on the amide I region during Maillard reaction due to protein structural changes[29]. Furthermore, Pirestani et al. (2018) reported an upshifting for amide I, II, and III regions[25]. Therefore, IR bands at 1455 and 1368 cm^{-1} could be related to the shifting of amide III region. NaCAS-GA mixtures data suggest that Maillard reaction significantly happened after 16 h. At this reaction time, protein structure changes and potential formation of MRPs such as Schiff's bases, β-diketones, and pyridine derivatives were detected though the IR spectral models. Furthermore, ICD values obtained for NaCAS-GA mixtures proved that Maillard reaction reached a plateau after 56 h (Figure 2.6). This trend was also detected by browning degree and conjugation efficiency tests.

Table 2.3 Variances explained (as cumulative percentages) of each factor and number of spectra rejected per class (PCA) for NaCAS-CMC 2-class SIMCA models.

2-class SIMCA models	Factor 1 (%)	Factor 2 (%)	Factor 3[a] (%)	Number Outliers[b]
0 h	59.5	86.3	96.5	4
8 h	77.1	94.8	98.8	2
0 h	60.1	85.7	95.3	3
16 h	58.9	92.5	97.1	3
0 h	57.8	80.3	94.5	2
24 h	69.6	90.4	96.5	0
0 h	65.9	89.8	95.1	3
32h	72.2	93.3	98.1	2
0 h	64.9	91.4	95.9	5
40 h	63.4	81.3	93.3	1
0 h	65.9	89.8	95.1	3
48 h	71.2	85.3	96.3	4
0 h	75.8	97.3	99.0	2
56 h	65.5	82.1	95.9	3
0 h	84.9	98.1	99.3	1
64 h	62.9	93.1	97.9	1
0 h	75.8	97.3	99.0	2
72 h	81.0	97.0	99.1	5

[a]The number of factors per model were selected to obtain at least 95% of explained variance.
[b]Sample residuals and Mahalanobis distances were used for outlier determination.

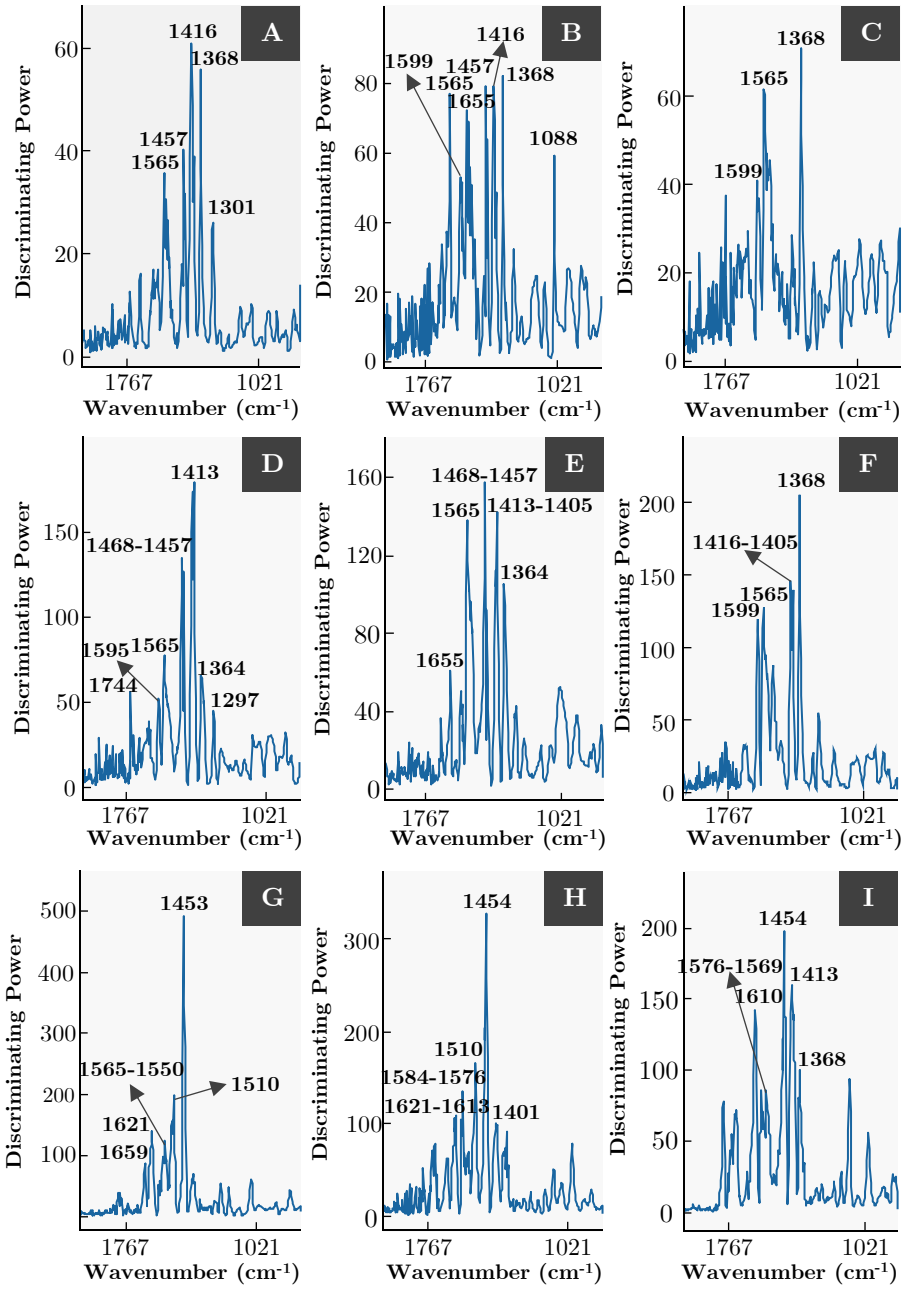


FIGURE 2.5 SIMCA discriminating power plots of transformed (second derivative, 25-point window) ATR-FTIR spectra of NaCAS-GA at different reaction times (A, B, C, D, E, F, G, H, and I for 8, 16, 24, 32, 40, 48, 56, 64, and 72 h, respectively).

TABLE 2.4 Variances explained (as cumulative percentages) of each factor and number of spectra rejected per class (PCA) for NaCAS 2-class SIMCA models.

2-class SIMCA models	Factor 1 (%)	Factor 2[a] (%)	No. of outliers[b]
0 h	96.9	—	0
8 h	95.7	—	4
0 h	96.9	99.9	0
24 h	81.7	98.5	0
0 h	97.3	99.4	1
48 h	68.6	99.3	1
0 h	92.9	98.4	1
72 h	94.3	98.6	3

[a]The number of factors per model were selected to obtain at least 95% of explained variance.
[b]Sample residuals and Mahalanobis distances were used for outlier determination.

TABLE 2.5 Variances explained (as cumulative percentages) of each factor and number of spectra rejected per class (PCA) for NaCAS-CMC 2-class SIMCA models.

2-class SIMCA models	Factor 1 (%)	Factor 2[a] (%)	No. of Outliers[b]
0 h	79.7	99.0	5
8 h	77.2	97.4	2
0 h	79.9	98.6	4
24 h	92.0	96.6	7
0 h	69.9	98.6	3
48 h	92.5	98.4	5
0 h	89.0	95.5	6
72 h	86.3	95.9	1

[a]The number of factors per model were selected to obtain at least 95% of explained variance.
[b]Sample residuals and Mahalanobis distances were used for outlier determination.

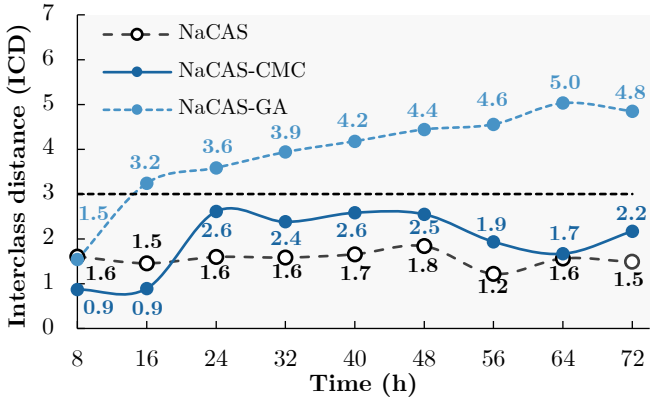


FIGURE 2.6 ICD values of NaCAS, NaCAS-CMC, and NaCAS-GA mixtures at different reaction times

2.4 Conclusions

ATR-FTIR combined with multivariate analysis has been proved to be a powerful tool to monitor the course of Maillard reaction rapidly and safely when producing two different conjugates, NaCAS-CMC and NaCAS-GA. Different functional groups linked to MRPs such as Schiff's base and pyridine derivatives were identified when two-classes SIMCA models were build up with IR data from NaCAS-GA mixtures. Changes in protein structure were also detected due to the covalent bond formed between amino and carbonyl groups of proteins and sugars, respectively. ICD values increased throughout the reaction, displaying the same trend detected with browning degree, conjugation efficiency, and ζ-potential tests. Moreover, turbidity measurements demonstrated the stabilizing effect of conjugates. On the other hand, even though NaCAS-CMC mixtures exhibited some changes in ζ-potential and BS values, no significant differences were detected in fluorescence and browning data. Two-classes SIMCA models of NaCAS-CMC mixtures confirmed this trend, with ICD values lower than 3.0. The results obtained suggested that NaCAS-CMC mixtures did not undergo Maillard reaction within the experimental conditions tested. Future research should focus on determining the amount of conjugate produced during Maillard reaction combining ATR-FTIR spectroscopy and multivariate regression modelling.

2.5 References

1. Akhtar, M. & Ding, R. Covalently cross-linked proteins & polysaccharides: Formation, characterisation and potential applications. *Current Opinion in Colloid and Interface Science* 28, 31–36 (2017).
2. Dickinson, E. Interfacial structure and stability of food emulsions as affected by protein-polysaccharide interactions. *Soft Matter* 4, 932–942 (2008).
3. Evans, M., Ratcliffe, I. & Williams, P. A. Emulsion stabilisation using polysaccharide-protein complexes. *Current Opinion in Colloid and Interface Science* 18, 272–282 (2013).
4. Dickinson, E. Colloids in Food: Ingredients, Structure, and Stability. *Annual Review of Food Science and Technology* 6, 211–233 (2015).
5. de Oliveira, F. C., Coimbra, J. S. dos R., de Oliveira, E. B., Zuñiga, A. D. G. & Rojas, E. E. G. Food Protein-polysaccharide Conjugates Obtained via the Maillard Reaction: A Review. *Critical Reviews in Food Science and Nutrition* 56, 1108–1125 (2016).
6. Liu, F., Wang, D., Xu, H., Sun, C. & Gao, Y. Physicochemical properties of β-carotene emulsions stabilized by chlorogenic acid-lactoferrin-glucose/polydextrose conjugates. *Food Chemistry* 196, 338–346 (2016).
7. Liu, F., Ma, C., Gao, Y. & McClements, D. J. Food-Grade Covalent Complexes and Their Application as Nutraceutical Delivery Systems: A Review. *Comprehensive Reviews in Food Science and Food Safety* 16, 76–95 (2017).
8. Silván, J. M., van de Lagemaat, J., Olano, A. & del Castillo, M. D. Analysis and biological properties of amino acid derivates formed by Maillard reaction in foods. *Journal of Pharmaceutical and Biomedical Analysis* 41, 1543–1551 (2006).
9. Geng, X. *et al.* Preparation and characterization of ovalbumin and carboxymethyl cellulose conjugates via glycosylation. *Food Hydrocolloids* 37, 86–92 (2014).
10. Liu, Q., Kong, B., Han, J., Sun, C. & Li, P. Structure and antioxidant activity of whey protein isolate conjugated with glucose via the Maillard reaction under dry-heating conditions. *Food Structure* 1, 145–154 (2014).

11. Su, J. F., Huang, Z., Yuan, X. Y., Wang, X. Y. & Li, M. Structure and properties of carboxymethyl cellulose/soy protein isolate blend edible films crosslinked by Maillard reactions. *Carbohydrate Polymers* 79, 145–153 (2010).
12. Temenouga, V. *et al.* Novel emulsifiers as products from internal Maillard reactions in okra hydrocolloid mucilage. *Food Hydrocolloids* 52, 972–981 (2016).
13. Karoui, R., Downey, G. & Blecker, C. Mid-infrared spectroscopy coupled with chemometrics: A tool for the analysis of intact food systems and the exploration of their molecular structure-quality relationships-A review. *Chemical Reviews* 110, 6144–6168 (2010).
14. Davidov-Pardo, G., Pérez-Ciordia, S., Marín-Arroyo, M. R. & McClements, D. J. Improving Resveratrol Bioaccessibility Using Biopolymer Nanoparticles and Complexes: Impact of Protein-Carbohydrate Maillard Conjugation. *Journal of Agricultural and Food Chemistry* 63, 3915–3923 (2015).
15. Yaylayan, V. A., Huyghues-Despointes, A. & Polydorides, A. A fluorescamine-based assay for the degree of glycation in bovine serum albumin. *Food Research International* 25, 269–275 (1992).
16. Bru, P. *et al.* Particle size and rapid stability analyses of concentrated dispersions: use of multiple light scattering technique. in *Particle Sizing and Characterization* (eds. Provder, T. & Texter, J.) vol. 881 45–60 (ACS Symposium Series, 2004).
17. Hunter, R. J., Ottewill, R. H. & Rowell, R. L. *Zeta Potential in Colloid Science: Principles and Applications* (Academic Press, 1981). doi:10.1016/0021-9797(82)90296-X.
18. Consoli, L. *et al.* Sodium caseinate-corn starch hydrolysates conjugates obtained through the Maillard reaction as stabilizing agents in resveratrol-loaded emulsions. *Food Hydrocolloids* 84, 458–472 (2018).
19. Wang, W. & Zhong, Q. Properties of whey protein-maltodextrin conjugates as impacted by powder acidity during the Maillard reaction. *Food Hydrocolloids* 38, 85–94 (2014).
20. Fayle, S. E. & Gerrard, J. A. *The Maillard Reaction.* (The Royal Society of Chemistry, 2007).
21. Liu, S. C., Yang, D. J., Jin, S. Y., Hsu, C. H. & Chen, S. L. Kinetics of color development, pH decreasing, and anti-oxidative activity reduction

of Maillard reaction in galactose/glycine model systems. *Food Chemistry* 108, 533–541 (2008).

22. Morales, F. J. & van Boekel, M. A. J. S. A study on advanced Maillard reaction in heated casein/sugar solutions: Colour formation. *International Dairy Journal* 8, 907–915 (1998).
23. Coultate, T. *Food: the chemistry of its components*. (The Royal Society of Chemistry, 2016).
24. Damodaran, S., Perkin, K. L. & Fennema, O. R. *Fennema's Food Chemistry*. (Taylor & Francis Group, LLC, 2007).
25. Pirestani, S., Nasirpour, A., Keramat, J., Desobry, S. & Jasniewski, J. Structural properties of canola protein isolate-gum Arabic Maillard conjugate in an aqueous model system. *Food Hydrocolloids* 79, 228–234 (2018).
26. Socrates, G. *Infrared and Raman characteristic group frequencies. Infrared and Raman characteristic group frequencies* (John Wiley & Sons, Inc., 2004). doi:10.1002/jrs.1238.
27. Yang, Y. *et al.* A soy protein-polysaccharides Maillard reaction product enhanced the physical stability of oil-in-water emulsions containing citral. *Food Hydrocolloids* 48, 155–164 (2015).
28. Estévez, M., Güell, C., de Lamo-Castellví, S. & Ferrando, M. Encapsulation of grape seed phenolic-rich extract within W/O/W emulsions stabilized with complexed biopolymers: Evaluation of their stability and release. *Food Chemistry* 272, 478–487 (2019).
29. Kosaraju, S. L., Weerakkody, R. & Augustin, M. A. Chitosan-glucose conjugates: Influence of extent of maillard reaction on antioxidant properties. *Journal of Agricultural and Food Chemistry* 58, 12449–12455 (2010).
30. Forina, M., Lanteri, S., Casolino, C. & Oliveri, P. Methods, Techniques, and Instruments. in *Handbook of food analysis* (eds. Nollet, L. M. L. & Toldrá, F.) 429–458 (CRC Press, 2015).
31. Dunn, W. J. & Wold, S. SIMCA Pattern Recognition and Classification. in *Chemometric Methods in Molecular Design* (ed. van de Waterbeemd, H.) 179–192 (VCH Publishers, Inc., 1995).
32. Su, W. H. & Sun, D. W. Fourier Transform Infrared and Raman and Hyperspectral Imaging Techniques for Quality Determinations of Powdery Foods: A Review. *Comprehensive Reviews in Food Science and Food Safety* 17, 104–122 (2018).

UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

Rapid discrimination and classification of edible insect powders using ATR-FTIR spectroscopy combined with multivariate analysis

This chapter is published as: Mellado-Carretero, J., García-Gutiérrez, N., Ferrando, M., Güell, C., García-Gonzalo, D., De Lamo-Castellví, S., 2019. Rapid discrimination and classification of edible insect powders using ATR-FTIR spectroscopy combined with multivariate analysis. Journal of Insects as Food and Feed, 6, p.141-148.

ABSTRACT

Insects are being proposed as an alternative way to ensure world food and feed security. Methods to determine edible insect powder origin and species will be needed for quality control purposes. Infrared spectroscopy has been extensively used in rapid chemical fingerprinting of food products. The present research explores a new approach to discriminate and classify commercial edible insect powders using attenuated total reflectance mid-infrared (ATR-FTIR) spectroscopy combined with multivariate analysis. Infrared spectra of seven commercial edible insect powders from different species (*T. molitor*, *A. diaperinus*, *G. sigillatus*, *A. domesticus*, and *L. migratoria*) and origins (the Netherlands and New Zealand) were collected to build up soft independent modelling of class analogy (SIMCA) models. SIMCA models clearly discriminated insects by species and origin linking their differences to lipids and chitin. SIMCA model performance was tested using five spectra of each class not used to build up the training set. 100% correct predictions were obtained for all the samples analysed except for one sample of *A. diaperinus*. Infrared spectroscopy coupled to multivariate analysis provided a powerful method for the assurance of insect powder authenticity.

3.1 Introduction

World population growth combined with an increasing demand for animal-derived products in both developed and developing countries requires finding other and more sustainable protein sources. Insects are being considered as an alternative way to ensure food and feed security[1–4]. Entomophagy, namely the habit of eating insects, arachnids, and arthropods, has been practiced by humans for centuries in several countries of Africa, Asia, Australia, and Latin America, with more than 2000 insects considered edible[1–7]. Edible insects are good sources of proteins, polyunsaturated fatty acids, vitamins, and minerals such as iron, calcium, and zinc. Even though insects can also produce greenhouse emissions and ammonia, most commercially reared edible insect species have lower environmental impact than conventional livestock[1,3,4,7,8]. In terms of productivity, insects have a higher growth rate and fecundity as well as higher feed conversion efficiencies since they are poikilothermic (i.e., they do not invest energy to maintain a constant body temperature). Additionally, insects can be reared on organic side streams and therefore they are able to transform waste into high value food and feed resources[3–5,8–10].

Even though the use of insects seems encouraging, its incorporation to the market could be slowed down by the long approval process stated by the current legislation and the poor acceptance of entomophagy in Western countries[1]. Concerning the low acceptance of entomophagy, innovative ways to incorporate insects into the human diet have been proposed. Edible insects can be processed into more palatable forms by drying and grinding them, thus obtaining a powder that can be used as an ingredient to increase the protein content of several food products[4].

Accordingly, methods are required to discriminate and classify edible insect powder's origin and species to prevent fraud and adulterations. Ulrich et al. (2017) have already discriminated several whole insects (*T. molitor*, *A. diaperinus*, *A. domesticus*, and *L. migratoria*) commercially available for human consumption through protein profiling using mass spectrometry (MS)[11]. Moreover, Köppel et al. (2019) developed a multiplex real-time PCR method for the detection of insect DNA and determination of contents of *T. molitor*, *L. migratoria*, and *A. domesticus* in several food matrixes[12] Even though these techniques have been gaining attention from both scientific and

industrial communities because of their accuracy and reliability, both present several drawbacks. In one hand, DNA-based molecular methods need to be cost-effective, as well as they require highly trained operators and set-up optimization, which is not easily achieved[13,14]. On the other hand, MS implies a high initial cost due to the purchase of the equipment and a fractionation of the sample must be done before collecting the spectra in some applications[15,16].

Infrared spectroscopy combined with multivariate analysis has also been applied for authentication of fruit juices, edible oils and dairy products, among others[17]. This technique is fast, cheap, non-destructive, robust, simple to use, and a minimum training for the operator is needed. However, infrared spectroscopy is highly affected by changes in sample's preparation[18], which is easily avoided when the procedure is properly standardized. As far as we know, no previous research has investigated the potential of using attenuated total reflectance Fourier transform mid-infrared spectroscopy (ATR-FTIR) combined with supervised pattern recognition techniques to discriminate and classify commercial edible insect powders by species and origin. The objective of the present work was to obtain mid-infrared spectroscopy profiles from commercial edible insect powders and to develop multivariate classification models to rapidly discriminate and classify edible insect powder species and origin.

3.2 Materials and methods

3.2.1 Edible insect samples

T. molitor (mealworm), *A. diaperinus* (Buffalo worm), *G. sigillatus* (banded cricket), *A. domesticus* (house cricket), and *L. migratoria* (grasshopper) powders were purchased from Kreca Ento-Food BV (Netherlands), DeliBugs (Netherlands), and Eat Crawlers (New Zealand) (more details are shown in Table 3.1).

3.2.2 ATR-FTIR spectroscopy

Insect powders were mixed thoroughly, and 4 mg were taken randomly (per each spectrum collected) and placed onto the sample stage of a portable spectrometer Cary 630 (Agilent Technologies, Spain), equipped with a single bounce ATR diamond crystal accessory and a deuterated triglycine sulphate (DTGS) detector. A pressure clamp was used to ensure optimal contact

between samples and the diamond crystal. A background scan was taken before every sample scan to prevent the environment from interfering the measurements. Spectra were obtained from 4000 to 800 cm^{-1} with 8 cm^{-1} of resolution. Data acquisition was controlled using MicroLab PC software (Agilent Technologies). Two different batches of each edible insect powder were analysed during three different days obtaining five spectra per day and batch (30 spectra in total per edible insect powder).

TABLE 3.1 Commercial edible insect powders tested in this research.

Insect powder	Supplier	Manufacturer	Amount per package[c] (g)	Sample tag
Mealworm	Kreca Ento-Food	1[a]	100	Mealworm
Buffalo	DeliBugs	1	50	Buffalo worm A
	Kreca Ento-Food	1	100	Buffalo worm B
Grasshopper	Kreca Ento-Food	1	100	Grasshopper
Banded cricket	Kreca Ento-Food	1	100	Banded cricket
House cricket	Eat Crawlers	2[b]	50	House cricket A
	Kreca Ento-Food	1	100	House cricket B

[a]Netherlands (origin).
[b]New Zealand (origin).
[c]Two different batches analysed per product tested.

3.2.3 Multivariate analysis

Multivariate analysis and data pre-processing were performed using a chemometric software (Pirouette version 4.5, Infometrix, US). Second derivative transformation (13-point window second order polynomial-fit Savitzky-Golay function) and multiplicative scatter correction (MSC) were performed on mean-centred data. A principal component analysis (PCA)-based pattern recognition method, soft independent modelling of class analogy (SIMCA), was performed to obtain classification models to discriminate the chemical differences among insect powders and for predicting unknown samples[19]. Sample residuals and Mahalanobis distances were used for outlier diagnostics[20]. Three SIMCA models were built up to assess differences between edible insect powders, a 7-class SIMCA model with all samples tested (whole model), a 3-class SIMCA model with Mealworm, Buffalo worm A, and Buffalo worm B (worm model), and a 3-class SIMCA model with Banded cricket, House cricket A and House cricket B (cricket model). SIMCA models were interpreted according to class projection, interclass distances, and discriminating power. Predicting ability of SIMCA models was tested using

five spectra of each class not used to build up the training set (25 spectra per class)[21].

3.3 Results and discussion

3.3.1 MIR raw spectra

Mid-infrared spectroscopy can provide information on the composition of complex chemical mixtures by studying the IR bands arising from their functional groups, whose assignment is known in most of the cases[22]. Raw IR spectra of several edible insect powders are shown in Figure 3.1. Different spectral regions linked to several components could be identified from insect powder spectra. In all samples, one broad band was observed at 3000-3500 cm^{-1} caused by the H-bonded O-H stretching of chitin, other polysaccharides, and residual water. The spectral region at 2800-3000 cm^{-1} represented the C-H stretching of methyl groups of lipids and chitin[23–25]. The IR band located at 1740 cm^{-1} was assigned to C=O stretching of esters of lipids. IR bands from amide I, II, and III regions linked to proteins or chitin were observed as well at 1700-1600, 1600-1500, and 1300-1200 cm^{-1}, respectively. Infrared spectra also showed an IR band around 1200-1000 cm^{-1} mainly attributed to carbohydrates[26–28].

3.3.2 Discrimination and classification of commercial edible insect powders by ATR-FTIR combined with SIMCA

Class projection plots, namely PCAs of the entire training set, were generated through SIMCA to visualize class separation among samples and spectra reproducibility (Figure 3.2). The ellipses represent the regions in which samples belonging to a certain class fall into with a 95% of confidence[29]. Each data point in the plot represents one insect powder spectrum. Tight clustering and good separation were obtained among powders of different insect species for all three SIMCA models. Spectra collected from insect powders made from same insect species but produced by different manufacturers (i.e., House cricket A and House cricket B, see Table 3.1) displayed clearly distinguishable clusters as well (Figure 3.2 B). On the other hand, Buffalo worm powders (i.e., Buffalo A and Buffalo B), which were elaborated by the same manufacturer, showed overlapping clusters and poor differentiation (Figure 3.2 C).

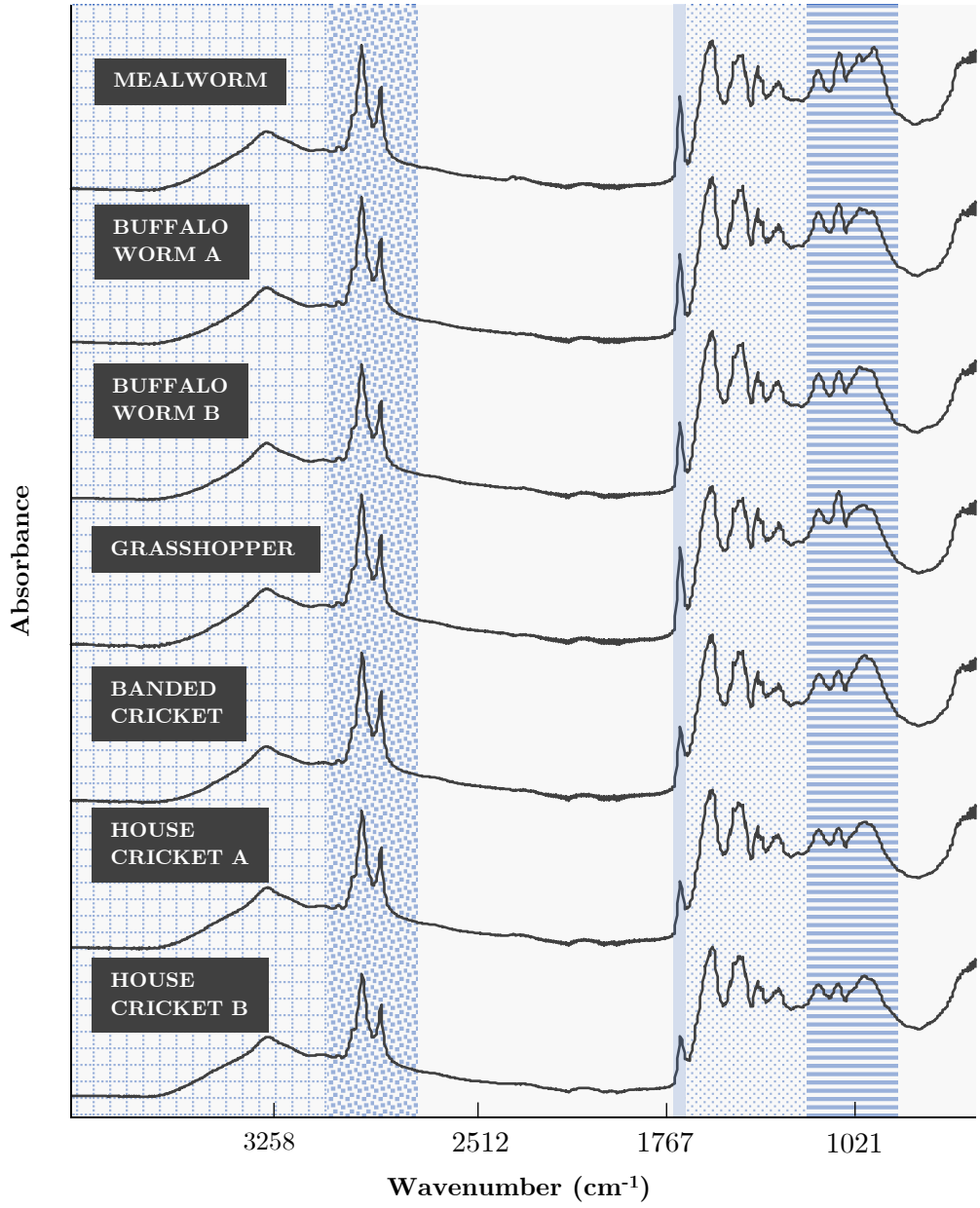


FIGURE 3.1 ATR-FTIR raw spectra of seven commercial edible insect powders.

Another way to evaluate class separation (i.e., distance between two classes) is the interclass distance (ICD) value. In SIMCA, PCA models are

built up for all the classes in the training set. Number of factors and outliers removed per class for the whole model are shown in Table 3.2. After building the PCA class models, the residuals are computed by fitting the objects of every class of the training set to the PCA model of each class. The overall standard deviations of the residuals (i.e., Euclidean distances) are used for calculating the ICD, which is a ratio of interclass to intraclass distance. ICD values close to 0 indicate that the two classes compared are almost identical and values larger than 1 show differentiation[19]. As a rule of thumb, ICDs above 3.0 are considered significant to discriminate two clusters of samples as distinct classes[30]. ICDs among insect powders for the whole model are displayed in Table 3.3 and are comparable to the clustering patterns obtained in class projection plots (Figure 3.2 A). ICD values above 3.0 were achieved when comparing powders made of different insect species, but lower values (ICD $<$ 3.0) were obtained when comparing products elaborated with the same insect species (Buffalo worm A-Buffalo worm B and House cricket A-House cricket B clusters). Despite that, the ICD for house cricket clusters was higher than the one obtained for Buffalo worm clusters. The nutritional and bioactive profiles of edible insects are affected by factors such as habitat, feed, sex, stage, and preparation/processing methods applied prior to consumption, among others[7,10]. Since house cricket powders were produced by different manufacturers (see Table 3.1), our data confirmed that product origin plays an important role in insect powder differentiation.

Table 3.2 Variances gathered by each factor and number of outliers removed per class (disjoint PCA model) of 7-class SIMCA whole model.

Class	Factor 1 (%)	Factor 2 (%)	Factor 3 (%)	Factor 4 (%)	Factor 5 (%)[a]	No. of outliers[b]
Mealworm	78.7	91.9	94.8	96.9	97.6	3
Buffalo worm A	76.5	84.9	90.3	94.1	95.5	0
Buffalo worm B	47.7	73.8	87.9	91.5	94.5	2
Grasshopper	83.1	94.7	96.9	97.8	98.4	2
Banded cricket	61.4	83.5	90.0	92.6	94.5	8
House cricket A	44.4	75.0	83.2	89.1	91.8	3
House cricket B	62.7	82.4	87.9	91.3	93.8	1

[a]The number of factors selected per model was to obtain at least 90% of variance. This criterion was established according to the number of optimal factors obtained through an F-test with 95% confidence on the reduced eigenvalues (i.e., latent factors' variances)[31].
[b]Sample residuals and Mahalanobis distances were used for outlier determination.

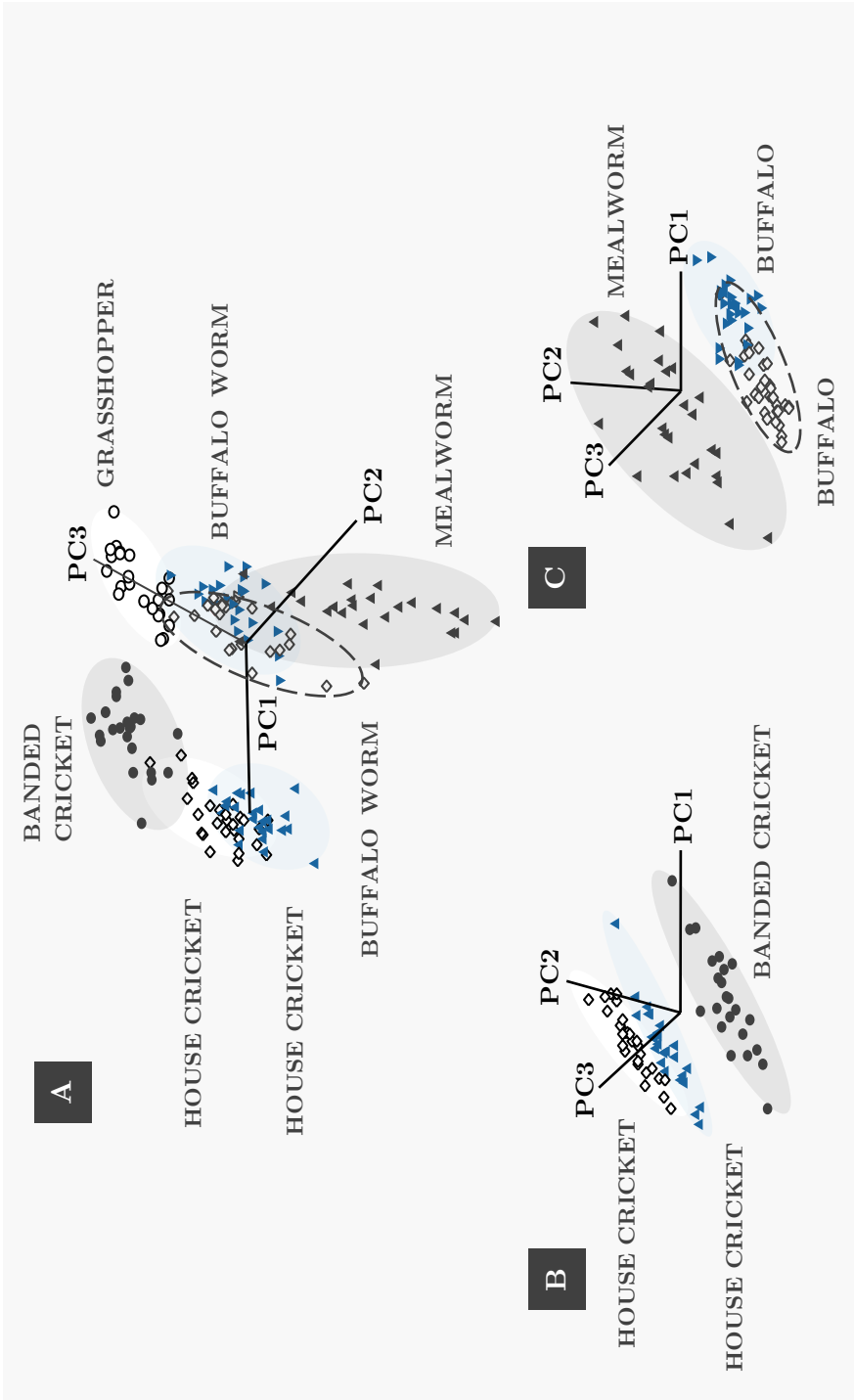


FIGURE 3.2 SIMCA class projection plots of transformed (second derivative, 13-points window) ATR-FTIR spectra(4,000-800 cm^{-1} region) of whole model (A), cricket model (B), and worm model (C).

TABLE 3.3 ICD values of transformed (second derivative, 13-point window) ATR-FTIR spectra (4000-800 cm^{-1} region) of all commercial edible insect powder (7-class SIMCA whole model)

Class	Meal-worm	Buffalo worm A	Buffalo worm B	Grass-hopper	Banded cricket	House cricket A	House cricket B
Meal-worm	0.0	—	—	—	—	—	—
Buffalo worm A	3.3	0.0	—	—	—	—	—
Buffalo worm B	3.1	0.9	0.0	—	—	—	—
Grass-hopper	4.5	3.3	3.6	0.0	—	—	—
Banded cricket	6.3	5.0	4.9	4.3	0.0	—	—
House cricket A	8.2	7.2	6.7	7.1	3.5	0.0	—
House cricket B	7.8	7.1	7.1	6.4	2.5	2.2	0.0

On the other hand, the ICD value between House cricket B-Banded cricket clusters was 2.5. Considering that both species belong to the same family (Gryllidae), low ICD values could be expected[32]. However, the same outcome was not observed in House cricket A-Banded cricket powders (ICD = 3.5), which were produced by different manufacturers, confirming again the assumption of differentiation by origin. From the results presented in both class projection plots and ICD values, it is possible to conclude that product's origin (i.e., manufacturer) was a key factor for the differentiation in powders made from same insect species, even though no significant differences were found among them (i.e., ICD < 3.0).

Wavenumbers in the spectral range analysed were plotted against their capability to classify and discriminate insect powders tested[19]. Discrimination power plots of all models are shown in FIGURE 3.3, displaying the IR bands responsible for the discrimination among classes. For the whole model (FIGURE 3.3 A), the IR bands at 2945, 2919, and 2851-2825 cm^{-1} could be attributed to CH_3 and CH_2 asymmetric stretching and CH_2 symmetric stretching of lipids and chitin, respectively[23–25,27,28]. The IR bands at 1744-1722 cm^{-1} might be linked to C=O stretching of lipids. All these IR bands were also present in the regions previously commented in the insect powders raw spectra (FIGURE 3.1). Nonetheless, SIMCA analysis revealed which of them were mainly responsible of the discrimination of the samples analysed, hardly detected from just observing their raw spectra. The data obtained suggests that the chemical

differences among the insect powders analysed were related to their chitin and fat fractions when all products were included in a single model (i.e., whole model).

Similar IR bands were obtained for cricket (Figure 3.3 B) and worm models (Figure 3.3 C), as well as new ones related to the insect protein and carbohydrate fractions, which were also present in the raw spectra. The IR band at 3114 cm^{-1} could be associated to O-H stretching of carbohydrates (chitin) or N-H stretching of amide A of proteins or chitin. Furthermore, several IR bands in the spectral region of 1700-1500 cm^{-1} were detected and could be attributed to amide I and II regions of proteins or chitin[23–28]. Once again, the chemical differences among cricket powders were mainly related to lipid and chitin components. Worm powders gave the same output as well.

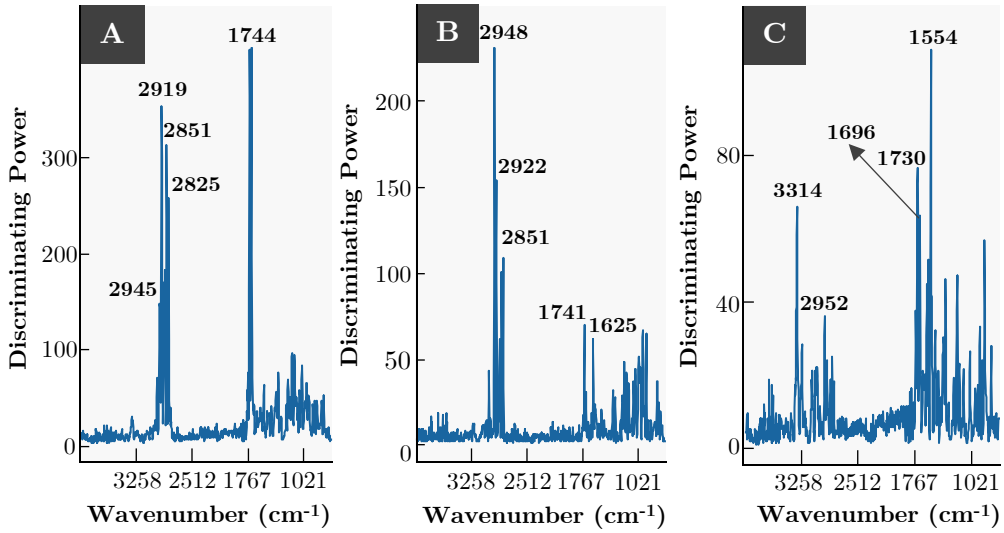


Figure 3.3 SIMCA discriminating power plots of transformed (second derivative, 13-point window) ATR-FTIR spectra (4000-800 cm^{-1} region) of 7-class SIMCA whole model (A), 3-class SIMCA worm model (B), and 3-class SIMCA cricket model (C).

3.3.3 Prediction of unknown samples

SIMCA model performance (whole model) was tested using five spectra per class not included in the training set. Every data point (spectrum) falling inside the 95% confidence interval boundary of a certain class (i.e., class PCA model) would be assigned as a member of that class, otherwise would be

rejected. Class prediction in SIMCA provides three possibilities, whether the observation belongs to one, more than one (next best prediction), or none of the predefined classes[33]. 100% of correct predictions into the appropriate class were obtained for most of the insect powders. Nonetheless, it is important to mention that for Buffalo worm clusters (A and B) some of the next best predictions were assigned to the other Buffalo worm cluster (see Table 3.4). These results, which are in line with both ICD values and class projection plots obtained through SIMCA, confirmed that there are not significant differences between Buffalo worm A and Buffalo worm B. Lastly, even though an ICD lower than 3.0 was obtained when comparing House cricket A and House cricket B clusters, 100% of the spectra tested in the validation set were correctly assigned to each class without obtaining next best predictions. These results confirmed that our SIMCA models could discriminate and properly classify the commercial insect powders tested by species and origin.

Table 3.4 Insect powders model predictions of all edible insect powder model (7-class whole SIMCA model) by internal validation.

Prediction	Meal-worm[1]	Buffalo worm A[2]	Buffalo worm B[3]	Grass-hopper[4]	Banded cricket[5]	House cricket A[6]	House cricket B[7]
Best[a]	100%	100%	80%	100%	100%	100%	100%
Next best[b]	-	60%[3] 20%[1]	75%[2]	-	-	-	-

[a]Percentages refer to spectra that were correctly identified by SIMCA model.
[b]Percentages were obtained from the number of correctly identified samples. Next best prediction is indicated by superscript numbers (1-7).

3.4 Conclusions

The current study has given the framework for a cheap, rapid, and easy technique in edible insect powders discrimination (i.e., origin and species) by using ATR-FTIR combined with SIMCA. Our findings are the first step towards a reliable and easy-to-implement way of preventing fraud and adulterations in the emerging insect sector. A further important implication is the possibility of extrapolating this approach to any insect-derived product. For this reason, future work should concentrate on the creation of models that include a wide range of products and, in collaboration with insect farmers, taking into consideration factors such as feed, sex, and stage when discriminating them. Additionally, to further our research we plan to

determine the feasibility of this technique for detecting insect powders in several food matrixes.

3.5 References

1. Belluco, S., Halloran, A. & Ricci, A. New protein sources and food legislation: the case of edible insects and EU law. *Food Security* 9, 803–814 (2017).
2. Belluco, S. *et al.* Edible insects in a food safety and nutritional perspective: A critical review. *Comprehensive Reviews in Food Science and Food Safety* 12, 296–313 (2013).
3. van Huis, A. & Oonincx, D. G. A. B. The environmental sustainability of insects as food and feed. A review. *Agronomy for Sustainable Development* 37, (2017).
4. van Huis, A. *et al.* Edible insects. Future prospects for food and feed security. *FAO Forestry Papers* 171, 1–201 (2013).
5. Premalatha, M., Abbasi, T., Abbasi, T. & Abbasi, S. A. Energy-efficient food production to reduce global warming and ecodegradation: The use of edible insects. *Renewable and Sustainable Energy Reviews* 15, 4357–4360 (2011).
6. Rumpold, B. A. & Schlüter, O. K. Potential and challenges of insects as an innovative source for food and feed production. *Innovative Food Science and Emerging Technologies* 17, 1–11 (2013).
7. Sun-Waterhouse, D. *et al.* Transforming insect biomass into consumer wellness foods: A review. *Food Research International* 89, 129–151 (2016).
8. Sánchez-Muros, M. J., Barroso, F. G. & Manzano-Agugliaro, F. Insect meal as renewable source of food for animal feeding: A review. *Journal of Cleaner Production* 65, 16–27 (2014).
9. Makkar, H. P. S., Tran, G., Heuzé, V. & Ankers, P. State-of-the-art on use of insects as animal feed. *Animal Feed Science and Technology* 197, 1–33 (2014).
10. Rumpold, B. A. & Schlüter, O. K. Nutritional composition and safety aspects of edible insects. *Molecular Nutrition and Food Research* 57, 802–823 (2013).
11. Ulrich, S. *et al.* Direct identification of edible insects by MALDI-TOF mass spectrometry. *Food Control* 76, 96–101 (2017).
12. Köppel, R. *et al.* Multiplex real-time PCR for the detection of insect DNA and determination of contents of Tenebrio molitor, Locusta

migratoria and Achaeta domestica in food. *European Food Research and Technology* 245, 559–567 (2019).

13. Ali, M. E., Razzak, M. A. & Hamid, S. B. A. Multiplex PCR in Species Authentication: Probability and Prospects—A Review. *Food Analytical Methods* 7, 1933–1949 (2014).
14. Levin, R. E., Ekezie, F.-G. C. & Sun, D. W. DNA-Based Technique: Polymerase Chain Reaction (PCR). in *Modern Techniques for Food Authentication* (ed. Sun, D. W.) 527–616 (Elsevier Inc., 2018). doi:10.1016/b978-0-12-814264-6.00014-1.
15. Singhal, N., Kumar, M., Kanaujia, P. K. & Virdi, J. S. MALDI-TOF mass spectrometry: An emerging technology for microbial identification and diagnosis. *Frontiers in Microbiology* 6, 1–16 (2015).
16. Sébédio, J. L. & Malpuech-Brugère, C. *Implementation of Foodomics in the Food Industry. Innovation Strategies in the Food Industry: Tools for Implementation* (Elsevier Inc., 2016). doi:10.1016/B978-0-12-803751-5.00013-1.
17. Rodriguez-Saona, L. E. & Allendorf, M. E. Use of FTIR for Rapid Authentication and Detection of Adulteration of Food. *Annual Review of Food Science and Technology* 2, 467–483 (2012).
18. Wenning, M. *et al.* Identification and differentiation of food-related bacteria: A comparison of FTIR spectroscopy and MALDI-TOF mass spectrometry. *Journal of Microbiological Methods* 103, 44–52 (2014).
19. Wold, S. & Sjöström, M. SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy. in *Chemometrics: Theory and Application* (ed. Kowalski, B. R.) 243–282 (ACS Symposium Series, 1977). doi:10.1021/bk-1977-0052.ch012.
20. Shah, N. K. & Gemperline, P. J. Combination of the Mahalanobis Distance and Residual Variance Pattern Recognition Techniques for Classification of Near-Infrared Reflectance Spectra. *Analytical Chemistry* 62, 465–470 (1990).
21. Supervised Pattern Recognition. in *Handbook of Chemometrics and Qualimetrics: Part B* (eds. Vandeginste, B. G. M. et al.) vol. 20 207–241 (Elsevier Science B.V., 1998).
22. Rodriguez-Saona, L. E. & Allendorf, M. E. Use of FTIR for Rapid Authentication and Detection of Adulteration of Food. *Annual Review of Food Science and Technology* 2, 467–483 (2012).

23. Ibitoye, E. B. *et al.* Extraction and physicochemical characterization of chitin and chitosan isolated from house cricket. *Biomedical Materials (Bristol)* 13, (2018).
24. Paulino, A. T., Simionato, J. I., Garcia, J. C. & Nozaki, J. Characterization of chitosan and chitin produced from silkworm crysalides. *Carbohydrate Polymers* 64, 98–103 (2006).
25. Marchessault, R. H., Pearson, F. G. & Liang, C. Y. Infrared spectra of crystalline polysaccharides. *Biochimica et Biophysica Acta* 45, 499–507 (2003).
26. Talari, A. C. S., Martinez, M. A. G., Movasaghi, Z., Rehman, S. & Rehman, I. U. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. *Applied Spectroscopy Reviews* 52, 456–506 (2017).
27. Stuart, B. H. Infrared Spectroscopy of Biological Applications: An Overview. *Encyclopedia of Analytical Chemistry* (2012) doi:10.1002/9780470027318.a0208.pub2.
28. Socrates, G. *Infrared and Raman characteristic group frequencies. Infrared and Raman characteristic group frequencies* (John Wiley & Sons, Inc., 2004). doi:10.1002/jrs.1238.
29. Kvalheim, O. M. & Karstang, T. V. SIMCA - Classification by means of disjoint cross validated principal components models. in *MultiVariate Pattern Recognition in Chemometrics: Illustrated by Case Studies* (ed. Brereton, R. G.) 209–248 (Elsevier, 1992).
30. Dunn, W. J. & Wold, S. SIMCA Pattern Recognition and Classification. in *Chemometric Methods in Molecular Design* (ed. van de Waterbeemd, H.) 179–192 (VCH Publishers, Inc., 1995).
31. Malinowski, E. R. StatisticalF-tests for abstract factor analysis and target testing. *Journal of Chemometrics* 3, 49–60 (1989).
32. Krinsky, W. L. Beetles (coleoptera). in *Medical and Veterinary Entomology* (eds. Mullen, G. R. & Durden, L. A.) 129–143 (Academic Press, 2019).
33. He, J., Rodriguez-Saona, L. E. & Giusti, M. M. Midinfrared spectroscopy for juice authentication-rapid differentiation of commercial juices. *Journal of Agricultural and Food Chemistry* 55, 4443–4452 (2007).

UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

Application of hand-held NIR and Raman spectrometers in surface treatment authentication of cork stoppers

This chapter is published as: Mellado-Carretero, J., Peren-Aykas, D., Puxeu, M., Varela, S., Rodriguez-Saona, L., García-Gonzalo, D., De Lamo-Castellví, S., 2021. Application of hand-held near-infrared and Raman spectrometers in surface treatment authentication of cork stoppers. Journal of Food Packaging and Shelf Life, 28, p.100680

ABSTRACT

The aim of this paper was to evaluate the potential of using near-infrared (NIR) spectroscopy and multivariate analysis as a rapid tool to non-destructively determine the presence of surface treatments applied to cork stoppers. Density and dimensions of 6 closure varieties were characterized and the extraction force was measured on those produced for still wines. Cork stoppers were also analysed using hand-held NIR and Raman spectrometers. Soft independent modelling of class analogy (SIMCA) models showed significant differences among treated and untreated samples, linked to components of the coating agents applied (silicone and paraffin). SIMCA model's classification performance was tested, and high sensitivity (93.33%) and specificity (100 %) values were obtained. Partial least squares regression (PLSR) model accurately predicted the extraction forces measured with low standard error of prediction (SEP = 4.0 daN). Our results are promising for the future application of this technology in cork industry, reducing time and economic losses.

4.1 Introduction

Cork, a natural material obtained from the bark of the cork oak (*Quercus suber L.*), has been used for stopping wine containers during more than 2000 years. Different physicochemical properties make cork suitable for its use as a wine closure, such as compressibility, resilience, high friction coefficient, imperviousness to liquids, chemical inertness, and air permeability[1–3]. Consequently, cork stoppers prevent any leakage, do not negatively change chemical and sensorial characteristics of wine, and allow an adequate oxygen transfer[2].

Because of the strong friction forces between cork and glass, cork stoppers need to be treated with several food grade coating agents, being paraffin and silicone the most commonly used. Moreover, such agents restore the imperviousness lost during the processing steps of cork manufacturing and improve their sealant properties by filling imperfections in the cork surface. The hydrophobic properties of paraffin help to reduce wine leaks and prevent the migration of potential taints from the stopper to the wine. However, this treatment can aggravate the stopper's extraction when applied without slipping additives. Therefore, paraffin is usually covered up with a layer of silicone (polydimethylsiloxane) in order to lubricate the stopper and simplify its insertion and extraction from the bottle[1,4,5]. Accordingly, it is imperative to check the presence of surface treatments applied to cork stoppers, otherwise the uncorking of those bottles sealed with nonproper treated stoppers becomes impossible, leading to their disposal.

At present, different methods have been proposed for the analysis of surface treatments applied to cork stoppers. The presence of surface treatments is indirectly assessed by measuring the extraction or torque forces for still wine stoppers and sparkling wine stoppers, respectively. Additionally, a simple test can be applied by examining the capillarity effect on the surface of cork stoppers by standing them upright in a coloured liquid[6]. Even though both tests are a simple way to check the presence of surface treatments, they cannot distinguish the different components applied on the cork and do not provide information about the homogeneity of their distribution[7].

Infrared spectroscopy has been used for cork quality assessments since it rapidly provides a chemical fingerprint of the different components that lie in its heterogeneous matrix[8]. Ortega-Fernández et al. (2006) could successfully

determine the dose and type of several surface treatments applied to cork stoppers by using attenuated total reflectance Fourier transform mid-infrared (ATR-FTIR) spectroscopy combined with powerful pattern recognition techniques[9]. Moreover, González-Gaitano and Ferrer (2013) provided several parameters obtained from ATR-FTIR spectra for the quantification of surface treatments applied in several cork stoppers[10]. Nonetheless, FTIR spectroscopy cannot be applied for on-line process monitoring since it implies the preparation and destruction of the samples, leading to time and economic losses for the manufacturers.

Near-infrared (NIR) spectroscopy has been used to quantitatively determine different physicochemical and mechanical parameters of cork (i.e., moisture, porosity, visual quality, etc.) and to classify cork stoppers by variety and origin[11–13]. Its low instrumentation costs (when compared to MIR technology) and the possibility to implement it on-line in a non-destructive way makes NIR spectroscopy a suitable technique for the assurance of the surface treatments applied on cork stoppers[14,15].

For this reason, the present work aims to validate NIR spectroscopy combined with multivariate analysis as a rapid and powerful tool used to non-destructively determine the presence of surface treatments applied to cork stoppers. Physical parameters to evaluate the surface treatments administered such as extraction force were also studied and compared with the technology proposed herein. Additionally, the feasibility of Raman spectroscopy for the surface treatment monitoring of cork stoppers was also tested. The results provided in this paper are promising for the future application of NIR spectroscopy for quality control assessments in the cork industry.

4.2 Materials and methods

4.2.1 Cork stopper samples

Six cork closure varieties were provided by a regional cork stopper manufacturer (Figure 4.1); including natural cork stoppers for still wines (N-S), technical cork stoppers consisting of agglomerated cork (particle size of 2-7 mm) with two end discs of natural cork for still and sparkling wines (AD-S and AD-SP, respectively), micro-agglomerated cork stoppers (particle size of 0.5-2 mm) for still wines and sparkling wines (M-S and M-SP, respectively), and cork stoppers made of a mixture of agglomerated and micro-agglomerated

FIGURE 4.1 Photographs of treated (A) and untreated (B) N-S cork stoppers, treated (C) and untreated (D) AD-S corks, and treated (E) and untreated (F) M-S stoppers.

cork for sparkling wines (A/M-SP). The composition of all body-agglomerated cork stoppers was 75% cork granules and 25% binder (polyurethane; information provided by the supplier).

Each variety consisted of 48 closures, 24 of them treated and the rest untreated in order to take them as a reference. Silicone (polydimethylsiloxane) treatments were applied with a pneumatic gun at room temperature, which projected the lubricating product on the moving corks in an industrial revolving stainless-steel drum (20 rpm for 1-2 min). The stoppers were subsequently left for 30 min more inside the drum. For N-S cork stoppers, paraffin was added prior to silicone and placed into the drum in a solid form (20 rpm at 35-40 °C for 30 min). Samples analysed in this study and the type of treatment applied to them are summarized in Table 4.1.

4.2.2 Physical assays

The mean diameter, ovalization, height, mass and apparent density of cork stoppers were measured according to the Spanish UNE standards[16–19]. Before performing the physical and mechanical assays, cork stoppers were conditioned for 24 h at 20 °C and 65% relative humidity. Diameters and heights were measured using a calliper with 0.01 mm precision. The mean diameter (D; expressed in mm) was obtained by averaging the diameters measured at parallel (D_1; in mm) and perpendicular directions (D_2; in mm) to the growth lines of the cork (Equation 4.1):

$$D = \frac{D_1 + D_2}{2} \tag{4.1}$$

The apparent density (d; in kg/m^3) was then obtained from Equation 4.2 (natural and agglomerated still wine cork stoppers):

$$d = \frac{4 \cdot 10^6 \cdot m}{\pi \cdot D^2 \cdot h} \tag{4.2}$$

where m is the mass in g, D is the mean diameter expressed in mm, and h is the height in mm. For natural cork stoppers, the ovalization (OV; in mm) was calculated by subtracting diameters D_1 and D_2 (in mm) and expressed in absolute value (Equation 4.3):

$$OV = |D_1 - D_2| \tag{4.3}$$

TABLE 4.1 Physical properties of cork stoppers analysed in this study.

Cork stopper	Treatment applied	Dose (mg/stopper)	Diameter (mm)	Height (mm)	Ovalization (mm)	Mass (g)	Density (kg/m^3)
N-S	Untreated	—	24.0 ± 0.1[a]	44.4 ± 0.1	0.1 ± 0.1	NA	177 ± 26
	Paraffin and silicone	9 (paraffin) 12 (silicone)	24.0 ± 0.1	44.1 ± 0.2	0.1 ± 0.1	NA	174 ± 17
AD-S	Untreated	—	23.3 ± 0.1	43.9 ± 0.1	NA	NA	272 ± 9
	Silicone[b]	12	23.4 ± 0.1	44.0 ± 0.1	NA	NA	276 ± 10
M-S	Untreated	—	23.9 ± 0.1	42.2 ± 0.2	NA	NA	283 ± 5
	Silicone	12	24.8 ± 0.1	43.8 ± 0.8	NA	NA	293 ± 10
AD-SP	Untreated	—	30.3 ± 0.1	48.1 ± 0.2	NA	9.2 ± 0.3	NA
	Silicone	37	30.8 ± 0.1	48.1 ± 0.1	NA	9.2 ± 0.2	NA
A/M-SP	Untreated	—	30.3 ± 0.1	48.0 ± 0.1	NA	9.6 ± 0.2	NA
	Silicone	37	29.8 ± 0.1	48.0 ± 0.1	NA	9.3 ± 0.2	NA
M-SP	Untreated	—	30.9 ± 0.0	48.2 ± 0.1	NA	9.6 ± 0.2	NA
	Silicone	37	30.3 ± 0.1	48.0 ± 0.1	NA	9.5 ± 0.2	NA

[a]Values are means of twenty-four replicates ± standard deviation (SD).
[b]Silicone oil.
Abbreviations used: N-S, natural cork-still wine; AD-S, agglomerated cork with discs-still wine; M-S, micro-agglomerated cork-still wine; AD-SP, agglomerated cork with discs-sparkling wine; A/M-SP, Mixture agglomerated/micro-agglomerated cork-sparkling wine; M-SP, micro-agglomerated cork-sparkling wine. NA, not applicable according to UNE standards.

4.2.3 Mechanical assays: Extraction force

The extraction force was measured following UNE standards with minor modifications[16–19]. First, corking was accomplished using a semi-automatic capping unit Officine Pesce model PG91 RN (Officine Pesce, Italy). The bottles (Bordeaux type, with 18.5 mm bottleneck) were previously filled with a hydroalcoholic solution (12% ethanol with 5 g/L of tartaric acid adjusted to pH = 3.5 with sodium hydroxide), and then left in an upright position. After 24 h, the cork stoppers were removed using an ExtraLab Portable device (EGITRON, Portugal) and the maximum value of the extraction force was subsequently recorded in daN. Twelve cork stoppers of each variety were analysed in this study.

4.2.4 Diffusive reflectance near-infrared (NIR) spectroscopy

NIR spectra were obtained from different spots of each cork stopper (see Figure 4.2; the spectra were taken considering the different sections of the cork plank from which the stopper came, since these could show differences on their physical properties) using a hand-held NIR analyser Polychromix microPHAZIR (Thermo Scientific, US) with a scanning window of 4 mm (sampling area of 13 mm^2) and equipped with a single InGaAs detector. NIR spectra were collected in the region of 1600-2400 nm with 8 nm resolution from the average of 64 scans to improve the signal-to-noise ratio. A background spectrum was collected with a highly reflective gold-coated reference material before every measurement to prevent the environment contribution. Four independent stoppers were randomly selected and analysed following the scheme depicted in Figure 4.2, except for untreated N-S stoppers, in which twelve of them were analysed for reproducibility testing purposes. The number of spectra (Table 4.2) in treated samples could vary (7-10 spectra taken) due to the presence of brand names printed with fire (Figure 4.1).

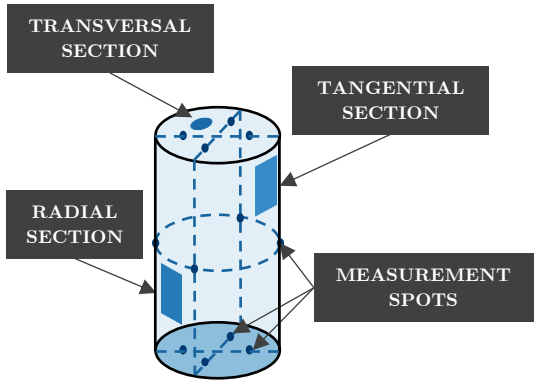


Figure 4.2 Scheme of a cork stopper and measurement spots taken for spectroscopic analyses.

TABLE 4.2 Variances explained (as cumulative percentages) for each principal component and number of spectra collected and rejected per class in cork pairwise models (2-class SIMCA models).

Pairwise model	Class (sample)	PC1	PC2	PC3	PC4	PC5	PC6[a]	Spectra collected	Outliers
N-S	Treated	48.3	74.1	83.7	89.4	92.4	—	36	3
	Untreated	63.6	80.9	89.7	93.3	95.3	—	144	11
AD-S	Treated	54.1	76.2	84.8	90.3	93.4	94.6	48	8
	Untreated	54.3	70.2	78.4	84.3	87.9	90.2	48	4
M-S	Treated	37.1	59.0	75.4	84.3	88.0	90.3	32	7
	Untreated	46.5	69.6	79.1	84.2	87.5	90.0	48	0
AD-SP	Treated	54.9	75.0	85.7	90.8	93.3	94.9	32	3
	Untreated	51.7	72.9	81.6	86.0	89.1	91.5	48	3
A/M-SP	Treated	43.9	79.4	86.5	91.8	94.5	96.3	32	2
	Untreated	58.2	77.1	82.9	86.8	89.6	91.5	48	5
M-SP	Treated	53.1	74.0	84.6	89.6	91.9	—	32	3
	Untreated	55.3	73.3	82.6	86.8	90.2	—	48	2
All but N-S	Treated	64.4	78.8	86.9	92.9	94.7	—	176	29
	Untreated	57.4	75.0	87.0	89.6	91.3	—	240	38

[a]The number of principal components (PC) selected per model was to obtain at least 90% of explained variance.
[b]Sample residuals and Mahalanobis distances were used for outlier determination.
Abbreviations used: N-S, natural cork-still wine; AD-S, agglomerated cork with discs-still wine; MD-S, micro-agglomerated cork with discs-still wine; AD-SP, agglomerated cork with discs-sparkling wine; A/M-SP, Mixture agglomerated/micro-agglomerated cork-sparkling wine; M-SP, micro-agglomerated cork-sparkling wine.

4.2.1 Raman spectroscopy

Raman spectra were obtained using a hand-held Raman Progeny analyser (Rigaku, US) interfaced with a 1064 nm (excitation wavelength) laser and an InGaAs detector (sampling area of 500 μm^2). Raman spectra were collected from 200-2500 cm^{-1} (8-11 cm^{-1} resolution) with a power of 150 mW and 1500 ms integration time. Only M-SP stoppers were analysed in the present work.

4.2.2 NIR and Raman spectra pre-processing

NIR data pre-processing was performed using Pirouette 4.5 (Infometrix, US). After mean-centring, NIR data was corrected using a Multiplicative Scatter Correction (MSC; the mean of the entire data set was taken as the reference spectrum) algorithm and transformed to its second derivative through a second polynomial Savitzky-Golay filter of 7 points width.

Raman pre-processing was first accomplished using MATLAB 9.5 (Mathworks, US). Spectra denoising was achieved using sym4 wavelets with two-level decomposition and soft minimax thresholding technique. Denoised spectra were then uploaded to Pirouette 4.5 to perform mean-centring and MSC algorithms (the mean of each sample data set was taken separately as the reference spectrum).

4.2.3 Multivariate analysis applied to Raman and NIR data

SIMCA models were built up to study the capability of NIR and Raman data to discriminate and differentiate the treated corks from those without treatment. Models were interpreted in terms of class projection plots [i.e., principal component analysis (PCA) score plots], interclass distances, and discriminating power[20,21]. Outliers' determination was performed according to sample residuals and Mahalanobis distances[22]. SIMCA models' classification performance was tested using an external validation set consisting of one stopper per variety analysed [5 treated closures (60 spectra in total) and 5 untreated (44 spectra); natural cork stoppers were not included in this set] and evaluated based on its sensitivity and specificity. The sensitivity is known as the percentage of samples from a certain class that are correctly identified by the model of that class, while specificity accounts for the percentage of samples belonging to other classes that are correctly rejected by that model[23]. All the mentioned outputs were based on the individual spectra.

Calibration models were constructed based on the PLSR method in order to correlate the reference values of extraction force with the NIR data. The number of latent variables selected was determined through an F test (probability level of 95%) to assess which model had the lowest and statistically different prediction residual error sum of squares (PRESS) value[24]. Models were interpreted in terms of regression vector and evaluated according to the coefficients of determination in calibration (R^2_{CAL}) and validation (R^2_{VAL}) and the standard error of cross-validation (SECV). In this regard, the cross-validation step was conducted using the leave-one-out approach. Studentized residuals and leverage values were used for outliers' assessment[25]. PLSR model's performance was tested using an external validation set consisting of one stopper per variety analysed and examined in terms of standard error of prediction (SEP).

4.2.4 Statistical analysis

A Welch's ANOVA test was conducted since the homoscedasticity assumption was not met (homoscedasticity and normality of the distributions were tested through the Levene's test and Shapiro-Wilk's test, respectively). A Dunnett's T3 post-hoc test was performed to obtain pairwise comparisons among sample means in extraction force data. The significance level was set at $P < 0.05$ in all tests. These statistical analyses were conducted using the software IBM SPSS Statistics for Windows, version 25.0 (IBM, US).

4.3 Results and discussion

4.3.1 Cork stoppers physical and mechanical characterizations

The UNE 56921 standard specifications for natural cork stoppers establishes ovalization values lower than 0.5 mm and apparent densities of 125-230 kg/m^3. Both ovalization and apparent density values obtained for N-S stoppers were within the aforementioned ranges (Table 4.1)[17]. Concerning micro-agglomerated cork stoppers for still wines, the UNE 56933 standard proposes an acceptance range of apparent densities from 240 kg/m^3 to 350 kg/m^3 [19], which is in agreement with the data obtained for M-S cork stoppers. AD-S cork stoppers also met the specifications of the UNE 56926 standard for apparent density values (235-315 kg/m^3, Table 4.1)[16]. Furthermore, the mass of cork stoppers made for sparkling wines was measured as an indirect estimator of the density. The mean values of masses for AD-SP, A/M-SP and M-SP were within the range stated by the UNE 56923 standard (8.4-10 g, Table 4.1)[18].

The extraction force measurements for the cork stoppers analysed are shown in Figure 3.2. The mean values for the extraction force in N-S, AD-S, and M-S samples were consistent with the data from literature (mean values of 19.6 daN and 34.7 daN for natural and technical cork stoppers, respectively)[26], as well as they met the UNE standard specifications (20-40 daN for N-S and AD-S samples and 15-40 daN for M-S samples)[16,17]. Significant differences were observed in extraction force values between treated and untreated corks in N-S and AD-S samples, confirming through this indirect method the presence of surface-coating agents in N-S and AD-S cork stoppers. However, no significant differences were detected between treated and untreated M-S samples. Sánchez-González and Pérez-Terrazas (2018) stated

that cork percentage and density provide a good indicator of the mechanical behaviour of agglomerated cork stoppers. In their work, when comparing micro-agglomerated stoppers with the same amount of cork, those having higher densities displayed higher extraction force values[27]. The fact that the treated and untreated M-S samples (same cork percentage) displayed differences in density (also in dimensions, Table 4.1) could be a reason why the extraction force values recorded were similar.

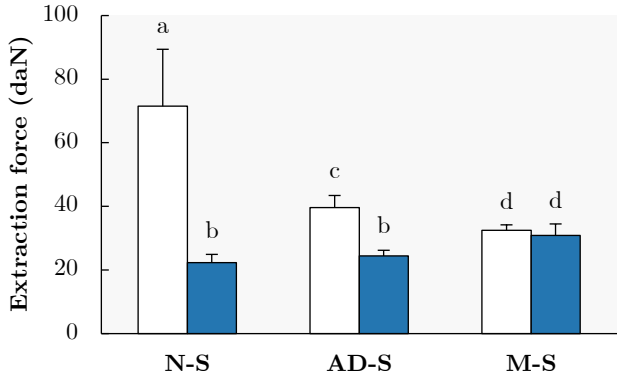


Figure 4.2 Extraction force measurements for untreated (blank) and treated (blue) still wine stoppers inserted in an 18.5 mm diameter Burgundy type bottle, with and without surface treatments. Error bars in the plot represent the standard deviation for 12 independent replicates. Means with different letters are significantly different ($P < 0.05$).

4.3.2 Interpretation of raw NIR and Raman spectra

Raw NIR and Raman (denoised) spectra of treated and untreated cork stoppers are shown in Figure 4.3. Since all NIR spectra looked very similar, only one obtained for one type of cork stopper is presented. The NIR spectra of cork stoppers (Figure 4.3 A) displayed the same profile as the ones reported previously[12,28]. Several NIR bands were identified and related to the main components of the corks. The NIR band at 1725 nm corresponds to the CH vibrations (first overtone stretching) of lignin or hemicellulose, as well as the NIR bands at around 2300 and 2345 nm could be assigned to CH (C–H stretching + C–H deformation) and CH_2 (C–H bending + C–H stretching) groups of cellulose or hemicellulose. Additionally, other NIR bands were observed at 1927 and 2145 nm, the former being attributed to OH groups of water (O–H stretching + O–H deformation) and the latter to the CH and CO combination band (C–H stretching + C=O stretching)[12,28–30].

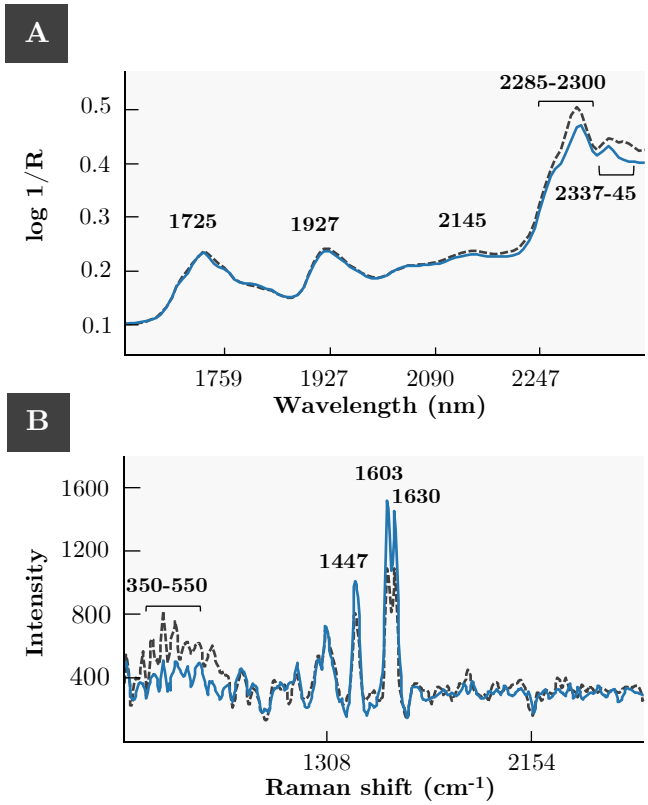


FIGURE 4.3 NIR (A) and Raman (B) spectra profiles of untreated (solid line) and treated (dashed line) cork stoppers (M-SP samples).

Regarding Raman spectra, the main Raman bands observed (FIGURE 4.3 B) were in accordance with previous results obtained for other wood species. The Raman bands located at 1603 and 1630 cm^{-1} could belong to the aryl ring stretching vibrations of lignin and the ring conjugated C=C stretching vibrations of coniferaldehydes, respectively[31,32]. Moreover, the band at 1447 nm might be linked to guaiacyl ring vibrations of lignin. Several Raman bands were also displayed in the region of 350-550 cm^{-1}, and could be associated to the skeletal deformation of aromatic rings, substituent groups, and side chains of lignin[32].

4.3.3 Surface treatment authentication by NIR spectroscopy combined with SIMCA

A PCA-based pattern recognition method, SIMCA, was performed to obtain classification models in order to discriminate treated and untreated cork stoppers and to evaluate the chemical components that could make them different. Class projections plots (i.e., PCA score plots of the entire data set) of 2-class SIMCA models built up for NIR data (Figure 4.4 and Table 4.2), each of them comprising treated and untreated corks for every type of stopper, were generated to visualize spectra reproducibility and class separation. The ellipses depicted in the plots represent the regions in which samples from a certain class fall into with a 95% of confidence[33]. Every data point in the figure (a 2D representation of a 3D graph) represents one sample's spectrum. All the classes were tightly clustered and well separated (not overlapped, Figure 4.4), indicating discrimination among treated and untreated cork stoppers regardless of their variety. However, some of the outliers detected (Table 4.2), which belonged to treated cork stoppers spectra, were classified as untreated samples. These results suggest that the surface treatments applied were not entirely homogeneous.

An alternative approach to study class separation is the interclass distance (ICD) value. In SIMCA, PCA models for every class in the training set are performed, and then the residuals are computed by fitting the objects of every training set's class to the PCA model of each class. The overall standard deviations of that residuals are used to calculate the ICD value, which is a ratio of interclass to intraclass distance. ICD values close to zero mean no differentiation while values greater than one indicates class separation[21]. In general, two clusters of samples are considered significantly different when an ICD value above 3.0 is reached[34]. ICDs among cork stopper samples are displayed in Figure 4.4 together with the class projection plots. ICD values greater than 3.0 (ranging from 4.3 to 6.8) were achieved for all compared cork stoppers, suggesting significant differences between treated and untreated cork samples for every type of closure.

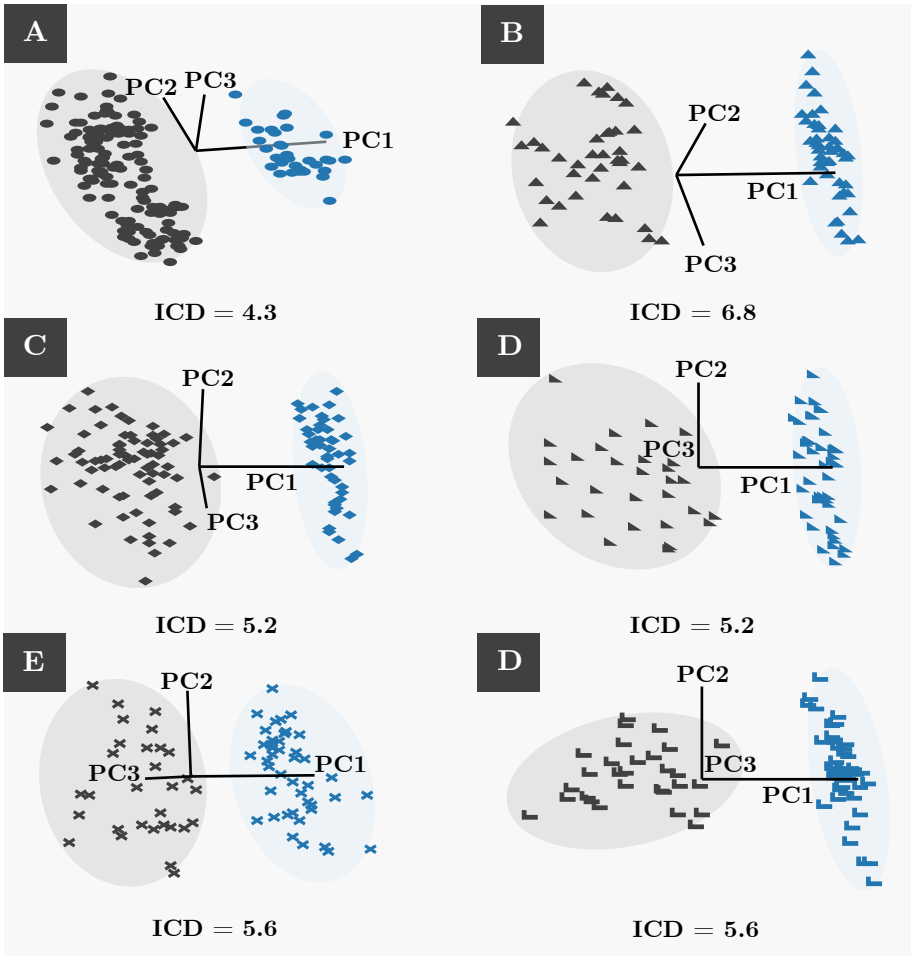


FIGURE 4.4 SIMCA class projection plots and interclass distance (ICD) values of transformed (MSC and second derivative, 7-point width) diffusive reflectance near-infrared spectra of N-S (A, ○), AD-S (B, △), M-S (C, ◇), AD-SP (D, ◺), A/M-SP (E, ⯃), and M-SP (F, ⌙) cork stoppers. Untreated and treated stoppers are represented by blue and grey filled symbols, respectively.

In addition to providing class separation, residuals give valuable information of the strength of any given variable to discriminate among classes[21]. The so-called discriminating power plots gather the spectral bands being responsible of the samples' differentiation. Discriminating power plots of all 2-class SIMCA models built up with NIR data are shown in FIGURE 4.5. For all models except that of the natural cork stopper, same NIR bands were obtained related to the chemical composition of the surface treatments applied.

The NIR bands located at 1933 and 2285-92 nm could be assigned to Si–O–H and Si–O–Si combination band (Si–O–H stretching + Si–O–Si deformation) and CH vibrations (second overtone bending) of silicone, respectively[29]. Other NIR bands were also obtained at 2337 and 2367 nm and may be attributed to CH combination band (C–H stretching + C–H deformation) of cellulose or silicone[30,35]. Minor NIR bands also found in 1650-1800 nm region could be linked to CH groups (first overtone C–H stretching) of cellulose and lignin or silicone[35]. Concerning natural cork stoppers, which were treated with a mixture of paraffin and silicone coating agents, other NIR bands were observed. The band at 2130 nm could be associated to CH and CO combination band (C–H stretching + C=O stretching) of cellulose[30], while the bands located in the region between 2200-2400 nm might be related to CH_2 groups of paraffin[29]. All the chemical groups already discussed were also displayed in paraffin and silicone-treated cork stoppers through mid-infrared (MIR) spectroscopy in previous conducted studies[7,9,10], supporting the suitability of NIR spectroscopy coupled to pattern recognition techniques for monitoring the surface treatment of cork closures.

Besides the characterization of surface treatments, another SIMCA model including only untreated samples confirmed that there were no chemical differences between technical and agglomerated corks (Figure 4.6 A and Table 4.3 and Table 4.4). Then, they were considered as a single class in a new model in order to assess its performance when obviating cork stopper's type (natural cork stoppers were excluded from the training set). Class projection and discriminating power plots of this model (Figure 4.6 B) proved that surface treatment can be authenticated through NIR analysis without considering closure's variety (except for natural cork stoppers), showing same separation and same NIR bands that in previous models (Figure 4.5), the latter being attributed to the main chemical groups of the surface coating agent applied (silicone).

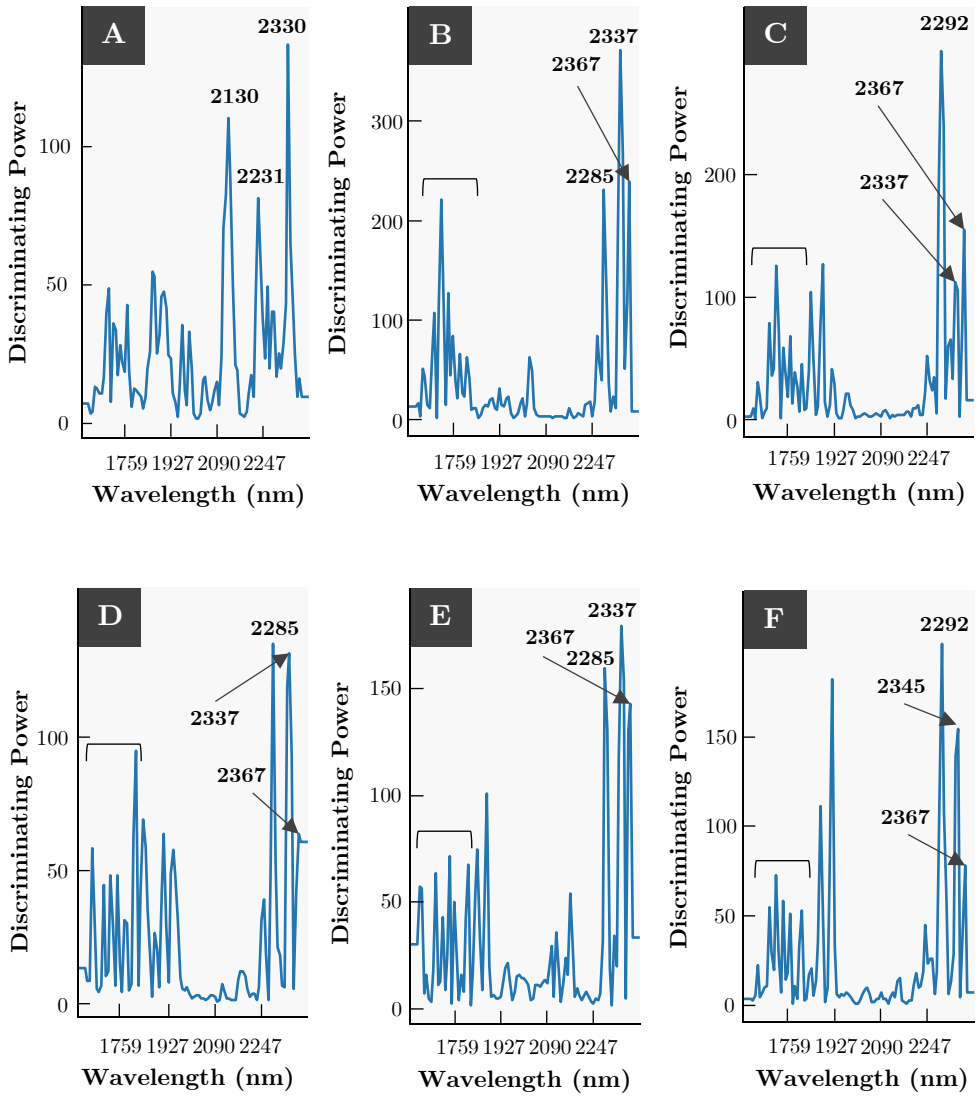


FIGURE 4.5 SIMCA discriminating power plots of transformed (MSC and second derivative, 7-point width) diffusive reflectance NIR spectra of N-S (A), AD-S (B), M-S (C), AD-SP (D), A/M-SP (E), and M-SP (F) cork stoppers.

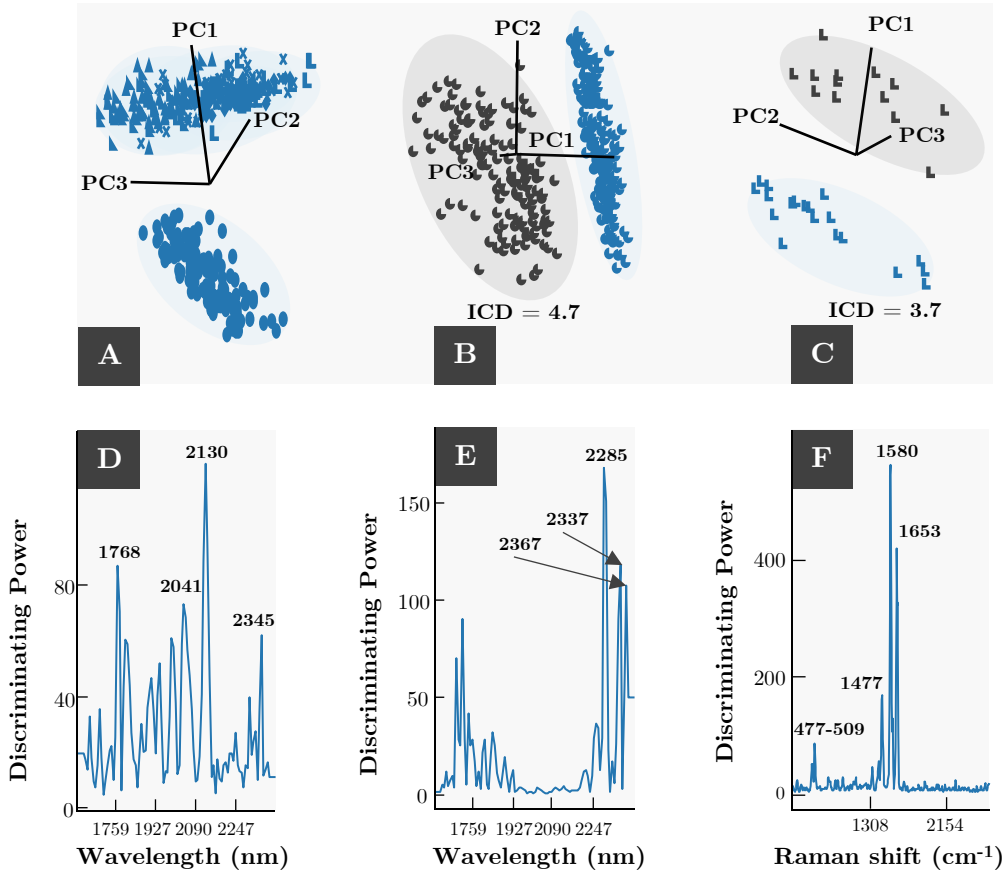


FIGURE 4.6 SIMCA class projection and discriminating power plots of transformed (MSC and second derivative, 7-point width) diffusive reflectance NIR spectra of untreated cork stoppers (A), all but not natural cork stoppers (B), and class projection and discriminating power plots of (denoised and MSC corrected) Raman spectra of M-SP cork stoppers. Untreated and treated stoppers are represented with blue and grey filled symbols, respectively. Interclass distance (ICD) values for (B) and (C) are also depicted.

Lastly, the classification performance of this model was tested using an external validation set. The model showed high specificity and sensitivity values, being 97.73% (43/44 spectra) and 96.67% (58/60 spectra) for the treated samples (5 samples analysed, one for each variety except for natural cork stoppers), respectively. Regarding untreated samples (5 samples analysed, one for each variety except for natural cork stoppers), the sensitivity and specificity were 93.33% (56/60 spectra) and 100.00% (44/44 spectra),

respectively. Our results confirmed that the model developed can discriminate the treated corks from those without treatment.

TABLE 4.3 Variances explained (as cumulative percentages) by each principal component and number of spectra rejected per class in the untreated cork stoppers model (6-class SIMCA model).

Class	PC1	PC2	PC3	PC4	PC5	PC6	PC7[a]	Outliers[b]
N-S	64.1	80.6	89.5	93.3	95.2	96.5	97.1	10
AD-S	54.3	70.2	78.4	84.2	87.8	90.2	92.3	4
MD-S	46.5	69.7	79.2	84.2	87.5	89.9	91.8	0
AD-SP	51.8	72.8	81.5	86.0	89.1	91.5	93.2	3
A/M-SP	58.2	76.9	82.7	86.7	89.5	91.5	93.1	5
M-SP	55.2	73.3	82.6	86.8	90.2	92.1	93.5	2

[a]The number of principal components (PC) selected per model was to obtain at least 90% of explained variance.
[b]Sample residuals and Mahalanobis distances were used for outlier determination.
Abbreviations used: N-S, natural cork-still wine; AD-S, agglomerated cork with discs-still wine; MD-S, micro-agglomerated cork with discs-still wine; AD-SP, agglomerated cork with discs-sparkling wine; A/M-SP, Mixture agglomerated/micro-agglomerated cork-sparkling wine; M-SP, micro-agglomerated cork-sparkling wine.

TABLE 4.4 Interclass distances of pre-processed NIR spectra (1600-2400 nm region) of untreated cork stoppers model (6-class SIMCA model).

Class	N-S	AD-S	M-S	AD-SP	A/M-SP	M-SP
N-S	0.0	—	—	—	—	—
AD-S	4.7	0.0	—	—	—	—
M-S	6.0	2.2	0.0	—	—	—
AD-SP	4.4	1.3	0.8	0.0	—	—
A/M-SP	5.0	0.9	1.0	0.8	0.0	—
M-SP	7.0	1.2	1.9	1.8	0.8	0.0

See TABLE 4.3 for abbreviations' descriptions.

4.3.4 Surface treatment authentication by Raman spectroscopy combined with SIMCA

The feasibility of Raman spectroscopy in surface treatment authentication was also studied in M-SP stoppers. FIGURE 4.6 C and 4.6 F show the class projection (with the ICD value) and discriminating power plots obtained through SIMCA's algorithm, respectively. Tight clustering and clear differentiation were achieved with an ICD value of 3.7, confirming that treated and untreated M-SP stoppers were statistically different. Raman bands displayed in discriminating power plots were related to chemical components of silicone, such as 1477 and 1580 cm^{-1} (aromatic C–H bending and ring

stretching of phenyl groups, respectively). The Raman bands around 477-509 cm^{-1} could be linked to Si–O–Si stretching of silicone, functional group also found in SIMCA models built up with NIR spectra[35]. Raman spectroscopy provided complementary information about the treatments applied to cork stoppers, fitting with the data obtained through NIR spectroscopy.

4.3.5 Development of a PLSR calibration model for predicting extraction forces in cork stoppers using NIR data

A factor-based regression technique was selected instead of Multiple Linear Regression (MLR) due to the highly correlated nature of spectroscopic measurements. In PLSR, new uncorrelated variables are first obtained by establishing linear combinations of the original ones. Then, a regression model is performed following the least squares' criterion (i.e., minimizing the error sum of squares)[36].

Good linear correlations were observed between measured and predicted extraction force values in the 5-factor PLSR model built up (Table 4.5), with a coefficient of determination in validation (R^2_{VAL}) of 0.97 (Figure 4.7). Moreover, the SECV and SEP values (i.e., estimators of the expected error of predicting the dependent variable of an unknown sample[37]) were low (3.8 daN and 4.0 daN, respectively). The fact that SECV and SEP were comparable indicated that the calibration model can predict accurately the extraction force from new NIR data (i.e., external validation set)[38]. The model displayed a better predictive performance than those described in previous literature[12].

Table 4.5 Extraction forces measured (reference method) and predicted using NIR data and PLSR models.

Class	N-S		AD-S		M-S	
	Untreated	Treated	Untreated	Treated	Untreated	Treated
Measured	71.5 ± 17.9[a]	22.3 ± 2.6	40.0 ± 3.8	24.4 ± 1.8	32.5 ± 1.7	30.8 ± 3.6
$Pred_{CAL}$	70.4 ± 4.0	22.4 ± 3.2	43.0 ± 2.9	24.3 ± 3.0	32.8 ± 2.8	30.2 ± 2.9
$Pred_{VAL}$	70.3 ± 4.2	22.4 ± 3.4	43.1 ± 3.1	24.3 ± 3.2	32.8 ± 2.9	30.2 ± 2.9
Predicted	69.1 ± 3.2	19.8 ± 3.9	41.8 ± 5.1	23.9 ± 4.3	32.7 ± 3.0	29.3 ± 5.1

a Values are displayed as mean ± standard deviation. See Table 4.3 for abbreviations' descriptions.

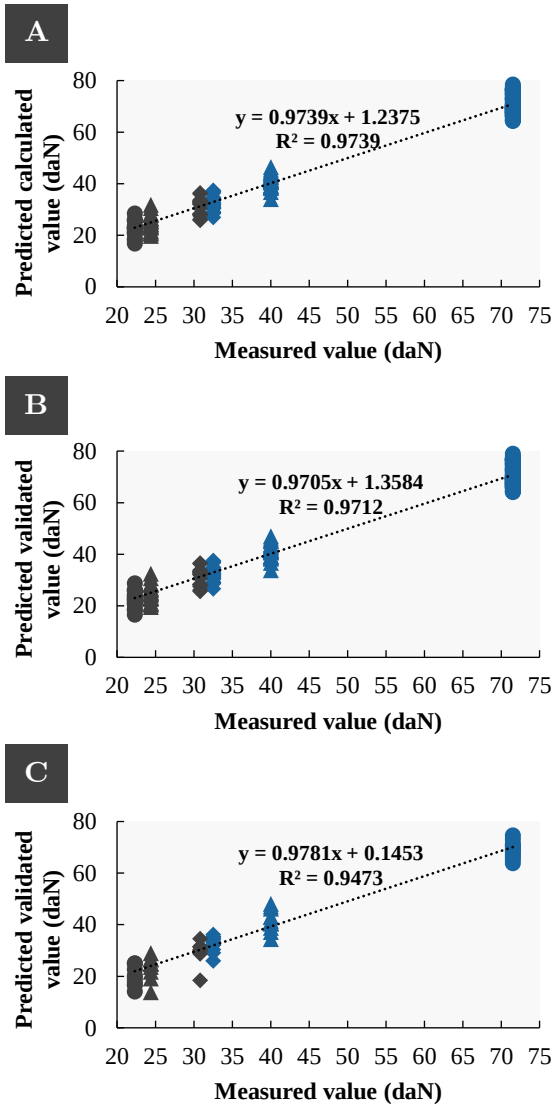


FIGURE 4.7 PLSR curves for the calibration step (A), the crossvalidation step (B), and the external validation step (C) of N-S (○), AD-S (△), M-S (◇) samples. Untreated and treated stoppers are represented with blue and dark grey filled symbols, respectively.

Additionally, the regression vector reveals which NIR bands are important in modelling the extraction force measurements[36]. As expected, the NIR bands observed in the regression vector (represented in FIGURE 4.8) were the same as those obtained in SIMCA's discriminating power plots, and were

attributed to silicone (1691, 1716, 1742, 1919-43, and 2285 nm), polyurethane binder (2130 nm; adhesive used to bind the particles in agglomerated corks) and paraffin (around 2230 nm)[29,30,35,39]. Our results suggest that the PLSR model built up can predict in a precise way the extraction force measurements based on the coating agents applied and the binder used to produce cork agglomerates.

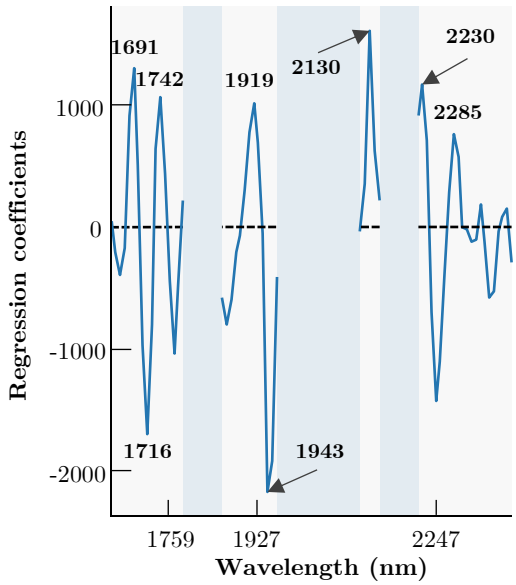


FIGURE 4.8 PLSR regression vector of transformed (MSC and second derivative, 7 points width) diffusive NIR spectra of treated and untreated still-wine cork stoppers (N-S, AD-S, and M-S samples). Spectral regions discarded when building up the PLSR model are displayed in light blue.

4.4 Conclusions

The evidence from the present work suggests that NIR spectroscopy combined with SIMCA provides a sensitive, rapid, and easy tool for the quality assurance of the surface treatments applied to cork stoppers. Significant chemical differences were detected between surface treated and untreated cork stoppers, which were linked to several functional groups belonging to paraffin and silicone coating agents. In this regard, it was found that a lack of homogeneity of the surface treatments applied could be assumed due to the presence of outliers detected in treated corks that were classified as untreated

ones. A 2-class SIMCA model was also developed including all cork stoppers treated in one class and those untreated in another, regardless of their variety. High classification rates for treated (97.73%) and untreated samples (93.33%) were displayed applying this model. Raman analysis was also conducted, and similar outcomes were obtained, confirming the band assignments attributed using NIR spectroscopy. Finally, PLSR models built up with NIR data showed good correlation with extraction force measurements performed, which were modelled through NIR bands linked to the coating agents applied and the binder used to produce the cork agglomerates. We believe that our findings will be valuable in solving the difficulty of assuring the presence and homogeneity of surface treatments applied to stoppers in a rapid way. This technique could also be implemented on-line, leading to a reduction of time and economic losses from the manufacturing perspective.

4.5 References

1. Silva, S. P. *et al.* Cork: properties, capabilities and applications. *International Materials Reviews* 53, 256–256 (2008).
2. Pereira, H. Wine and cork. in *Cork: Biology, Production and Uses* (ed. Pereira, H.) 305–327 (Elsevier B.V., 2007). doi:10.1016/B978-044452967-1/50016-9.
3. Jackson, R. S. Post-Fermentation Treatments and Related Topics. in *Wine Science* 535–676 (Elsevier Inc., 2014). doi:10.1016/B978-012373646-8.50011-1.
4. Pereira, H. Production of cork stoppers and discs. in *Cork: Biology, Production and Uses* (ed. Pereira, H.) 263–288 (Elsevier B.V., 2007). doi:10.1016/B978-044452967-1/50014-5.
5. Fugelsang, K. C., Callaway, D., Toland, T. & Muller, C. J. Coating Agents for Corks. *Wine Industry Journal* 12, 185–187 (1997).
6. Jung, R. & Schaefer, V. Reducing cork taint in wine. in *Managing Wine Quality* (ed. Reynolds, A. G.) 388–417 (Woodhead Publishing Limited, 2010). doi:10.1533/9781845699987.388.
7. Gonzalez-Adrados, J. R. *et al.* Control by ATR-FTIR of surface treatment of cork stoppers and its effect on their mechanical performance. *Wood Science and Technology* 46, 349–360 (2012).
8. Kizil, R. & Irudayaraj, J. Spectroscopic Technique: Fourier Transform Raman (FT-Raman) Spectroscopy. in *Modern Techniques for Food Authentication* (ed. Sun, D. W.) 193–217 (Elsevier, 2018). doi:10.1016/B978-0-12-814264-6.00006-2.
9. Ortega-Fernández, C., González-Adrados, J. R., García-Vallejo, M. C., Calvo-Haro, R. & Cáceres-Esteban, M. J. Characterization of Surface Treatments of Cork Stoppers by FTIR-ATR. *Journal of Agricultural and Food Chemistry* 54, 4932–4936 (2006).
10. González-Gaitano, G. & Ferrer, M. A. C. Definition of QC Parameters for the Practical Use of FTIR-ATR Spectroscopy in the Analysis of Surface Treatment of Cork Stoppers. *Journal of Wood Chemistry and Technology* 33, 217–233 (2013).
11. Prades, C., García-Olmo, J., Romero-Prieto, T., García de Ceca, J. L. & López-Luque, R. Methodology for cork plank characterization (Quercus suber L.) by near-infrared spectroscopy and image analysis. *Measurement Science and Technology* 21, 065602 (2010).

12. Prades, C., Gómez-Sánchez, I., García-Olmo, J., González-Hernández, F. & González-Adrados, J. R. Application of VIS/NIR spectroscopy for estimating chemical, physical and mechanical properties of cork stoppers. *Wood Science and Technology* 48, 811–830 (2014).
13. Sánchez-González, M., García-Olmo, J. & Prades, C. Correlation between porosity of cork planks before and after boiling using near infrared spectroscopy. *European Journal of Wood and Wood Products* 74, 509–517 (2016).
14. Porep, J. U., Kammerer, D. R. & Carle, R. On-line application of near infrared (NIR) spectroscopy in food production. *Trends in Food Science & Technology* 46, 211–230 (2015).
15. McClure, W. F. Introduction. in *Near-Infrared Spectroscopy in Food Science and Technology* (eds. Ozaki, Y., McClure, W. F. & Christy, A. A.) 1–10 (John Wiley & Sons, Inc., 2006). doi:10.1002/9780470047705.ch1.
16. Asociación Española de Normalización y Certificación. *UNE 56926. Three pieces cork stoppers. Test methods and specifications (in Spanish).* (AENOR, 2001).
17. Asociación Española de Normalización y Certificación. *UNE 56921. Natural cork stoppers for still wines. Test methods and specifications (in Spanish).* (AENOR, 2003).
18. Asociación Española de Normalización y Certificación. *UNE 56923. Agglomerated cork stoppers with natural cork discs for sparkling wines. Test methods and specifications (in Spanish).* (AENOR, 2006).
19. Asociación Española de Normalización y Certificación. *UNE 56933. Micro-agglomerated cork stoppers for still wines.* (AENOR, 2019).
20. Wold, S. Pattern recognition by means of disjoint principal components models. *Pattern Recognition* 8, 127–139 (1976).
21. Wold, S. & Sjöström, M. SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy. in *Chemometrics: Theory and Application* (ed. Kowalski, B. R.) 243–282 (ACS Symposium Series, 1977). doi:10.1021/bk-1977-0052.ch012.
22. Shah, N. K. & Gemperline, P. J. Combination of the Mahalanobis Distance and Residual Variance Pattern Recognition Techniques for Classification of Near-Infrared Reflectance Spectra. *Analytical Chemistry* 62, 465–470 (1990).

23. Bevilacqua, M. *et al.* Classification and Class-Modelling. in *Chemometrics in Food Chemistry* (ed. Marini, Federico.) vol. 28 171–192 (Elsevier B.V., 2013).
24. Osten, D. W. Selection of optimal regression models via cross-validation. *Journal of Chemometrics* 2, 39–48 (1988).
25. Weisberg, S. Weights, Lack of Fit, and More. in *Applied Linear Regression* 96–114 (John Wiley & Sons, Ltd, 2005). doi:10.1002/0471704091.ch5.
26. Giunchi, A., Versari, A., Parpinello, G. P. & Galassi, S. Analysis of mechanical properties of cork stoppers and synthetic closures used for wine bottling. *Journal of Food Engineering* 88, 576–580 (2008).
27. Sánchez-González, M. & Pérez-Terrazas, D. Assessing the percentage of cork that a stopper should have from a mechanical perspective. *Food Packaging and Shelf Life* 18, 212–220 (2018).
28. Prades, C., Gómez-Sánchez, I., García-Olmo, J. & González-Adrados, J. R. Discriminant Analysis of Geographical Origin of Cork Planks and Stoppers by Near Infrared Spectroscopy. *Journal of Wood Chemistry and Technology* 32, 66–85 (2012).
29. Workman, Jerry. & Weyer, Lois. Spectra–Structure Correlations—Labeled Spectra from 7200 cm−1 to 3800 cm−1 (1389 nm to 2632 nm). in *Practical guide to interpretive near-infrared spectroscopy* 165–206 (CRC Press, 2008).
30. Schwanninger, M., Rodrigues, J. C. & Fackler, K. A Review of Band Assignments in near Infrared Spectra of Wood and Wood Components. *Journal of Near Infrared Spectroscopy* 19, 287–308 (2011).
31. Gierlinger, N. & Schwanninger, M. Chemical Imaging of Poplar Wood Cell Walls by Confocal Raman Microscopy. *Plant Physiology* 140, 1246–1254 (2006).
32. Tshabalala, M. A., Jakes, J., VanLandingham, M. R., Wang, S. & Peltonen, J. Surface Characterization. in *Handbook of Wood Chemistry and Wood Composites* (ed. Rowell, R. M.) 217–254 (CRC Press, 2012). doi:10.1201/b12487-13.
33. Kvalheim, O. M. & Karstang, T. V. SIMCA - Classification by means of disjoint cross validated principal components models. in *MultiVariate Pattern Recognition in Chemometrics: Illustrated by Case Studies* (ed. Brereton, R. G.) 209–248 (Elsevier, 1992).

34. Dunn, W. J. & Wold, S. SIMCA Pattern Recognition and Classification. in *Chemometric Methods in Molecular Design* (ed. van de Waterbeemd, H.) 179–192 (VCH Publishers, Inc., 1995).
35. Cai, D., Neyer, A., Kuckuk, R. & Heise, H. M. Raman, mid-infrared, near-infrared and ultraviolet–visible spectroscopy of PDMS silicone rubber for characterization of polymer optical waveguide materials. *Journal of Molecular Structure* 976, 274–281 (2010).
36. Martens, Harald. & Næs, Tormod. *Multivariate calibration.* (John Wiley & Sons Ltd., 1989).
37. Westad, F., Bevilacqua, M. & Marini, F. Regression. in *Chemometrics in Food Chemistry* (ed. Marini, F.) vol. 28 127–170 (Elsevier B.V., 2013).
38. Santos, P. M., Pereira-Filho, E. R. & Rodriguez-Saona, L. E. Application of hand-held and portable infrared spectrometers in bovine milk analysis. *Journal of Agricultural and Food Chemistry* 61, 1205–1211 (2013).
39. Miller, C. E. & Eichinger, B. E. Analysis of Rigid Polyurethane Foams by Near-Infrared Diffuse Reflectance Spectroscopy. *Applied Spectroscopy* 44, 887–894 (1990).

Use of a portable ATR-FTIR spectrometer for microbiological applications

ABSTRACT

This chapter introduces the use of portable ATR-FTIR device combined with multivariate analysis in food microbiology. Spectra of *Escherichia coli* and yeast strains were collected and analysed by means of principal component analysis (PCA) in order discriminate them and analyse the biochemical differences among them. Soft independent modelling of class analogy (SIMCA) models were built up to classify *S. cerevisiae* and non-*Saccharomyces* yeasts. PCA models successfully discriminated *E. coli* isogenic variants and wine-making yeasts. Loading plots revealed that the main components of their differentiation were related to protein and DNA for *E. coli* mutants and β-glucans for yeast strains. Besides the good separation (interclass distance of 6.1), yeasts' SIMCA model displayed a good performance when classifying *S. cerevisiae* and non-*Saccharomyces* strains, with a high sensitivity and specificity (100%). Portable ATR-FTIR spectroscopy constitutes a promising and powerful technique for food microbiological applications.

5.1 Introduction

The emergence of new rapid and sensitive equipment such as gas chromatography and vibrational spectroscopy together with the development of molecular techniques has provided new approaches and tools for microbial characterization. Since the first mention to its use for biological applications in 1911 by Coblentz, FTIR spectroscopy remained dormant until the advent of the FT instrumentation in the 1980s and 1990s[1]. After the publication of the pioneer works of Naumann et al. (1991) and Helm et al. (1991), FTIR spectroscopy has been widely used (when combined with chemometric methods) as a tool for microbial characterizations[1–3]. When compared to cultured-based methods and molecular techniques, FTIR spectroscopy constitutes a fast technique, which can be easily implemented for quality control purposes since a minimum training for the operator is required[1,4]. However, microbial characterization by means of portable and miniaturized spectroscopic devices has been scarcely reported in the literature. In this chapter, two different applications will be introduced and discussed: 1) to get new insights in microbial resistance against essential oils (EOs) by the study of *E. coli* mutants and 2) to discriminate native yeast species isolated from wine must spontaneous fermentations.

EOs are complex mixtures of hundreds of individual constituents (ICs), which are synthetized by plants as a defence mechanism to different stressing conditions. Among the properties of EOs and their ICs (e.g., antifungal, insecticidal, antiviral, etc.), their antibacterial properties are the most reported, either when tested alone or in combination with other food processing technologies. Due to consumer concerns about food safety along with an upward trend towards sustainable and green consumerism, EOs have captured most of the attention of the scientific community[5–8].

Even though the development of bacteria resistance against EOs and ICs was initially discarded, several studies have confirmed its existence[9–12]. The emergence of resistant mutants against EOs as well as against other food preservation technologies may compromise food safety, thus their early detection would allow to design the most suitable food hygienisation techniques.

In previous work, three hyper-resistant mutants were isolated after subjecting *E. coli* MG1655 to long time exposures with subinhibitory

concentrations of the ICs carvacrol (CAR), citral (CIT), and limonene oxide (LIM), respectively[11]. Those evolved strains showed direct- and cross-resistance to other antimicrobials and food preservation treatments. Whole genome sequencing (WGS) of those isogenic strains allowed the identification of genetic changes responsible for the increased resistance. Some of those changes were minimal: for example, only 1 single nucleotide polymorphism (SNP), a 3-nucleotide deletion and 1-nucleotide insertion were detected in three different genomic regions of LIM strains[10]. The objectives of this study were 1) to differentiate wild-type *E. coli* (WT) and its hyper-resistant isogenic strains CAR, CIT, and LIM using a portable FTIR spectrometer, and 2) to evaluate the chemical changes on those strains and the relationship with their resistance increase.

The same portable FTIR spectrometer was tested for yeast characterization. Yeasts are unicellular fungi that play important roles in clinical microbiology since they can act as pathogens under some circumstances. In food science, the characterization of yeasts becomes a key factor when analysing the production of wine and bread or monitoring food spoilage. As in the study of bacterial resistance, FTIR spectroscopy becomes a good candidate for replacing the standard culture plating and molecular techniques used in yeast identification, which are time-consuming and require trained personnel[13,14]. FTIR spectroscopy has already been applied in the wine-making field for analysing and predicting parameters such as pH, acidity, and alcoholic strength[15]. It has also been used for monitoring fermentation processes and characterizing the autolysis of *S. cerevisiae* in sparkling wines[16,17]. Moreover, the type of yeast (most importantly in *non-Saccharomyces* yeasts), constitutes an important factor since it confers specific and singular organoleptic properties to the final product[18,19]. Therefore, the use of portable FTIR spectroscopic devices for yeast typing would be interesting for the wine sector as well as easy to implement. In this study, several yeasts were isolated from autochthonic spontaneous wine fermentations. The objectives of this work were 1) to discriminate all yeast strains isolated, and 2) to classify these strains into *Saccharomyces* and *non-Saccharomyces* species by means of a portable FTIR spectrometer combined with multivariate analysis.

5.2 Materials and methods

5.2.1 Strains

E. coli MG1655 wild type and its EO resistant mutants [isolated by Chueca et al. (2016)[11]] were analysed in this study. Briefly, *E. coli* mutant strains were isolated by culturing them daily in tryptic soy broth (Oxoid, Spain) and 0.6% yeast extract (TSBYE) containing sublethal concentrations [between 0.5- and 0.75-times the minimum inhibitory concentration (MIC)] of carvacrol, citral, and limonene oxide for 24 h at 37 °C. After 10 days, samples were pour-plated and single colonies were selected and stored at -80 °C for further analyses. The genotypic differences or mutations of the different derivative strains compared to the wild type are described in Table 5.1.

Table 5.1 Mutations of *E. coli* MG1655 derivative strains [adapted from Chueca et al. (2018)].

Strain	WT base	Mutant base(s)	Mutated genes	Result of mutation
CAR	—	—	Δ*(yhhJ-mokA)*	Deletion of 75 genes
	G	T	*soxR*	SoxR missense mutation
CIT	—	+A	*marR*	MarR frameshift
	A	C	Intergenic *eptB-yhjX*	-
	—	-GGT	*gyrB*	GyrB in-frame mutation
LIM	—	+C	*acrR*	ArcR frameshift
	—	—	Δ(*cheA-insA*)	Deletion of 3 genes
	—	—	Δ(*nrdA-enrB*)	Deletion of 9 genes

Yeast strains were kindly provided by Centre Tecnològic del VI (VITEC, Spain). Wild yeasts were isolated from spontaneous fermentations under simulated industrial conditions and using different grape varieties corresponding to different Spanish wine regions. The complete description of the methodology as well as the yeast typing procedure is described in previous work[20]. Ten strains (stored at -80 °C), either *S. cerevisiae* or non-*Saccharomyces*, were selected for this work, and are displayed in Table 5.2.

TABLE 5.2 Wild yeast strains analysed in this work.

Code	Yeast strain	PDO	Variety	Origin
SC1	*S. cerevisiae*	La Rioja	Tempranillo	Fermentation
SC2	*S. cerevisiae*	La Rioja	Tempranillo	Fermentation
SC3	*S. cerevisiae*	La Rioja	Tempranillo	Lees
SC4	*S. cerevisiae*	Terra Alta	Garnatxa Blanca	Lees
SC5	*S. cerevisiae*	Terra Alta	Garnatxa Blanca	Fermentation
SC6	*S. cerevisiae*	Rías Baixas	Mencía	Fermentation
NS	*non-Saccharomyces*	Penedès	Xarel·lo	Fermentation
MP	*Metschnikowia pulcherrima*	Priorat	Garnatxa Negra	Fermentation
DH	*Debaryomyces hansenii*	Terra Alta	Garnatxa Blanca	Fermentation
KT	*Kluyveromyces thermotolerans*	Terra Alta	Garnatxa Blanca	Fermentation

Abbreviations used: PDO, protected designation of origin.

5.2.2 Growth conditions

For *E. coli* strains, 5 mL tubes of sterile TSBYE were inoculated with one single colony from a tryptic soy agar (Oxoid) and 0.6% yeast extract (TSAYE) plate (incubated for 24 h at 37 °C) and incubated overnight at 37 °C. TSBYE flasks (250 mL) were then inoculated with the resulting subcultures to a final concentration of 10^5 CFU/mL for 24 h at 37 °C (final concentration of 10^9 CFU/mL). Four independent cultures of each *E. coli* strain were used for the analyses.

S. cerevisiae strains were recovered in yeast peptone dextrose (YPD) agar media (Panreac, Spain) plates and incubated for 48 h at 28 °C. Non-*Saccharomyces* yeasts were incubated in Lysine Agar medium (Oxoid) under the same conditions. Tubes with 10 mL of YPD broth were then inoculated with a single colony and incubated at 28 °C for 18 h more. Different volumes of the resulting subcultures were taken, and 7.5 mL of filter-sterilized Tempranillo red grape must (Sant Josep Agricultural Cooperative, Spain) were inoculated to obtain a concentration of 10^6 CFU/mL, which were then incubated for 48 h at 28 °C (final concentration of 10^8 CFU/mL). Two independent cultures of each yeast strain were used.

5.2.3 Sample preparation and FTIR analyses

Aliquots of 1 mL were taken from growth cultures and centrifuged at 10000 g for 5 min. The obtained pellets were washed twice with 0.1% peptone

water (prepared with bacteriological peptone, Oxoid) and centrifuged under the same conditions. Supernatants were discarded and the remaining pellets were used for IR analysis.

Aliquots of 0.6 µL from washed pellets were placed onto the sample stage of a portable infrared spectrometer Cary 630 (Agilent Technologies, Spain) equipped with a single bounce ATR diamond crystal accessory and a deuterated triglycine sulphate (DTGS) detector. Samples were vacuum dried prior to IR analysis to reduce the contribution of water. IR spectra were collected in the 4000-800 cm^{-1} region with 8 cm^{-1} resolution and 128 coadditions to improve the signal-to-noise ratio. Ten spectra were collected per pellet for bacteria analyses. Four spectra were acquired per pellet for yeast analyses.

5.2.4 Raman analyses

An aliquot of 1 µL washed pellets was placed onto a silicon support and allowed to dry at room temperature. Samples were analysed by inVia™ Confocal Raman Microscope (Renishaw, Spain) interfaced with an Argon laser (514 nm) and equipped with a motorized x-y stage and a Peltier-cooled CCD detector. The silicon supports were placed on the microscope stage and specific positions of the aliquots were selected with the assistance of a live camera (Leica DM 2500). Raman spectra were collected from 950-1800 cm^{-1} with a power of 25 mW (50% attenuated) and 50x magnification. A bleaching process of 100 s was performed before each spectrum collection (acquisition time of 60 s). The dispersive spectrometer was controlled using the software WiRE (Renishaw).

5.2.5 Spectra pre-processing

After mean-centring, FTIR spectra were transformed to their second derivative by applying a 13-point polynomial-fit Savitzky-Golay function. Vector length normalization was then performed to minimize the influence of signal intensity variation.

Raman spectra were uploaded to MATLAB 9.5 (Mathworks, US) and then denoised by using sym4 wavelets with two-level decomposition and soft minimax thresholding technique

5.2.6 Multivariate analysis

For both yeast and *E. coli* strains, unsupervised models were built up by means of Principal Component Analysis (PCA) to visualize data clustering patterns and to identify the chemical components responsible of their separation through the inspection of the principal components extracted.

Moreover, supervised models based on Soft Modelling of Class Analogy (SIMCA) algorithm were developed to classify *S. cerevisiae* and non-*Saccharomyces* yeasts. Class residuals and Mahalanobis distances were analyzed for outlier detection in both approaches. Pirouette software (Infometrix, US) was used to build up all chemometric models. The number of components selected for the SIMCA analysis was based on an F-test with 95% confidence on the reduced eigenvalues (i.e., latent factor variances)[21]. The performance of the models was also tested by means of an internal validation approach. The training and validation sets were randomly selected through an algorithm provided by the software. The training and validating sets consisted of 39 and 9 spectra for *S. cerevisiae* strains, and 26 and 6 spectra for non-*Saccharomyces* strains, respectively.

5.3 Results and discussion

5.3.1 Raw spectra of *E. coli* and yeast strains

Raw MIR spectra of *E. coli* and *yeast* strains are displayed in Figure 5.1 and Figure 5.2, respectively. Spectra of microbiological organisms can be divided into different spectral windows that display bands related to characteristic biochemical compounds[1].

Regarding bacterial strains, the region between 3100 and 2800 cm^{-1} displays several stretching vibrations (either symmetric or asymmetric) of methyl and methylene groups, which are related to fatty acid chains from various membrane amphiphiles (e.g., phospholipids). A weak band can be observed at around 1740 cm^{-1} assigned to carboxylic esters of lipids. The region between 1700 and 1500 cm^{-1} exhibits two intense bands assigned to the amide I and amide II bands of proteins. Complex patterns are obtained in between 1500 cm^{-1} and 1200 cm^{-1} since several compounds are active in that region, such as lipids and proteins (e.g., amide III band of proteins). Stretching vibrations of phosphate groups are also found at around 1230 cm^{-1} which can be attributed to the phosphodiester groups of DNA and RNA and to the head

groups of phospholipids. Contributions from nucleic acids and polysaccharides are observed in the region between 1200 and 900 cm^{-1}, with stretching vibrations of phosphate and C-O-C groups, respectively. Finally, a variety of weak and complex features can be detected in the region of 900-600 cm^{-1}, which are attributed to several amino acid sidechains (e.g., tyrosine, phenylalanine, etc.), nucleotides (ring vibrations) and fatty acids (rocking vibrations of methylene groups)[1,4,22–26].

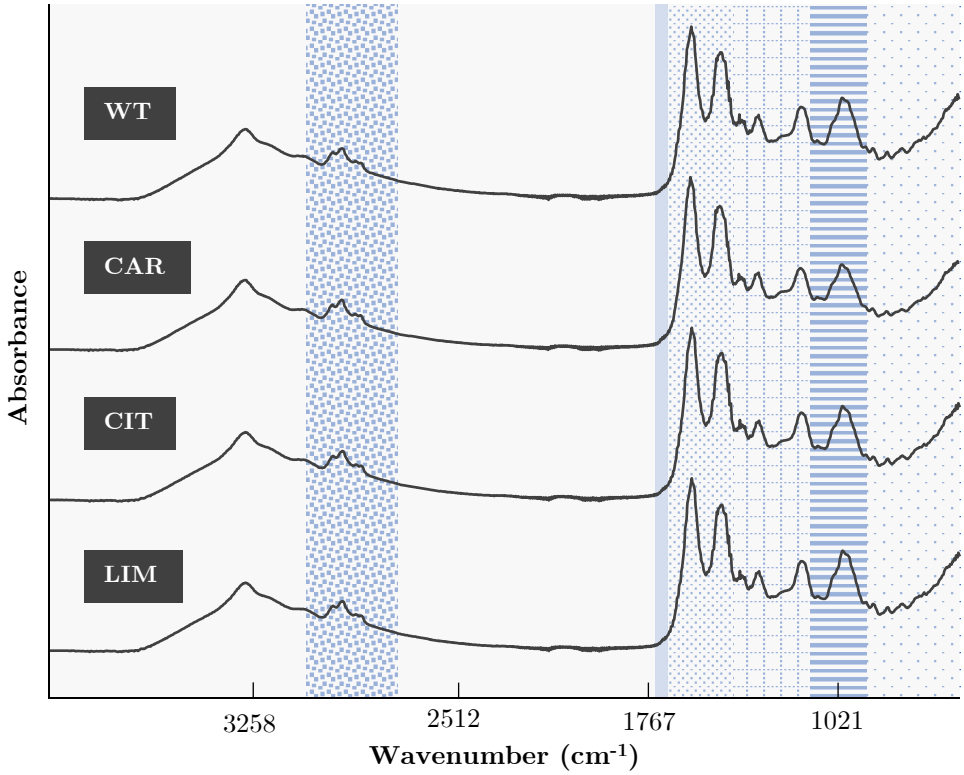


FIGURE 5.1 Raw MIR spectra of *E. coli* MG1655 wild-type (WT) and EO resistant isogenic variants (CAR, CIT and, LIM).

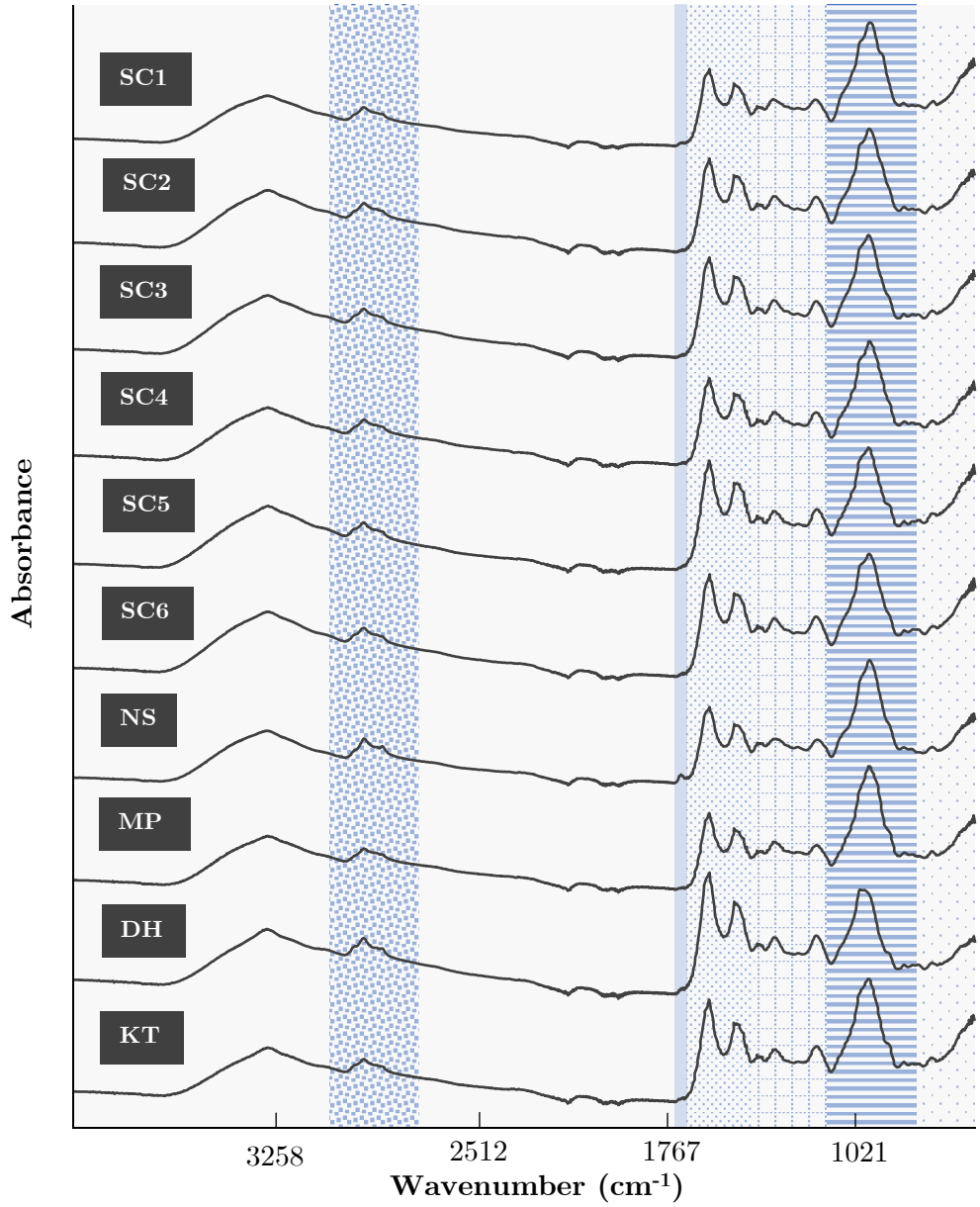


FIGURE 5.2 Raw MIR spectra of wine-making yeast strains.

Raman spectra of *E. coli* strains also provided bands linked to the same biochemical compounds (Figure 5.3) but arising from different vibrations, as mentioned before. The Raman bands located at 1655, 1580 and 1328 cm^{-1} can be associated with the amide I, amide II, and amide III regions of proteins, respectively. Moreover, other bands can be found at 1178 and 1003 cm^{-1}, which can be assigned to protein amino acids such as phenylalanine. Other component contributions can also be observed at 1452 cm^{-1}, corresponding to the methylene deformation Raman bands of lipids. Lastly, the bands centered at 1580 (note the double assignation) and 1484 cm^{-1} may be attributed to nucleic acids, while the Raman band at 1091 cm^{-1} to phosphodiester groups of DNA[22,27–30].

Similar IR spectral features can be observed in yeast raw spectra (Figure 5.2). However, a major distinction can be made when comparing both types of microorganisms. Yeast spectra are dominated by an intense band in the region of 1200-900 cm^{-1}, while the amide I and amide II bands have a weaker signal[31]. This is in agreement with differences in cell structure between bacteria and yeasts. Yeasts present a cell wall composed of complex polysaccharides namely β-glucans that consist of β-(1→3) and β-(1→4) linked mannose and glucose monosaccharides (other polysaccharides found in yeast cell walls are chitosan and chitin). Some proteins can also be found in the cell wall linked to mannose units. On the other hand, the cell envelope of *E. coli* cells (gram-negative bacteria) is constituted of two different membranes. The inner or cytoplasmic membrane consists of phospholipids, while the outer one differs from the former regarding the high content of lipopolysaccharides facing outwards with proteins embedded in between them[1].

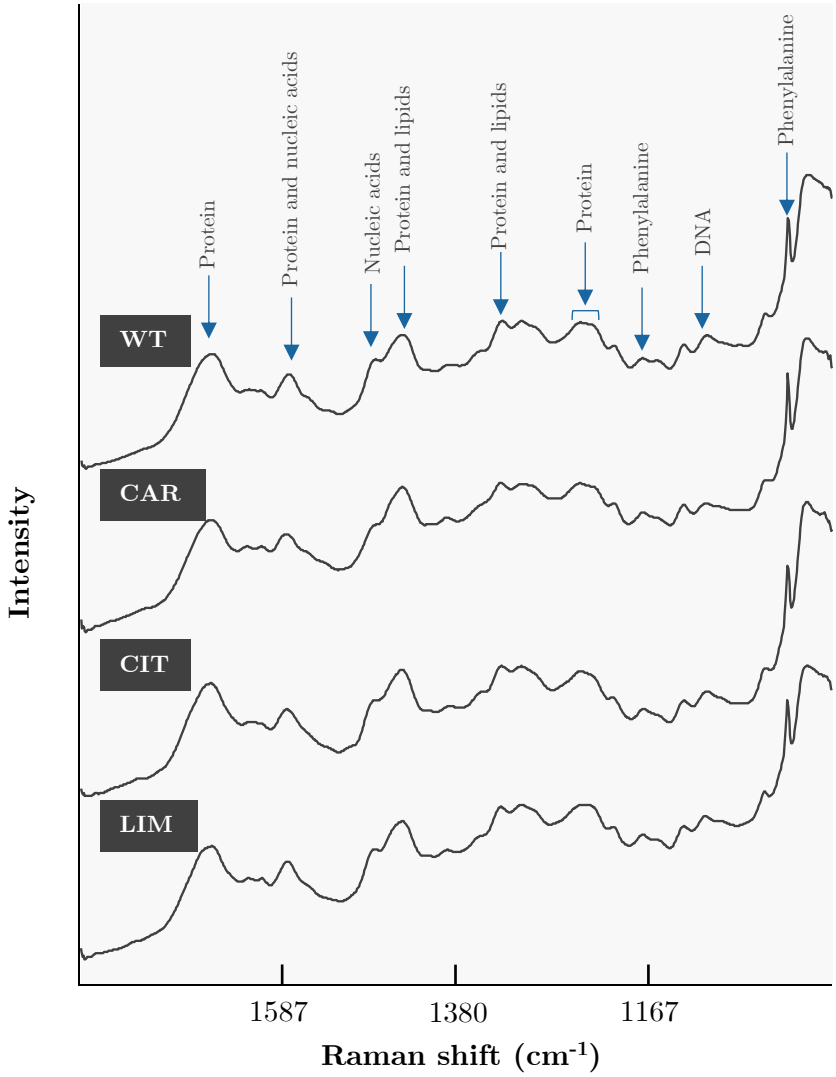


FIGURE 5.3 Raman spectra of *E. coli* MG1655 wild-type (WT) and EO resistant isogenic variants (CAR, CIT and, LIM).

5.3.2 Exploratory analysis of bacterial and yeast strains

PCA score plots of *E. coli* parental strain (WT) versus each resistant mutant (CAR, CIT, and LIM) are depicted in FIGURE 5.4 (A, B, and C, respectively). Such plots aided to visualize sample separation, with each data point representing one sample spectrum. Good separation from WT was

obtained with respect to CAR and CIT. However, LIM and WT exhibited overlapping clusters and thus poor discrimination. Even though not all the mutants were successfully discriminated, our results are in line with the genetic changes experimented by each of them (Table 5.1). CIT and LIM exhibited small changes in their genomes corresponding to genes that control the expression of proteins involved in antibiotic resistance and multidrug transport (*marR* and *acrR*, respectively). On the contrary, CAR presented a major deletion of 75 genes and a mutation in the *soxR* gene, which encodes for a protein responsible of activating the transcription of many other genes related to defence against oxidative damage. Therefore, it would be expected that the phenotype (i.e., biochemical composition) of CAR cells differs significantly from the WT. This is in line with our results, in which CAR and WT show the greatest separation among all the models developed. Moreover, it can also be inferred from Figure 5.4 that LIM exhibits less phenotypical changes than CIT since the separation between CIT-WT is higher than LIM-WT[10].

Useful information can also be obtained from the loading plots. Loadings give information about which variables are more influential for sample separation. Figure 5.5 shows the loadings for the first principal component of each pairwise model. Interestingly, most of the bands obtained were the same for all models. Therefore, the phenotypical changes for all mutants can be linked to similar biochemical components of the cell. The IR bands located at 1711 and 1215, and 961 cm^{-1} and the region of 900-800 cm^{-1} can be assigned to several functional groups of DNA/RNA and nucleic acids, respectively. Moreover, bands attributed to proteins can be found, such as the ones in 1645-1610 cm^{-1} and 1530 cm^{-1} that can be assigned to the amide I and amide II regions, respectively. Lipid contributions can also be observed, displaying bands centered in 1744, 1379, and 1215 cm^{-1} (note the double assignation) mainly attributed to membrane lipids. Finally, carbohydrate bands are found in the region of 1200-900 cm^{-1}, with main contributions being the band located at around 1150 cm^{-1} and the one at 1050 cm^{-1}[1,4,23–26,32]. Therefore, genetic differences attributed to the bands assigned to DNA, RNA, and nucleic acids were detected, which in turn triggered structural and biochemical changes on the cells. The structural and biochemical modifications observed in resistant mutants could be linked to components of the cell envelope, such as phospholipids. Modification of cell envelopes would avoid access of antimicrobial molecules into the bacterial cytoplasm, making the mutants more resistant against antimicrobials. Moreover, due to the changes

in the DNA coding sequence, changes in protein structures were also identified, which could be either responsible for the control of the expression over other genes or be building blocks of efflux pumps which regulate the inner cell concentration of certain drugs and antibiotics.

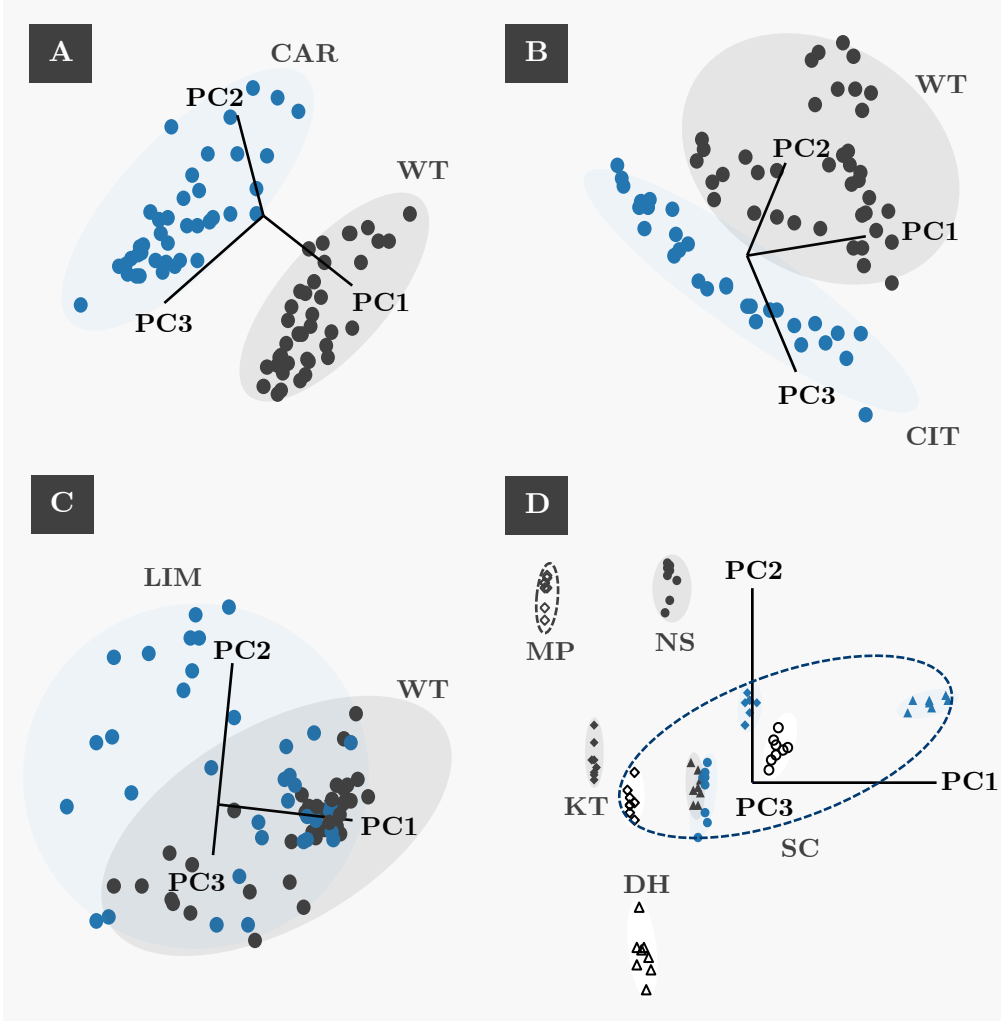


FIGURE 5.4 SIMCA class projection plots of transformed (second derivative, 13-points window) ATR-FTIR spectra of CAR (A), CIT (B), LIM (C) versus *E. coli* MG1655 WT, and wild yeast strains (D).

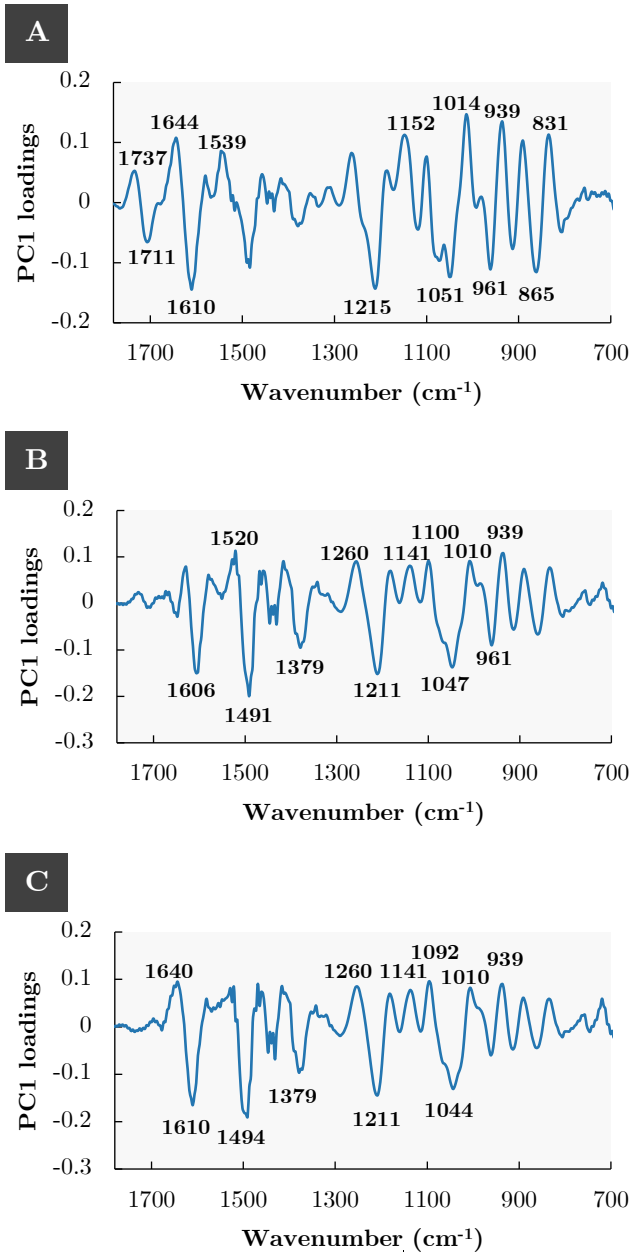


FIGURE 5.5 Loading plots corresponding to the first factor of CAR (A), CIT (B), and LIM (C) vs. WT PCA pairwise models.

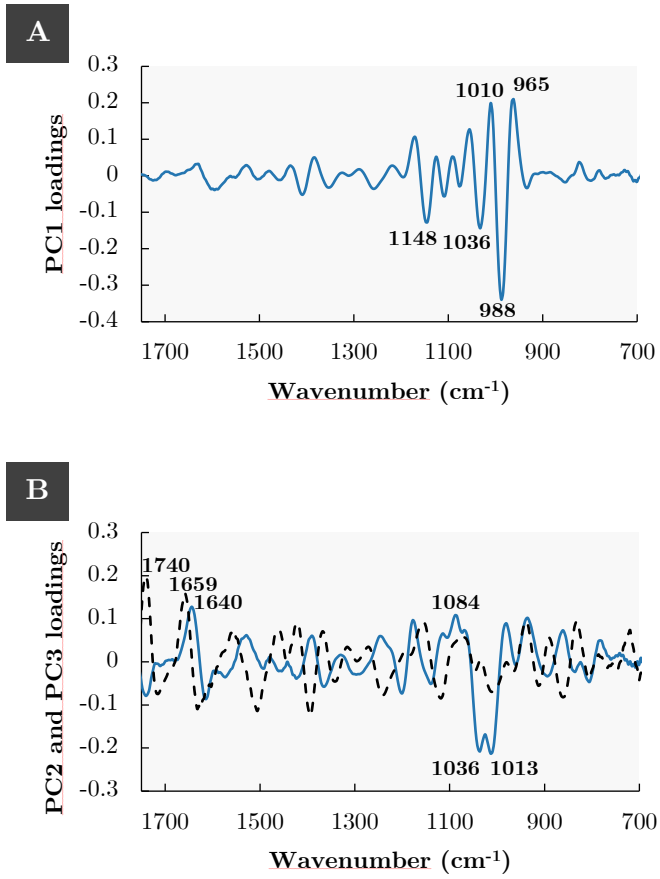


FIGURE 5.6 Loading plots for the yeast model corresponding to the first factor (A), and second (blue line) and third factors [dashed line, (B)].

Considering yeast strains, all clusters were well separated, as it is shown in FIGURE 5.4. A certain degree of separation between *S. cerevisiae* and non-*Saccharomyces* strains was also observed, which gives the option of developing classification models for their identification. Loading plots of the three first principal components are depicted in FIGURE 5.6. The first principal component gathers most of the variance of the model with the bands being mainly related to polysaccharides from the cell wall (i.e., 1200-900 cm^{-1} region). The most prominent band, centered at 988 cm^{-1}, is related to β-(1→6) glucans. Other bands, such as the ones located at 1036 and 1010 cm^{-1}, are attributed to β-(1→4) glucans. Other band linked to carbohydrates can also be found at 1148

cm^{-1}, and the one centered at 1084 cm^{-1} that can be assigned to β-(1→3) glucans. By examining the loadings of the second and third principal components several IR bands linked to other compounds than sugars can be observed, such as the amide I region bands (1660-1640 cm^{-1}) and the band of carboxylic esters of lipids band 1740 cm^{-1}[16,17,33,34]. These results suggest that the main differences among all strains studied in this work might come from cell wall components such as glucans and mannans.

5.3.3 Discriminating *S. cerevisiae* and non-*Saccharomyces* yeasts by a supervised chemometric approach

A SIMCA model was built up to classify *S. cerevisiae* and non-*Saccharomyces* yeast strains. In SIMCA, PCA models are developed for each class (i.e., *S. cerevisiae* and non-*Saccharomyces* yeasts). Samples from every class are fitted into each of the models and the residuals are then computed, which are used for calculating the ICD value[35]. The so-called Coomans plot depicts the residuals for each sample (i.e., spectrum). Figure 5.7 A displays the Coomans plot obtained by SIMCA for *S. cerevisiae* and non-*Saccharomyces* yeast strains. Low residuals are expected in data points that belong to the class of the model while high ones are expected for the rest. Good separation was obtained between *S. cerevisiae* and non-*Saccharomyces* yeasts, with smaller groupings depicting the different strains analysed in this work (i.e., SC1, SC2, and so on). Moreover, the interclass distance (ICD) value, which is a ratio of interclass to intraclass distance, had a value of 6.1. As a rule of thumb, clusters with ICD values higher than 3 are considered significantly different[36]. Therefore, it can be concluded that *S. cerevisiae* and non-*Saccharomyces* strains were successfully discriminated by a portable ATR-FTIR device combined with SIMCA.

Furthermore, as in PCA, the variables of importance can be assessed by examining the discriminating power plots (Figure 5.7 B). The most important variable for the discrimination of *S. cerevisiae* from non-*Saccharomyces* strains is the one located at 988 cm^{-1}, which is attributed to β-(1→6) glucans, as aforementioned. The rest of the bands displayed in Figure 5.7 B can linked to the lipid content of the yeasts. Bands at 2940, 2874 and 2844 cm^{-1} can be attributed to stretching vibrations of methyl and methylene groups of lipids[16,17,33,34]. Moreover, the region of 1770-1744 cm^{-1} could be assigned to carboxylic groups of lipids. These results suggest that, besides the carbohydrate composition of the cell wall, lipids are the major biochemical

constituent being responsible of discriminating between *S. cerevisiae* and non-*Saccharomyces* yeasts, which is in accordance with previous literature[18].

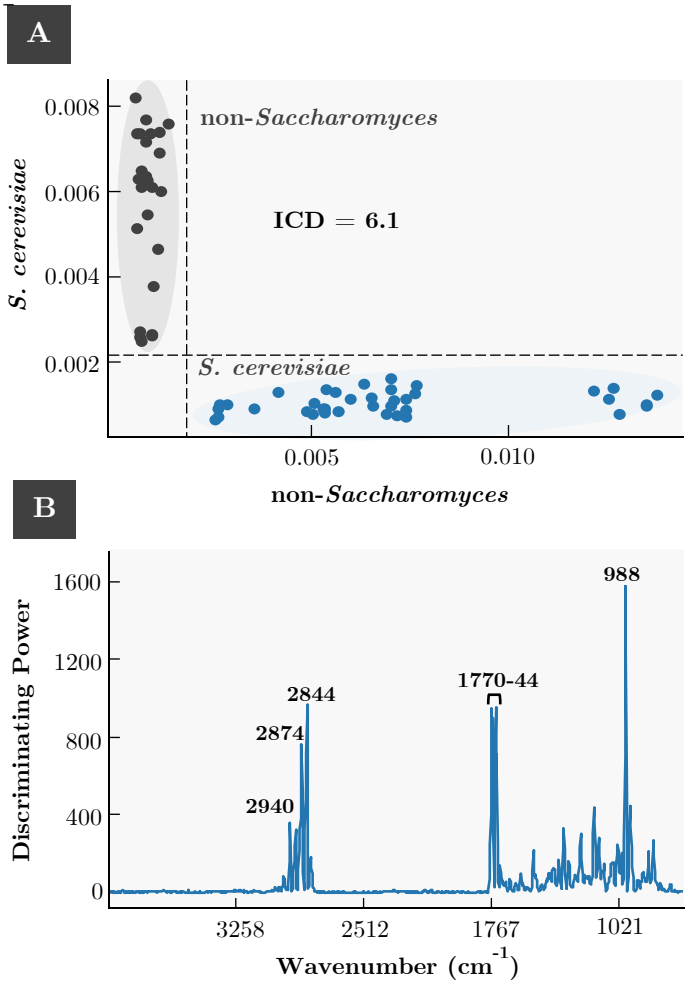


FIGURE 5.7 SIMCA Coomans (A) and discriminating power (B) plots of transformed (second derivative, 13-points window) ATR-FTIR spectra of yeast strains.

Finally, the performance of the SIMCA model was tested through internal validation. Sensitivity and specificity values were considered as figures of merit for assessing the quality of the model. Sensitivity is described as the percentage of objects belonging to a certain class that are correctly classified by its own model, while specificity accounts for the percentage of objects from other classes that are correctly rejected by the same model[37]. Sensitivity and

specificity values of 100% were obtained for *S. cerevisiae* (9/9 spectra) and non-*Saccharomyces* (6/6 spectra). Therefore, it is possible to conclude that the model developed could discriminate *S. cerevisiae* and non-*Saccharomyces* yeasts.

5.4 Conclusions

In this work, a portable ATR-FTIR equipment combined with powerful chemometric algorithms was tested for two different applications in the microbiological field. Such approach demonstrated to be partially successful for the discrimination of bacteria beyond the strain level, which is scarcely reported in the literature. Even though not all strains were discriminated through PCA, the differences observed from IR data among the EO resistant *E. coli* mutants were consistent for all of them with the WGS studies previously reported, confirming the genetic changes and expression of different proteins when compared to the wild type. These results provide complementary information regarding the mechanism of resistance in *E. coli*, which could be valuable for designing the most suitable food preservation techniques. Considering the application in yeast typing, good separation was obtained by means of PCA for all the strains tested, with the components being responsible of such differentiation mainly related to biochemical components of the cell wall. Moreover, SIMCA models were built up to determine the capability of the portable ATR-FTIR spectrometer for classifying *S. cerevisiae* and non-*Saccharomyces* yeasts. Through an internal validation approach, 100% sensitivity and specificity values were obtained. Even though the results presented seem promising, further work needs to be done by including a higher number of strains and proving the performance of the models by external validation.

5.5 References

1. Naumann, D. Infrared Spectroscopy in Microbiology. *Encyclopedia of Analytical Chemistry* (2006) doi:10.1002/9780470027318.a0117.
2. Naumann, D., Helm, D. & Labischinski, H. Microbiological characterizations by FT-IR spectroscopy. *Nature* 351, 81–82 (1991).
3. Helm, D., Labischinski, H., Schallehn, G. & Naumann, D. Classification and identification of bacteria by Fourier-transform infrared spectroscopy. *Microbiology (N Y)* 137, 69–79 (1991).
4. Yu, C. & Irudayaraj, J. Spectroscopic characterization of microorganisms by Fourier transform infrared microspectroscopy. *Biopolymers* 77, 368–377 (2005).
5. Burt, S. Essential oils: Their antibacterial properties and potential applications in foods - A review. *International Journal of Food Microbiology* 94, 223–253 (2004).
6. Pandey, A. K., Kumar, P., Singh, P., Tripathi, N. N. & Bajpai, V. K. Essential oils: Sources of antimicrobials and food preservatives. *Frontiers in Microbiology* 7, (2017).
7. Hyldgaard, M., Mygind, T. & Meyer, R. L. Essential oils in food preservation: Mode of action, synergies, and interactions with food matrix components. *Frontiers in Microbiology* 3, (2012).
8. Calo, J. R., Crandall, P. G., O'Bryan, C. A. & Ricke, S. C. Essential oils as antimicrobials in food systems - A review. *Food Control* 54, 111–119 (2015).
9. Berdejo, D., Pagán, E., Merino, N., García-Gonzalo, D. & Pagán, R. Emerging mutant populations of Listeria monocytogenes EGD-e under selective pressure of Thymbra capitata essential oil question its use in food preservation. *Food Research International* 145, (2021).
10. Chueca, B. *et al.* Whole-Genome Sequencing and Genetic Analysis Reveal Novel Stress Responses to Individual Constituents of Essential Oils in Escherichia coli. *Applied and Environmental Microbiology* 84, (2018).
11. Chueca, B., Berdejo, D., Gomes-Neto, N. J., Pagán, R. & García-Gonzalo, D. Emergence of hyper-resistant Escherichia coli MG1655 derivative strains after applying sub-inhibitory doses of individual constituents of essential oils. *Frontiers in Microbiology* 7, (2016).

12. Berdejo, D. *et al.* Sub-Inhibitory doses of individual constituents of essential oils can select for staphylococcus aureus resistant mutants. *Molecules* 24, (2019).
13. Santos, C., Fraga, M. E., Kozakiewicz, Z. & Lima, N. Fourier transform infrared as a powerful technique for the identification and characterization of filamentous fungi and yeasts. *Research in Microbiology* 161, 168–175 (2010).
14. Kümmerle, M., Scherer, S. & Seiler, H. Rapid and Reliable Identification of Food-Borne Yeasts by Fourier-Transform Infrared Spectroscopy. *Applied and Environmental Microbiology* 64, 2207–2214 (1998).
15. Puxeu, M., Andorra, I. & de Lamo-Castellví, S. Monitoring Saccharomyces cerevisiae Grape Must Fermentation Process by Attenuated Total Reflectance Spectroscopy. *Food and Bioprocess Technology* 8, 637–646 (2015).
16. Cavagna, M., Dell'Anna, R., Monti, F., Rossi, F. & Torriani, S. Use of ATR-FTIR microspectroscopy to monitor autolysis of Saccharomyces cerevisiae cells in a base wine. *Journal of Agricultural and Food Chemistry* 58, 39–45 (2010).
17. Burattini, E. *et al.* A FTIR microspectroscopy study of autolysis in cells of the wine yeast Saccharomyces cerevisiae. *Vib Spectrosc* 47, 139–147 (2008).
18. Mbuyane, L. L., Bauer, F. F. & Divol, B. The metabolism of lipids in yeasts and applications in oenology. *Food Research International* 141, (2021).
19. Grangeteau, C. *et al.* FT-IR spectroscopy: A powerful tool for studying the inter- and intraspecific biodiversity of cultivable non-Saccharomyces yeasts isolated from grape must. *Journal of Microbiological Methods* 121, 50–58 (2016).
20. Puxeu, M. Application of infrared spectroscopy in mid-infrared range combined with multivariate analysis to study yeasts involved in wine production. (Universitat Rovira i Virgili, 2015).
21. Malinowski, E. R. StatisticalF-tests for abstract factor analysis and target testing. *Journal of Chemometrics* 3, 49–60 (1989).
22. Socrates, G. *Infrared and Raman characteristic group frequencies. Infrared and Raman characteristic group frequencies* (John Wiley & Sons, Inc., 2004). doi:10.1002/jrs.1238.

23. Garip, S., Gozen, A. C. & Severcan, F. Use of Fourier transform infrared spectroscopy for rapid comparative analysis of Bacillus and Micrococcus isolates. *Food Chemistry* 113, 1301–1307 (2009).
24. Al-Qadiri, H. M., Lin, M., Cavinato, A. G. & Rasco, B. A. Fourier transform infrared spectroscopy, detection and identification of Escherichia coli O157:H7 and Alicyclobacillus strains in apple juice. *International Journal of Food Microbiology* 111, 73–80 (2006).
25. Alvarez-Ordóñez, A., Mouwen, D. J. M., López, M. & Prieto, M. Fourier transform infrared spectroscopy as a tool to characterize molecular composition and stress response in foodborne pathogenic bacteria. *Journal of Microbiological Methods* 84, 369–378 (2011).
26. Al-Qadiri, H. M. *et al.* Rapid detection and identification of Pseudomonas aeruginosa and Escherichia coli as pure and mixed cultures in bottled drinking water using fourier transform infrared spectroscopy and multivariate analysis. *Journal of Agricultural and Food Chemistry* 54, 5749–5754 (2006).
27. Huang, W. E., Griffiths, R. I., Thompson, I. P., Bailey, M. J. & Whiteley, A. S. Raman microscopic analysis of single microbial cells. *Analytical Chemistry* 76, 4452–4458 (2004).
28. Huang, W. E., Li, M., Jarvis, R. M., Goodacre, R. & Banwart, S. A. Shining Light on the Microbial World: The Application of Raman Microspectroscopy. in *Advances in Applied Microbiology* vol. 70 153–186 (Academic Press Inc., 2010).
29. Maquelin, K. *et al.* Raman spectroscopic method for identification of clinically relevant microorganisms growing on solid culture medium. *Analytical Chemistry* 72, 12–19 (2000).
30. Xie, C. & Li, Y. Q. Confocal micro-Raman spectroscopy of single biological cells using optical trapping and shifted excitation difference techniques. *Journal of Applied Physics* 93, 2982–2986 (2003).
31. Shapaval, V. *et al.* FTIR spectroscopic characterization of differently cultivated food related yeasts. *Analyst* 138, 4129–4138 (2013).
32. Maquelin, K. *et al.* Identification of medically relevant microorganisms by vibrational spectroscopy. *Journal of Microbiological Methods* 51, 255–271 (2002).
33. Mihoubi, W., Sahli, E., Gargouri, A. & Amiel, C. FTIR spectroscopy of whole cells for the monitoring of yeast apoptosis mediated by p53 over-

expression and its suppression by Nigella sativa extracts. *PLoS ONE* 12, (2017).

34. Kuligowski, J., Quintás, G., Herwig, C. & Lendl, B. A rapid method for the differentiation of yeast cells grown under carbon and nitrogen-limited conditions by means of partial least squares discriminant analysis employing infrared micro-spectroscopic data of entire yeast cells. *Talanta* 99, 566–573 (2012).
35. Wold, S. & Sjöström, M. SIMCA: A Method for Analyzing Chemical Data in Terms of Similarity and Analogy. in *Chemometrics: Theory and Application* (ed. Kowalski, B. R.) 243–282 (ACS Symposium Series, 1977). doi:10.1021/bk-1977-0052.ch012.
36. Dunn, W. J. & Wold, S. SIMCA Pattern Recognition and Classification. in *Chemometric Methods in Molecular Design* (ed. van de Waterbeemd, H.) 179–192 (VCH Publishers, Inc., 1995).
37. Bevilacqua, M. *et al.* Classification and Class-Modelling. in *Chemometrics in Food Chemistry* (ed. Marini, Federico.) vol. 28 171–192 (Elsevier B.V., 2013).

Conclusions and future work

UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

6.1 Conclusions

A main conclusion can be formulated from the results presented in this thesis: it has been proven that miniaturized vibrational spectroscopy equipment combined with multivariate analysis can successfully be implemented as a powerful discriminating and predicting tool in several food applications.

More specific conclusions can also be described:

1. Portable ATR-FTIR spectroscopy combined with soft independent modelling of class analogy successfully monitored the course of Maillard reaction when producing NaCAS-GA conjugates. Different functional groups linked to Maillard Reaction Products (MRPs) such as Schiff's base and pyridine derivatives were identified during the reaction
2. IR data confirmed that NaCAS-CMC mixtures did not undergo Maillard reaction under the conditions tested (60 °C at 76-79% humidity for 72 h), which were in accordance with the results obtained through the standard methods of reference (browning degree, stability tests, and conjugation efficiency)
3. The chemometric analysis based on SIMCA of MIR data from seven commercial edible insect powders successfully discriminated them by species and origin. Their main differences were linked to components such as lipids and chitin
4. Internal model validation gave 100% correct predictions for *T. molitor*, *Gryllodes sigillatus*, *Acheta domesticus*, and *Locusta migratoria* species
5. SIMCA models developed with NIR data successfully determined the presence of surface treatments applied to wine cork stoppers, regardless of their variety. Handheld NIR spectroscopy combined with chemometrics could also detect those stoppers that were not properly treated due to homogeneity issues, thus saving time and economic losses for the industry. Raman further confirmed the chemical agents

responsible for the discrimination between treated and untreated cork stoppers

6. The extraction force was accurately predicted for still wine cork stoppers by means of a PLSR model built up with NIR data, with a high coefficient of determination and a low standard error
7. *E. coli* evolved mutants were partially separated from the wild type through PCA of FTIR data. PCA loading plots revealed the biochemical changes mainly linked to DNA, proteins, and lipids, which were consistent to the genetic mutations observed in the derivative strains by Wide Genome Sequencing (WGS) techniques
8. A PCA model created with MIR data successfully differentiated wine-making yeast strains. The components responsible of such differentiation were related to glucans and mannans from the yeasts' cell wall
9. Good classification performances were obtained for *S. cerevisiae* and non-*Saccharomyces* yeast strains, achieving sensitivity and specificity values of 100% during the validation step. Biochemical differences between *S. cerevisiae* and non-*Saccharomyces* yeasts were observed and linked to the lipid composition of yeast cell membranes, which is in accordance with previous literature

6.2 Future work

As a result of the research conducted in this thesis, possible future work is outlined below:

1. To explore new applications of portable FTIR spectroscopy equipment in Maillard reaction characterization by determining the amount of conjugate produced during the process, which can be used as an emulsifying agent in food products
2. The insect sector can benefit from vibrational spectroscopy as a rapid method for insect typing and characterization, as well as for the quantification of its nutritional components. Future

work should concentrate on considering factors such as feed, sex, and stage when discriminating them

3. Since insects are being considered as a novel ingredient in food design, the feasibility of this FTIR spectroscopy can also be further evaluated for the detection of insect powders as well as their protein and lipid fractions in several food matrixes
4. To explore the online implementation of NIR spectroscopy to control the quality of the surface treatments applied to closures in the cork industry
5. Even though the capabilities of portable FTIR spectroscopy and chemometrics for microbial typing (beyond the species level) have been proven, new multivariate models should be developed by including a higher number of strains

UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

UNIVERSITAT ROVIRA I VIRGILI
EVALUATION OF PORTABLE VIBRATIONAL SPECTROSCOPY EQUIPMENT COMBINED WITH MULTIVARIATE ANALYSIS FOR FOOD APPLICATIONS
Jorge Mellado Carretero

UNIVERSITAT
ROVIRA i VIRGILI