We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS
BOOK CITATION INDEX
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Chapter

Discrimination by Infrared Spectroscopy: Application to Micronized Locust Bean and Guar Gums

Catherine Rébufa, Fabien Girard, Jacques Artaud and Nathalie Dupuy

Abstract

The authentication of locust bean and guar powder gums requires usually the use of sophisticated and time-consuming analytical techniques. There is a need for fast and simple analytical techniques for the objective of a quality control methodology. Commercial locust bean and guar micronized powder gums present characteristic MIR spectra. Principal component analysis of the infrared spectra of these micronized powder gums allowed to distinguish locust bean from guar samples and to perform good classification results. The prediction of the two varieties was done without any ambiguity with a partial least square regression-discriminant analysis (PLS-DA). A simplex approach was used to generate binary blends mathematically taking into account the intrinsic variability of chemical composition of commercial products. The simulated spectral profiles allowed to develop predictive model of the percentage of gums in blends.

Keywords: locust bean, carob, guar, infrared, chemometrics, classification

1. Introduction

In food industry, it is common to add ingredients that have in aqueous solution thickening, stabilizing, or gelling properties to the formulation of food products. These ingredients are interesting for the appearance and organoleptic characteristics but can also become innovative ingredients in the search for original textures or to support the use of alternative chemical additives. In bioprocessing, the field of ingredients or additives authorized in this framework is more limited than in conventional processing. However, a number of gums, starches, fibers, or flours can be used. Among them, two vegetable gums, extracted from the endosperm of legume seeds, were commonly used with similar properties: guar gum and locust bean gum. These natural polysaccharides are extensively used in a wide range of applications in food, medical, pharmaceutical, textile, paper, hydraulic fracturing, explosives, agriculture, cosmetic, bioremediation, and petroleum industries because of their ability to modify the rheological properties and in the aim of green chemistry approach [1–5]. Guar gum (GG), named E412 in European

additive list, is a polysaccharide of natural origin, extracted from the seed of *Cyamopsis tetragonoloba* L. (Fabaceae), a plant also called guar or guar bean and native. The world's production of guar is concentrated in India, Pakistan, and the United States with limited amounts grown in South Africa and Brazil. The annual guar plant grows to about 0.6–1 m in height and produces seed pods growing in clusters giving guar pods the common name cluster bean (6–9 guar beans per pod). Locust bean gum (LBG) or E410 is the crushed endosperm of locust bean tree seeds, a dioecious evergreen tree grown in a Mediterranean climate, botanically known as *Ceratonia siliqua* L., belonging to the subfamily *Caesalpinioideae* of the *Fabaceae* family. It is abundant in Spain, Italy, and Cyprus and is also found in other Mediterranean countries, in various regions of North Africa, South America, and Asia. It can reach 8–15 m in height and live up to 500 years. Its fruits are long thick and tough pods containing 10–15 oval-shaped locust bean seeds or kernels. The locust tree can produce annually 300–800 kg of locust bean seeds from which LBG will be produced, also referred as locust bean seed gum, locust bean flour, or *ceratonia*. The guar seed is smaller than the locust bean seed but they both have the same structure. They consist of four elements (**Figure 1**): the tegument (outer husk or seed coat), a translucent endosperm representing 40–50% of the weight of the locust bean seed and 35–45% of the weight of the guar seed, two cotyledons, and an embryo (or germ) [6].

The number of tissues within the locust bean would even be 10 according to microscopic studies [7]. These seeds are a basic material for the manufacture of gum and contain hydrocolloids (polysaccharides), called galactomannans, which serve as a reserve for the embryo during germination. With the help of various thermal, mechanical, or chemical processes [2, 5], the seeds are dehusked without damaging the endosperm and the embryo (germ). After this peeling process, the endosperm is split from the cotyledons and then it is ground to produce gum. To eliminate protein content and impurities for certain industrial applications, the gum is purified by washing with solvent or dispersing in boiling water, followed by filtering, evaporation, and drying [2, 5, 6, 8]. Galactomannans are heterogeneous polysaccharides with a high molecular weight, composed by linear chain of β-(1-4)-D-mannopyranosyl units with a single α-D-galactopyranosyl (1-6) linked residue, a conformation similar to that of cellulose. The structure of these gums differs according to the distribution and the number of galactose residue along the mannose chain, randomly arranged in pairs and triplets in the case of the GG [5] leading to regions of low or high substitution, or in blockwise while LBG presents a random, blockwise, and ordered distribution of α-D-galactopyranosyl residues along the β-D-mannose backbone [9]. The degree of galactosyl substitution is responsible for water solubility differences of galactomannans; an increase in the substitution leads to higher solubility through steric effects, whereas galactose-poor regions are less soluble and can involve both inter- and intramolecular associations. Then, GG is dissolved in cold water, while heating is needed to solubilize LBG. The
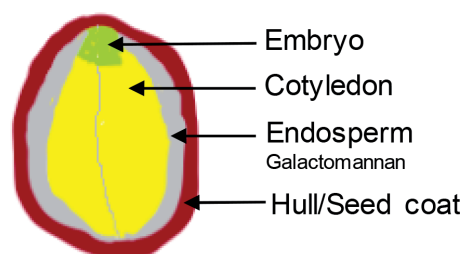

Figure 1.
Scheme of gum seed cross-section.

ratio of mannose to galactose can vary because of the extraction process (particularly purification conditions depending of the end of the desired product) and plant source (geographic location with various climate). The average ratio of galactose to mannose has been estimated to be 1/2 for GG (typically in the range 1/1.4–1.8) [2, 10, 11] and 1/4 (found in the range 1/2.3–6.0) [2, 11, 12] for LBG according to the different chemical techniques (high-performance liquid chromatography, gas chromatography, ^{13}C NMR spectroscopy, or enzymatic method with β-D-mannase).

Because of its good sensitivity and its simplicity in sample preparation, Fourier Transform InfraRed (FTIR) spectroscopy has been common to differentiate GG and LBG and mixtures [10, 12, 13], to study galactomannose interaction with solids [14], to control the gum quality after chemical treatments modifying their properties [15, 16], and to predict the origin [17] with a partial least squares regression-discriminant analysis (PLS-DA).

The aim of this work is to confirm the potential of FTIR technique for the discrimination and the classification of the nature of LBG and GG with the help of chemometric treatments such as principal component analysis (PCA) and partial least squares regression. Moreover, linear-discriminant analysis (LDA) was used to predict the percentage of adulteration by using Scheffé's simplex network to generate simulated binary blends taking into account the variability of the chemical composition of GG and LBG because of their different geographic origins and manufacturing processes.

2. Materials and method

2.1 Sampling

Guar (n = 74) and locust bean (n = 25) commercial gums were obtained from different suppliers without information about their geographic origins, manufacturing processes, mannose/galactose ratios, and the mesh size of particles (**Table 1**).

These powders were freeze-dried before spectroscopic characterization to eliminate the available water interactions. Mathematical binary blends were also built with simplex approach in different GG percentages (varying between 0 and 100% in weight) from simplex method to take into account the variability of spectral signature of GG and LBG.

Pure sugars (D-mannose and D-galactose) were purchased from Sigma-Aldrich (99% of purity) to obtain its FTIR-ATR profile in the same conditions of the gum sample.

2.2 Attenuated total reflectance (ATR) characterization

The technique of attenuated total reflectance (ATR) is making easier the solid and liquid analysis by reducing the sample preparation time and increasing spectral

| Gum	Suppliers
LBG	Alliance Gums & Industries, ARLES, Chemcolloids Ltd, Iranex, Pharmacie des Rosiers, Santeflor, SEATH International, Sigma Aldrich, Tassy & Cie, Viscogum FA
GG	Alliance Gums & Industries, ARLES, Chemcolloids Ltd, Associated Dichem corporation, Iranex, Laviosa MPC, Nitrochemie, SEATH International, Sigma Aldrich, Santeflor, Starlight, ROTH, Tassy & Cie, Viscogum FA

Table 1.
Gum suppliers.
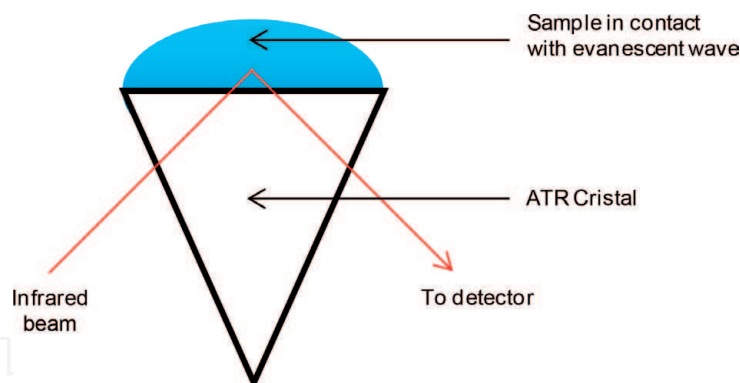
Figure 2.
A simple reflection ATR system.

reproducibility by depositing the sample on the crystal of the attenuated total reflection accessory. By crossing the optical dense crystal (with a high refractive index), the infrared beam will undergo the phenomenon of total internal reflection creating an evanescent wave that extends beyond the surface of the crystal and penetrates a few microns into the sample (the least dense medium). If the sample absorbs light, a part of light energy is retained and the total reflection is attenuated (**Figure 2**).

The penetration depth D_p of the evanescent wave at any specific wavelength is a function of the angle ϕ of incidence of the internally reflected beam and of the ratio of the refractive index n_1 of the crystal to the index n_2 of the sample:

$$D_p = \lambda / \left[2\pi\, n_1 \left(\sin^2 \phi - (n_2/n_1)^2 \right) \right]^{\frac{1}{2}} \quad (1)$$

The refractive index of a diamond crystal is $n_1 = 2.4$ and for the organic compound $n_2 = 1.5$ on the average. For an angle of incidence of 45°, the penetration depth is approximated by $D_p = 0.2\lambda$ (i.e., between 0.5 and 5/µm approximately for the mid-infrared range). This ATR experiment supposes a very good optical contact between the crystal and the sample. To improve this contact, a press is used.

GG and LBG powders were directly deposited on the attenuated total reflectance (ATR) accessory (Specac's "Golden Gate") equipped with a diamond crystal prism (brazed in only one tungsten carbide part), four mirrors, and two ZnSe focusing lenses to reflect the optical path. The sample was pressed on the crystal area with the pressure arm. FTIR-ATR spectra were recorded with a Thermo Nicolet IS10 spectrometer equipped with a Mercury Cadmium Telluride (MCT) detector, an Ever-Glo source, and a KBr/Ge beam-splitter, at room temperature. Data acquisition was done in absorbance mode from 4000 to 650 cm^{-1} with a 4 cm^{-1} nominal resolution (OMNIC 8.1 software). For each spectrum, 100 scans were co-added. A background scan in air (in the same resolution and scanning conditions used for the samples) was carried out before the acquisition. The ATR crystal was carefully cleaned with ethanol to remove any residual trace of the previous sample. Three spectra were recorded for each sample.

2.3 Spectral corrections

The spectral range of the absorption of the carbon dioxide was removed (between 2400 and 1900 cm^{-1}) and then a baseline correction was used to adjust the spectral offset. A unit vector normalization was applied to compensate for additive and/or multiplicative effects (**Figure 3**).
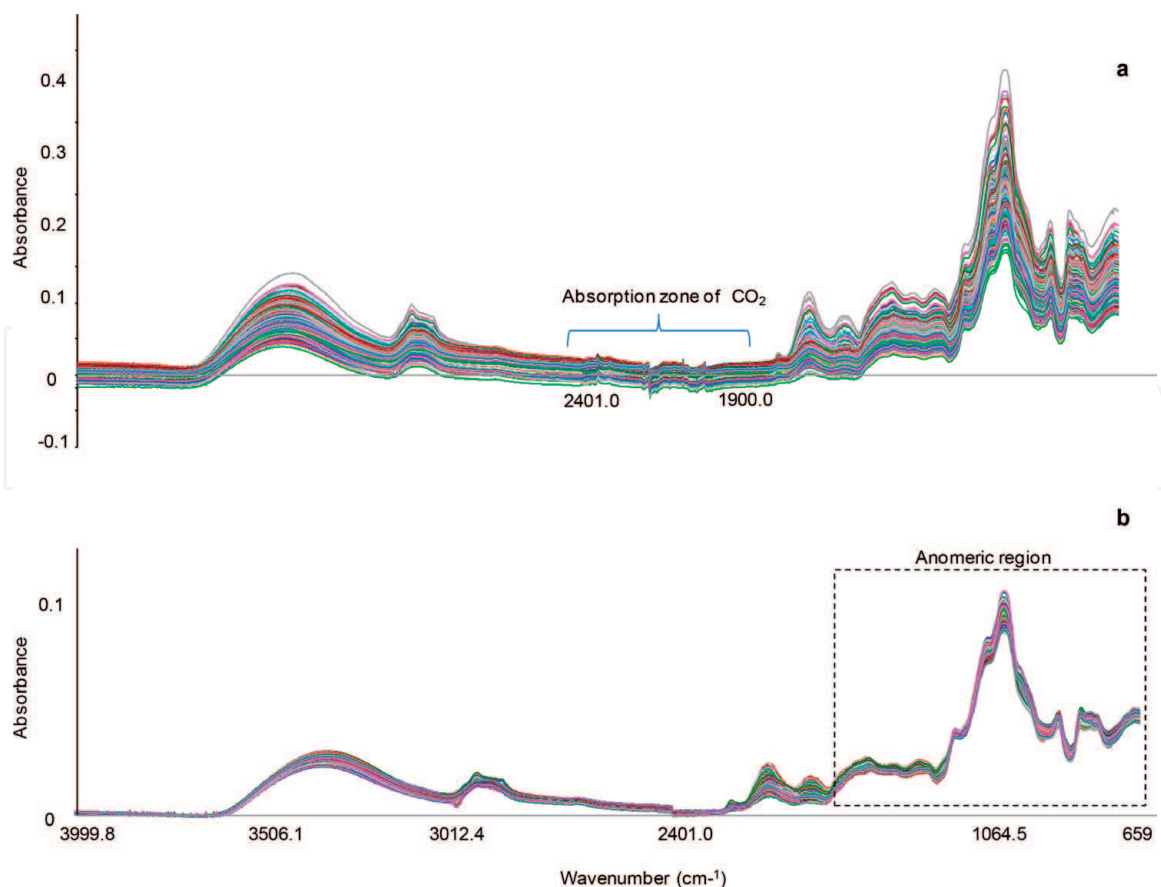
Figure 3.
Guar and locust bean gum FTIR-ATR spectra: original data (a) and corrected data (b).

A selection of variable was used between 1450 and 650 cm^{-1} for the chemometric treatments to keep only the fingerprint of gums where the anomeric region was located.

2.4 Chemometrics

Principal component analysis (PCA) and partial least squares discriminant analysis (PLS-DA) regression are two tools currently used in chemometrics, they are described in a previous study [18, 19]. The optimal number of principal components and latent variables was determined by a cross validation. A binary codification was used to build PLS-DA models which will predict continuous values (positive or negative) and arbitrary intervals of predicted values were chosen as zones delimited the predicted samples well and no-recognized. Also samples with a reference value of one (expected positive) were considered true positive if their predicted value was between 0.7 and 1.3, false negative if predicted under 0.5, and uncertain if predicted between 0.5 and 0.7 or over 1.3. On the contrary, samples with a reference value of zero (expected negative) were considered true negative if predicted between −0.3 and 0.3, false positive over 0.5, and uncertain between 0.3 and 0.5 or under −0.3. The quality of the model was evaluated by the total percentage of correct classification, as well as sensitivity and specificity, calculated according to the following equations:

$$\%\text{correct classification of variety} = \frac{\text{true positive + true negative}}{\text{total number of predicted samples}} \times 100 \quad (2)$$

$$\%\text{sensitivity} = \frac{\text{true positive}}{\text{expected positive}} \times 100 \quad (3)$$

$$\%specificity = \frac{\text{true negative}}{\text{expected negative}} \times 100 \tag{4}$$

Linear discriminant analysis (LDA) is the simplest of all possible classification methods that provide a linear transformation of n-dimensional samples into an m-dimensional space (m < n). LDA allowed to develop a model based on predefined classes (GG proportion-LBG proportion) and corresponding FTIR-ATR data of pure gums and gum blends. Here calibration model was built with gum blends generated mathematically from Scheffé's simplex approach (described in the following section) in order to obtain a large number of combinations of gum proportions with different increments (between 0 and 100%) and determine the best fit parameters for classification of gums. These models were then used to classify pure gums or blends not considered in the calibration step. For models, the percentage of recognition (or correct classification) of gum proportions was obtained by calculating first the absolute error on the proportion of each gum as follows:

$$\text{Absolute error} = \left|\text{proportion}^{\text{predicted}} - \text{proportion}^{\text{reference}}\right|_{\text{guar or locust bean gum}} \tag{5}$$

A correct classification was considered when absolute error on gum proportion was inferior to the increment and a value of one was attributed to the sample. Otherwise, a zero value was affected to samples with badly predicted gum proportion. Then a percentage of correct classification was calculated.

2.5 Simplex mixture system

The dataset of binary mixtures was built with a simplex approach [20] where the studied factors were the proportions of the components. FTIR-ATR spectral data of GG and LBG powders were combined in several proportions given by Scheffé's mixture design [21, 22] to create artificial blends. Considering q components and a constant constraint on the sum of component proportions, Scheffé used a regular (q − 1) dimensional simplex to describe the experimental region of the possible component combinations. In our case, a binary system (q = 2), the required simplex is a straight line, where q apexes of simplex space corresponded to pure component (GG or LBG). On this linear space, each of the *N* calculated mixtures was characterized by a weight (w_j) satisfying the following equations:

$$\sum_{j=1}^{q=2} w_j = 1 \tag{6}$$

$$N = \frac{(w + q - 1)!}{(q - 1)!w!} \tag{7}$$

This method favored a uniform distribution of each mixture between the two pure components and made sure that all these immediate neighbors were at the same distance from the latter. The constant w represented here the mesh step.

An approach similar to that developed by Semmar et al. [20] was applied to take into account the variability of the spectral profile of pure components due to different geographic origins or different manufacturing processes. Spectral profiles of pure component were randomly chosen between the dataset of pure GG or LBG samples to obtain an average profile subsequently used to build mixture components. The number of spectral data of pure components was equal to w, defined previously. k iterations (400) of Scheffé's simplex design were carried out to ensure variability in average spectral profiles. A matrix of (N × k) mixtures was obtained with variables corresponding to the spectral wavenumbers and the associated values

of w_j. These average mixture profiles were used as input data to perform a linear discriminant analysis to build a calibration model to predict further the pair of points (w_{GG}, w_{LBG}) representing the proportions of each gum in the blends.

2.6 Software

The Unscrambler Version 10.3 from CAMO (Computer Aided Modeling, Trondheim, Norway) was used to perform chemometric analyses (PCA, PLS, and LDA). The Scheffé approach was developed using MATLAB R2014b.

3. Results and discussion

3.1 Spectral signature of GG and LBG

Galactomannans are polysaccharides formed by a linear (1-4)-β-D-mannan backbone with a D-galactose side chain (**Figure 4**). On average, GG has a single α-D-galactopyranosyl unit connected by (1–6) linkages to every second main chain unit. In the case of the LBG, unsubstituted or sparingly (1–4) substituted regions of mannopyranose units and regions heavily substituted with α-D-galactopyranosyl residues attached by (1–6)-bonds have been observed.

The galactose substitution that influences the intrinsic flexibility of mannan backbone causes solubility differences and controls the rheological properties. The FTIR-ATR signatures of GG and LBG (**Figure 5**) show essentially difference in intensity due to the mannose/galactose ratio and the presence of residual chemical compounds from thermo-mechanical and chemical dehusking pretreatments (remaining germ particles, products of thermal degradation of endosperm, etc.). In GG, the endosperm is composed of 75% of galactomannose and the rest consists of pentosan, protein, pectin, phytin, ash, and dilute acid insoluble residues [5]. High protein content (like albumin,
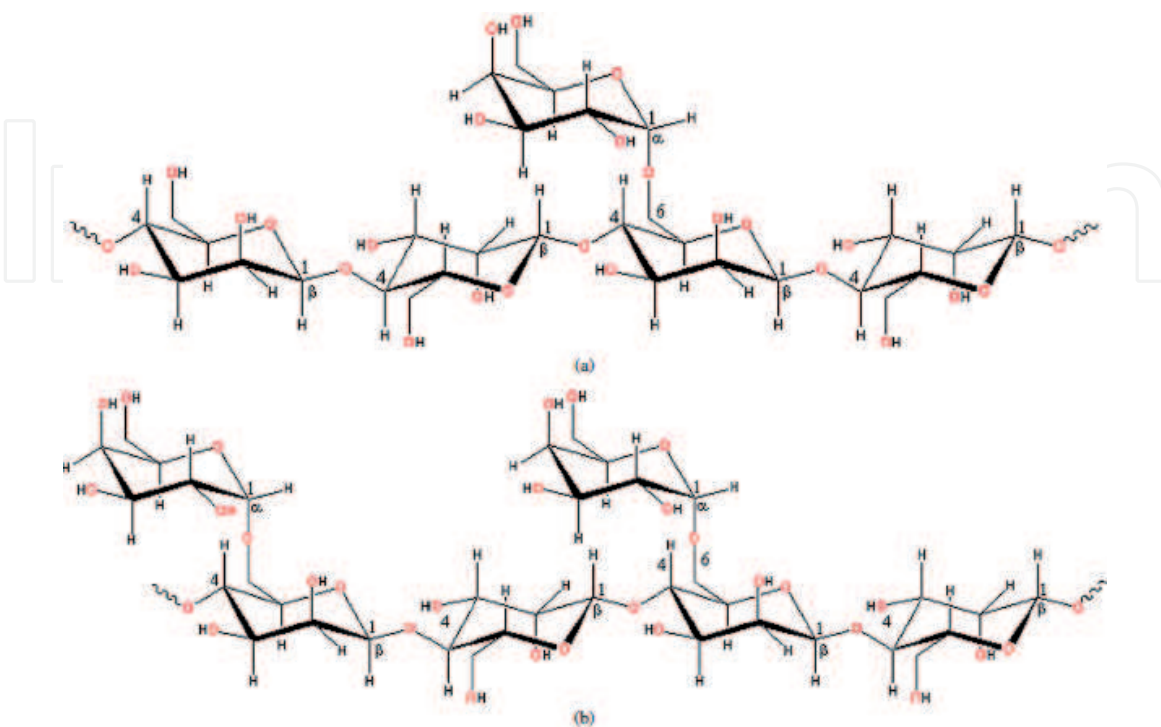

Figure 4.
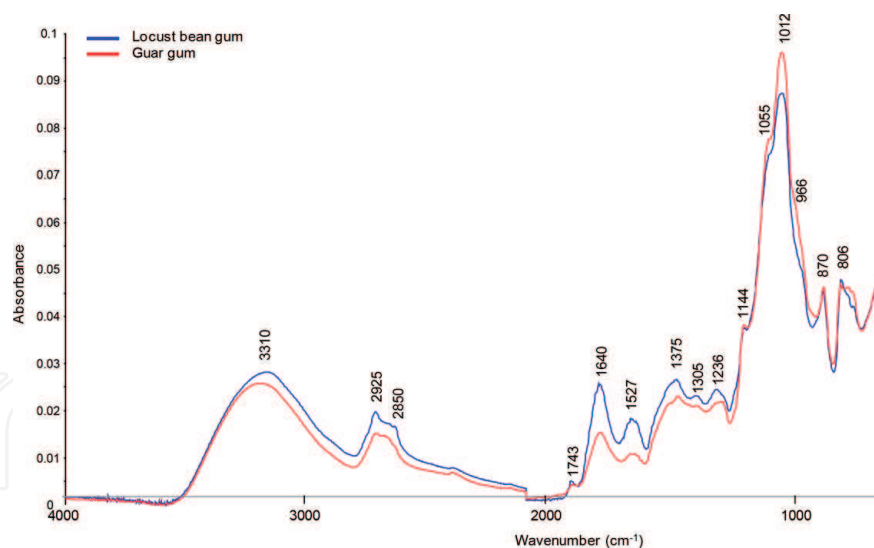*Molecular structure of LBG (a) and GG (b) from [14].*

Figure 5.
FTIR-ATR spectra of LBG and GG.

globulin, and glutelin [2]) could be found in LBG because of a greater contamination by germ particles; lipids were also detected, just like the presence of rhamnose, arabinose, xylose, and glucose, contaminants proceeding from the seed coat [23].

The broadband between 3700 and 3000 cm^{-1} describes the O−H and N−H stretching vibration (hydrogen bonding) attributed to water, amide, and carbohydrates. The spectral zone between 3000 and 2800 cm^{-1} contains two bands assigned to C−H stretching vibration in methylene groups of rings (ν_{as} CH_2 at 2925 cm^{-1} and ν_s CH_2 at 2871 cm^{-1}). The carbonyl signal at 1743 cm^{-1} (C=O stretching vibration) could be in relation with the amide group of amino acids. A spectrum of GG reference material does not reveal the bands of these contaminants [10]. The band pointed at 1640 cm^{-1} is due to the presence of bound water (O−H bending of absorbed water), which cannot be eliminated despite the freeze-drying. This last band could also be attributed to the axial deformation of C=O bond (amide band I) and the one pointed at 1527 cm^{-1} could correspond to the angular deformation of N−H bond (amide band II) or to amine N−H deformation vibrations due to the presence of impurities such as proteins and amino acids found in the germ and seed coat, badly remaining during the purification process [24].

The peaks observed in the spectra between 1480 and 1190 cm^{-1} represented C−H, C−OH (1236 cm^{-1}), and H−C−H deformation vibration (bending). Depending on the level of protein impurities, a band at 1236 cm^{-1} could be assigned to the amide band III (C−N) vibration mode [24]. The following spectral region (1190–900 cm^{-1}) presents different bands contributing to the skeletal vibrations and glycosidic bonds (νCC, νCOC, νCCO, and δOCH) of galactomannans' sugar composition. The C−O stretching mode of pyranose ring exists as a small shoulder at 1055 cm^{-1} and a broad peak at 1012 cm^{-1}, while the shoulder at 966 cm^{-1} is a characteristic contribution of C−OH bending [25]. At lower wavenumbers (900–700 cm^{-1}), peaks at 806 and 870 cm^{-1} appear that are related with anomeric C−H deformation bands (CCH and OCH) of structural isomers (α or β-pyranose compounds), equatorial C−H deformation bands (nonglycosidic), and skeletal symmetric and asymmetric ring vibrations (CCO, COC, and OCO) [26, 27].

The FTIR-ATR spectrum of each sugar is presented in **Figure 6**.

These compounds have a CH_2OH labile group capable of generating intra- and intermolecular hydrogen bonding. This CH_2OH labile group and the arrangement of the other OH group on pyranose ring affect differently the spectral region between 3500 and 3000 cm^{-1}. The reversed hydroxyl group orientation in the C-2 and C-4 atom in the
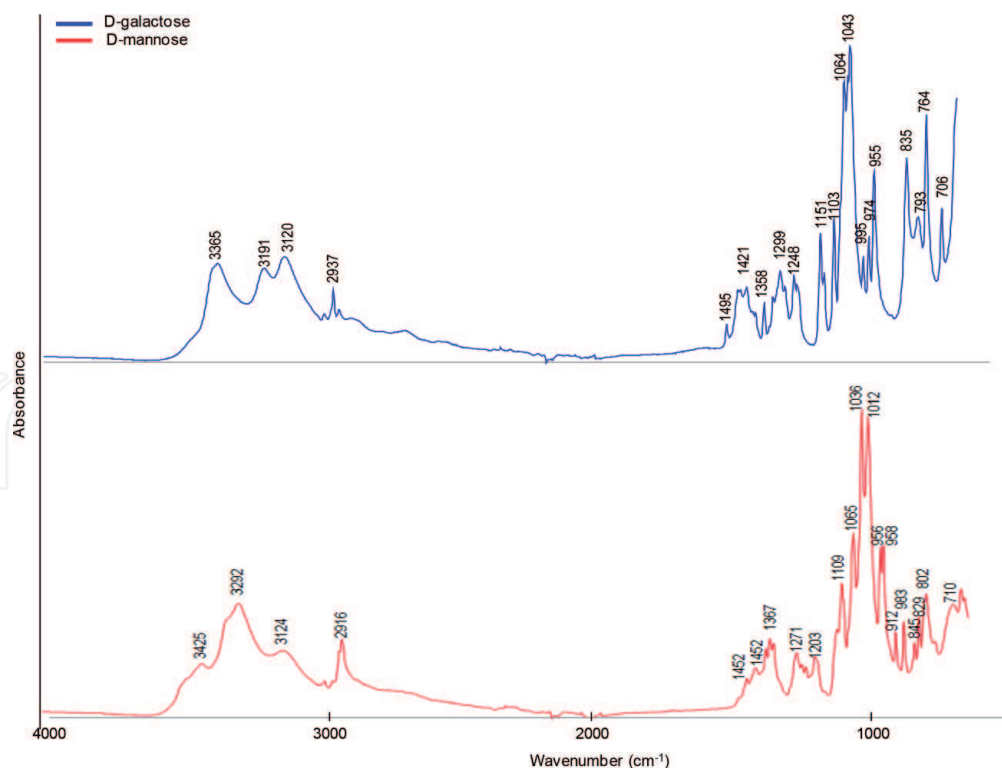
Figure 6.
FTIR-ATR spectra of D-mannose and D-galactose.

structure results in differences in positions and intensities of several bands in the region of CH$_2$ scissoring (1495–1420 cm^{-1}), between 1400 and 1100 where bands of CH$_2$ wagging and twisting, C–O and C–C stretching vibration of backbone, and COH bending of primary and secondary bonded alcohol appear. Because of the equatorial and axial OH position, the glycosidic and anomeric regions (1000–700 cm^{-1}) show also spectral differences about skeletal stretching (CO, CC, and COC) vibration coupled with (COH, CCH, and CCO) deformation bands (**Table 2**) [27, 28].

3.2 PCA of gum FTIR-ATR spectra

PCA, carried out on all of the normalized infrared spectra of gums, shows that on the components 1 and 2, which represent respectively 57 and 24% of the total spectral variance, the samples of GG in red and LBG in blue form two groups perfectly separated from each (**Figure 7**). The best separation between clusters was obtained with a selection of variables in anomeric region, between 1450 and 700 cm^{-1}.

The mid-infrared technique allows characterizing the commercial samples of GG and LBG. Globally, all GG samples are negatively projected; LBG samples are positively projected along PC1, while PC2 is a representative of the intragroup variability. The spectral band characteristics of GG and LBG powders are observable on PC1 loading (**Figure 8**).

The negative part of PC1 loading characterize more LBG samples by exalting spectral bands pointed at 1386, 1315, and 1236 cm^{-1} (amide bands) characterizing the proteins whose content appears more important in the case of LBG. The band at 1170 could be attributed to C–O vibrations of locust bean galactomannan. The large dispersion of locust bean samples along PC1 could be due to the presence of impurities in gum powders coming from husk and germ, (like insoluble matter, proteins, amino acid, etc.), residual compounds resulting from the various steps of extraction, and purification processes. Different maturity stages of seed, geographic origin, and climatic conditions could also be responsible for the variability of the chemical composition of these galactomannan gums. The positive part of PC1 reveals

Spectral region	Wavelength (cm⁻¹)		Chemical group assignment
	D-Galactose	**D-Mannose**	
3500–3000	3365, 3191, 3120	3425	O—H stretching vibration, hydrogen bonded
3000–2800	2937	2916	C—H stretching vibration of CH_2 group
1500–1200	1495, 1421, 1358, 1299, 1248	1452, 1419, 1367, 1271, 1203	CH_2 symmetric deformation, C—OH deformation
1200–950	1151, 1103, 1064, 1043, 995, 974	1109, 1065, 1036, 1012, 956, 958, 912	C—C skeletal, C—O—C, CO—O—C, C—O stretching vibration, C—O bending of C—OH group, OCH bending
950–700	835, 793, 764, 706	845, 829, 802	Anomeric C—H deformation (α or β), equatorial C—H deformation (nonglycosidic) and asymmetric and symmetric ring vibration

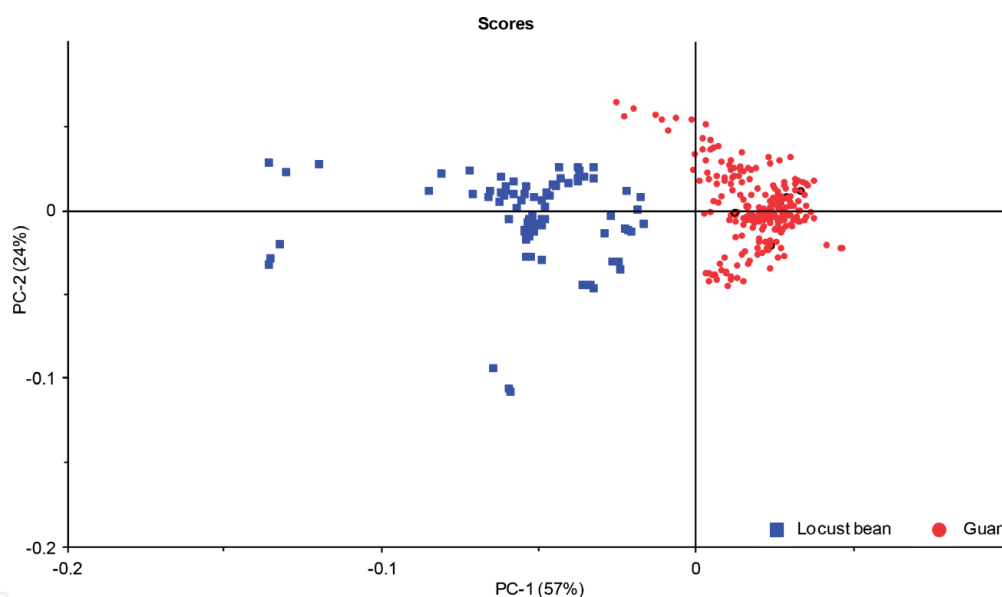**Table 2.**
Spectral interpretations of pure sugars.

Figure 7.
Representation of GG and LBG samples in (PC1 and PC2) plan (normalized spectral data).

the most intense spectral bands in GG samples at 1066, 1012, 964, 863, 819, and 771 cm⁻¹. Bands at 1012 and 964 cm⁻¹ are attributed to C–O–H and C–O vibrations, respectively. The band at 771 cm⁻¹ is due to ring stretching and ring deformation of β-D-(1-4) and α-D-(1-6) linkages. These last ones are specific to the anomeric region where C–O stretching bands are more representative because of the largest number of galactose units (1-6) linked to β-D-mannopyranosyl backbone in GG.

While GG was richer in galactosyl residue, no specific bands of D-galactose were found in positive part of PC1 loading representing the GG samples. The comparison with the spectral bands of pure D-galactose or D-mannose is not a good way because of their crystalline structure (free form) that is not the case in their polymer form. Another explanation could be a possible interference with the presence of water that modifies the band's resolution in the anomeric region as it is observable in **Figure 9** presenting sugar profiles under crystalline and hydrated forms.
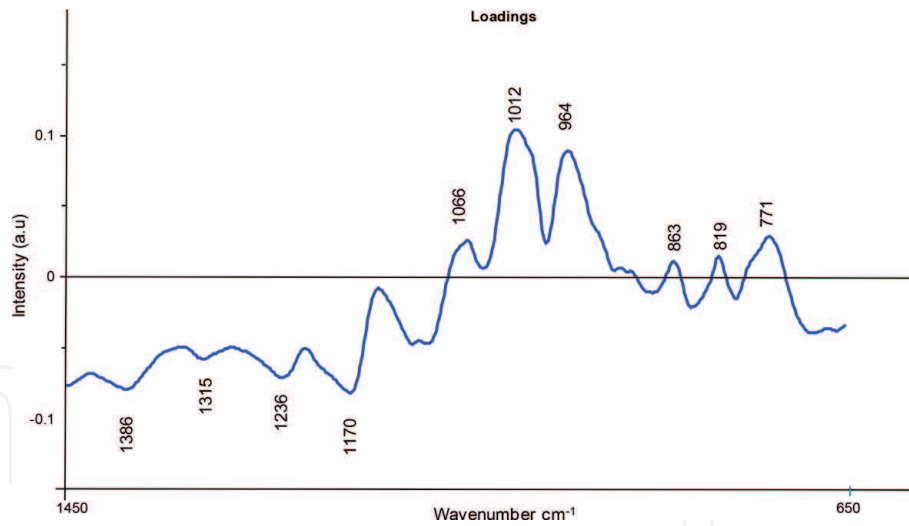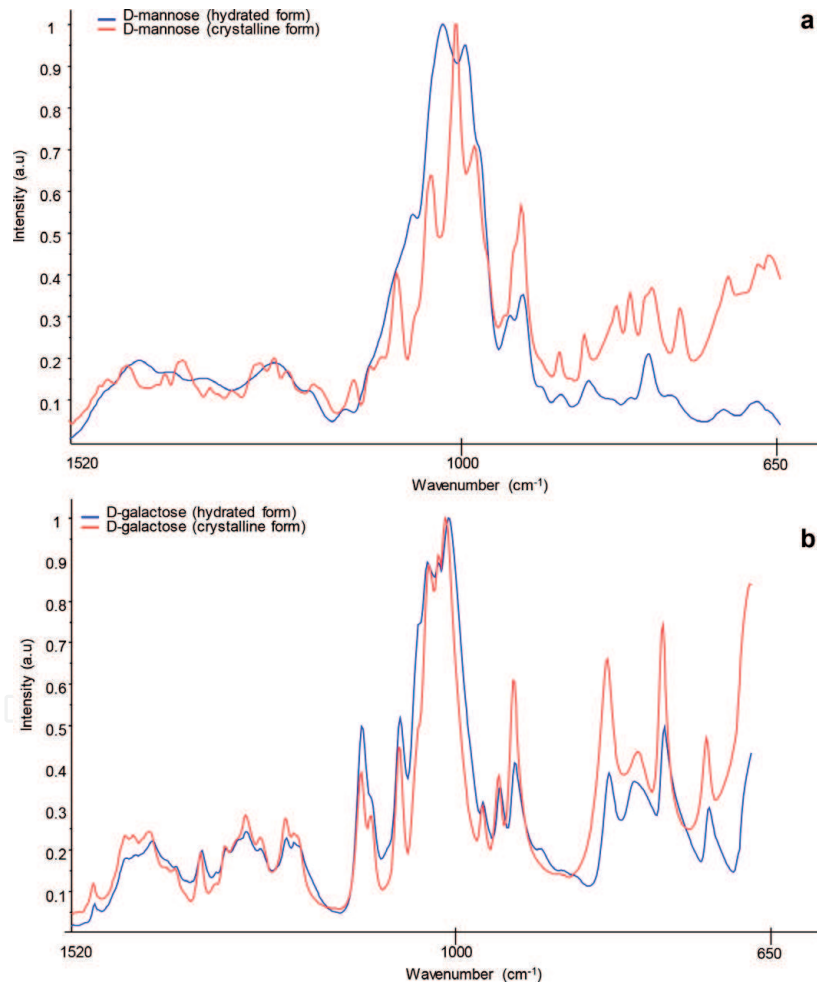
Figure 8.
PC1 loading.

Figure 9.
Spectral profiles of D-mannose (a) and D-galactose (b) under crystalline and hydrated forms.

3.3 Prediction of species origin by PLS-1-DA regression

The classification into varietal origin (GG or LBG) was performed using PLS-1-DA analysis. The calibration dataset was composed of 37 GG and 13 LBG (n = 50 samples × 3 spectra = 150). Different samples of calibration set, 12 LBG and 37 GG (n = 49 samples × 3 spectra = 147), have constituted the validation set.

Model parameters	Values
Spectra number for calibration	150
Spectra number for validation	147
LV	3
SEC	0.08
SEP	0.11
R^2	0.97
Q^2	0.94
%Sensibility	100
%Specificity	100
%Correct classification	100

SEC or SEP: standard error of calibration or prediction, R^2 or Q^2 correlation coefficients for calibration or prediction.

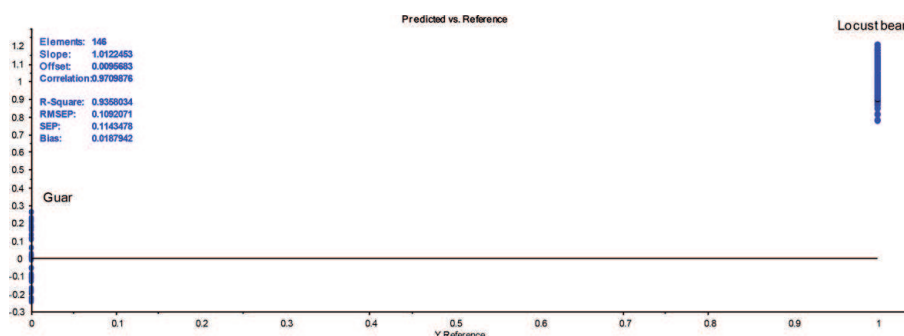**Table 3.**
PLS-1-DA parameters for LBG.

Figure 10.
Predicted versus reference value of gum classification (validation step).

The calibration model had very good quality parameters, as shown in **Table 3**, and very good validation results for LBG were obtained using normalized spectra with a selection of wavelength region from 1450 to 700 cm^{-1}.

For GG validation, the different statistic parameters were also closed to 100%.

The following graph (**Figure 10**) shows that predicted values from validation data are closed to zero for GG and 1 for LBG.

Calibration model has satisfying quality parameters, as shown in **Figure 10**, with RMSEP ranging from 0.11 for LBG to 0.94 for Q^2 (R-square), and 100% good classification is obtained. The predicted species origins are never given by zero or one results because the different rates of carbohydrates in the samples vary according to the origins. As a matter of fact, there is a natural variation of the carbohydrate rates that can notably be a function of geographic origins and harvest dates.

Contrary to the results of Prado et al. [12] who published that diffuse reflectance (DRIFT) method was better suited for differentiation of gum type, ATR technique showed here a very good classification. It was noted that these authors have realized spectra with a ZnSe multiple bounce ATR on gum aqueous solutions, heated for their preparation. Nevertheless, the work of Wang et al. [14] showed that the computer-simulated molecular space filling structure of GG and LBG was different in no solvent and aqueous environments. In the aqueous environment, GG form presented a more complicated structure than LBG form because of the increase in galactose units on the mannose backbone [14]. In conclusion, spectral data in solid or liquid environment were difficult to compare as done by Prado et al. [12] because the intermolecular interactions in the structure of gums were not the same.

3.4 Quantitative analysis of gum mixtures

As PLS-DA allowed easily the prediction of the botanical origin of galacto-mannans, LDA was adapted to the prediction of the proportions (or weights, w_i) of pure compounds in blends with the advantage to be governed by a constant sum of w_j (equal to 100%). But the fact to provide ordinal predictions (or class) leads to a strict response of model, which considers a wrong classification even if the predictive weights are slightly different from reference data. A selection of variables between 1900 and 650 cm^{-1} and an average reduction (by two) of variables in this spectral zone were performed to make simplex iterative operations possible, and because in LDA classification method, the number of objects (or pairs of weights) should be larger than the number of variables (wavenumbers). In this way, 325 variables, 11 pairs of weights between 0 and 1 (with a constant increment of 0.1), and a value of k (number of iterations in Scheffé's simplex) equal to 400 to be superior at the variable number (the constraint of LDA calibration step) were used to generate 4400 artificial bends (11 × 400) from simplex design.

An example of the repartition of simulated blends generated with this method is given in **Figure 11** where a step of 10% was chosen to clarify the graphic representation. The graph has been obtained after realizing a PCA on blends' binary data.

Dispersed experimental samples were placed at the extrema of PC1 axis: LBG at the left in the negative part of PC1 and GG at the right in the positive part, respectively. It well appears that simulated blends well take into account the intrinsic variability of pure components, but in certain regions, an inevitable overlapping originating from the high variability of chemical composition of pure gums is also observed.

Five different LDA calibration models were built with five values of weight step (0.100, 0.050, 0.04, 0.02, and 0.007). The robustness of each calibration model has been tested with four validation sets obtained with steps of 0.10, 0.083, 0.067, and 0.002, without constraint about the number of blends. The results are resumed in **Table 4**.

All weights of GG and LBG in pure state and mixtures containing 2–10% of GG were well predicted in the validation step. A percentage of 61% was obtained when the increment chosen in validation was lower than the calibration step.
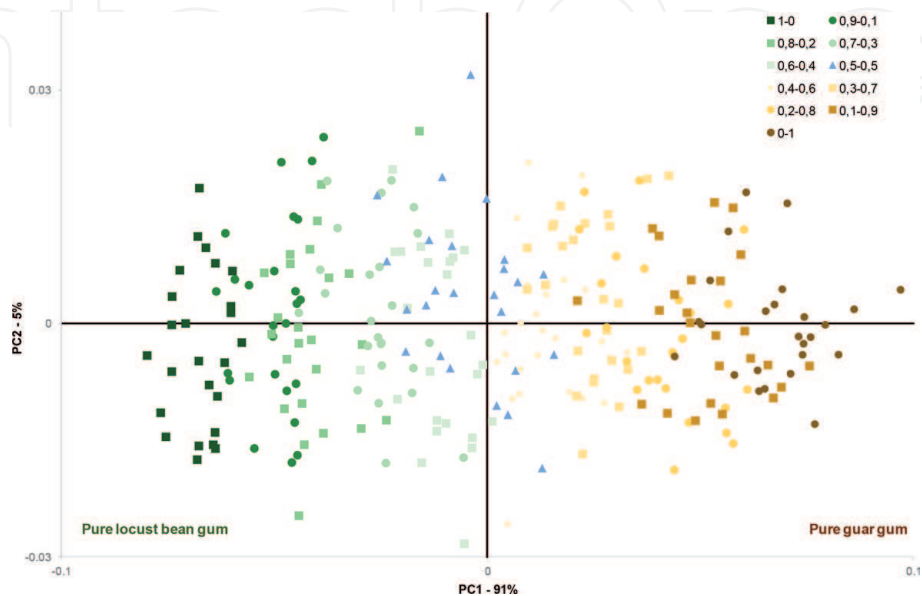

Figure 11.
Location of the binary blends in the simplex space defined by GG and LBG proportions (step of 10%).

Blends number		Calibration model				
		4400	8400	10400	20400	60400
	Weight step	0.100	0.050	0.040	0.020	0.0067
1100	0.100	100%	100%	100%	100%	100%
780	0.083	100%	100%	100%	100%	100%
1200	0.067	100%	100%	100%	100%	100%
5100	0.020	61%	100%	100%	100%	100%

Table 4.
Prediction of correct weight for different validation sets.

4. Conclusions and outlook

Guar and locust bean gums are galactomannans with similar chemical structure because only the galactose/mannose ratio differentiates them. Depending on the geographic origins of seeds and their industrial manufacturing process, this ratio can be a variable and improves certain variability in their chemical composition, weakly discernible according to the analytical techniques used to characterize these gums. Despite the similarity of the infrared spectral signatures of guar and locust bean gums and the overlapping of IR bands due to inter- and intramolecular inter-actions, this work showed the feasibility of using FTIR-ATR technique to discriminate GG and LBG with the help of chemometric treatments. Best results have been obtained with a variable selection in the anomeric spectral region (between 1450 and 650 cm^{-1}) to differentiate gum samples using a principal component analysis and to predict the species origin with a partial least square regression. A particular approach has been proposed to quantify the proportion of gums in blends with a good accuracy with the advantage to detect an adulteration of LBG by GG from their spectral profiles. The Scheffé's simplex approach allowed taking into account the variability of the chemical composition of gum samples due to different environmental parameters and manufacturing processes and generating a lot of simulated blends to increase the robustness of LDA calibration model. The approach with computational blends is a rapid and a low cost way to generate an FTIR-ATR dataset used in favor of quality control and prediction of adulteration.

Author details

Catherine Rébufa, Fabien Girard, Jacques Artaud and Nathalie Dupuy*
Aix Marseille Univ, Avignon Université, CNRS, IRD, IMBE, Marseille, France

*Address all correspondence to: nathalie.dupuy@univ-amu.fr

IntechOpen

References

[1] Aqualon. Guar and Guar Derivatives Oil and Gas Field Applications. 2007. Available from: http://www.ashland.com/Ashland/Static/Documents/AAFI/PRO_250-61_Guar.pdf

[2] Barak S, Mudgil D. Locust bean gum: Processing, properties and food applications—A review. International Journal of Biological Macromolecules. 2014;**66**:74-80

[3] Thombare N, Jha U, Mishra S, Siddiqui MZ. Guar gum as a promising starting material for diverse applications: A review. International Journal of Biological Macromolecules. 2016;**88**:361-372

[4] Hasan AMA, Abdel-Raouf ME. Applications of guar gum and its derivatives in petroleum industry: A review. Egyptian Journal of Petroleum. 2018;**27**:1043-1050

[5] Gupta S, Variyar PS. Chapter 12: Guar gum: A versatile polymer for the food industry. In: Grumezescu AM, Holban AM, editors. Handbook of Food Bioengineering. Vol. XX: Biopolymers for Food Design. Cambridge: Elsevier, Academic Press; 2008. pp. 383-407

[6] Gillet S, Simon M, Paquot M, Richel A. Synthèse bibliographique de l'influence du procédé d'extraction et de purification sur les caractéristiques et propriétés d'une gomme de caroube. Biotechnology, Agronomy, Society and Environment. 2014;**18**:97-107

[7] Srečec S, Dunkić V, Bezić N, Kremer D, Erhatić R. Some doubts controversies about anatomy of locust bean (*Ceratonia siliqua* L.) seed coat. Genetics, plant breeding and seed production. In: Proceeding of 53rd Croatian & 13th International Symposium on Agriculture; 18-23 February 2018; Vodice, Croatia. 2018

[8] Vishwakarma RK, Shivhare US, Nanda SK. Novel method and machine for dehulling of guar seeds and optimisation of dehulling process. Food and Bioproducts Processing. 2016;**99**:51-57

[9] Pitkänen L, Tuomainen P, Mikkonen K, Tenkanen M. The effect of galactose side units and manna chain length on the macromolecular characteristics of galactomannans. Carbohydrate Polymers. 2011;**86**:130-1235

[10] Liyanage S, Abidi N, Aud D, Moussa H. Chemical and physical characterization of galactomannan of extracted from guar cultivars (*Cyamopsis tetragonolobus* L.). Industrial Crops and Products. 2015;**74**:388-396

[11] Prajapati VD, Girish KJ, Moradiya NG, Randeria NP, Nagar BJ, Naikwadi NN, et al. Galactomannan: A versatile biodegradable seed polysaccharide. International Journal of Biological Macromolecules. 2013;**60**:83-92

[12] Prado BM, Kim S, Ozen BF, Mauer LJ. Differenciation of carbohydrate gums and mixtures using Fourier transform infrared spectroscopy and chemometrics. Journal of Agricultural and Food Chemistry. 2005;**53**:2823-2829

[13] Ćopíková J, Synystsya A, Ćerná M, Kaasová J, Novotná M. Application of FT-IR spectroscopy in detection of food hydrocolloids in confectionery jellies and food supplements. Czech Journal of Food Sciences. 2001;**19**:31-36

[14] Wang J, Somasundaram P. Study of galactomannose interaction with solid using AFM, IR and allied techniques. Journal of Colloid and Interface Science. 2007;**309**:373-338

[15] Mudgil D, Barak S, Khatkae BS. X-ray diffraction, IR spectroscopy and

thermal characterization of partially hydrolysed guar gum. International Journal of Biological Macromolecules. 2012;**50**:1035-1039

[16] Yen SN, Choi SM, Phillips DL, Ma CY. Raman and FTIR spectroscopic study of carboxymethylated non-starch polysaccharides. Food Chemistry. 2009;**114**:1091-1098

[17] Alabdi F, Elharfi N, Balouki A, Kzaiber F, Oussama A. Locust bean origin classification by FTIR spectroscopy and chemometrics. Journal of Chemistry and Chemical Engineering. 2001;**5**:1020-1029

[18] Lafhal S, Vanloot P, Bombarda I, Valls R, Kister J, Dupuy N. Raman spectroscopy for identification and quantification analysis of essential oil varieties: A multivariate approach applied to lavender and lavandin essential oils. Journal of Raman Spectroscopy. 2015;**46**:577-585

[19] Bombarda I, Dupuy N, Le Van Da JP, Gaydou EM. Comparative chemometric analyses of geographic origins and compositions of lavandin var. Grosso essential oils by mid-infrared spectroscopy and gas chromatography. Analytica Chimica Acta. 2008;**613**:31-39

[20] Semmar N, Artaud A. A new simplex-based approach predicting olive oil blend compositions from fatty acid data. Journal of Food Composition and Analysis. 2015;**43**:149-159

[21] Scheffé H. Experiments with mixtures. Journal of the Royal Statistical Society: Series B. 1958;**20**:344-360

[22] Scheffé H. Simplex centroid designs for experiments with mixtures. Journal of the Royal Statistical Society: Series B. 1963;**25**:235-263

[23] Dakia A, Blecker C, Robert C, Wathelet B, Paquot M. Composition and physicochemical properties of locust bean gum extracted from whole seeds by acid or water dehulling pre-treatment. Food Hydrocolloids. 2008;**22**:807-818

[24] Figueiró SD, Góes J, Moreira RA, Sombra ASB. On the physico-chemical and dielectric properties of glutaraldehyde crosslinked galactomannan-collagen films. Carbohydrate Polymers. 2004;**56**:313-320

[25] Cerquiera MA, Souza BWS, Simões J, Teixeira JA, Domingues MRM, Coimbra MA, et al. Structural and thermal characterization of galactomannans from non-conventional sources. Carbohydrate Polymers. 2011;**83**:179-185

[26] Talari ACS, Martinez MAG, Movasaghi Z, Rehman S, Rehman IU. Advances in Fourier transform infrared (FTIR) spectroscopy of biological tissues. Applied Spectroscopy. 2017;**52**:456-506

[27] Socrates G. Infrared and Raman Characteristic Group Frequencies Tables and Charts. 3rd ed. New York: John Wiley & Sons, Ltd; 2001. p. 347

[28] Wiercigroch E, Szafraniec E, Czamara K, Pacia MZ, Majzner K, Kochan K, et al. Raman and infrared spectroscopy of carbohydrates: A review. Spectrochimica Acta, Part A: Molecular and Biomolecular Spectroscopy. 2017;**185**:317-335