

Characterization and discrimination of phenolic compounds using Fourier transform Raman spectroscopy and chemometric tools

Darly R. Pompeu^(1,2), Yvan Larondelle⁽²⁾, Hervé Rogez⁽³⁾, Ouissam Abbas⁽⁴⁾, Juan Antonio Fernández Pierna⁽⁴⁾, Vincent Baeten⁽⁴⁾

⁽¹⁾ Universidade do Estado do Pará. Centre of Natural Science and Technology. Department of Food Technology. Trav. Éneas Pinheiro 2626. BR-66.095-100 Belém-PA (Brazil).

⁽²⁾ Université catholique de Louvain. Institut des Sciences de la Vie. Croix du Sud, 2/8. BE-1348 Louvain-la-Neuve (Belgium).

⁽³⁾ Universidade Federal do Pará & Centre for Agro-food Valorisation of Amazonian Bioactive Compounds (CVACBA). Faculty of Food Engineering. Av. da Ciência, km 01. BR-66.095-780 Belém-PA (Brazil).

⁽⁴⁾ Walloon Agricultural Research Centre (CRA-W). Valorisation of Agricultural Products Department. Chaussée de Namur, 24. BE-5030 Gembloux (Belgium). E-mail: o.abbas@cra.wallonie.be, v.baeten@cra.wallonie.be

Received 9 May 2017, accepted 17 October 2017, available online 24 October 2017.

This article is distributed under the terms and conditions of the CC-BY License (<http://creativecommons.org/licenses/by/4.0>)

Description of the subject. Phenolic compounds (PCs) are the most abundant secondary metabolites in plants. This work was part of a study that sought to develop rapid screening FT-Raman methods for identifying and quantifying classes and/or types of PCs in the dry extracts of plant products.

Objectives. Fourier transform Raman spectroscopy (FT-Raman) coupled with chemometric tools was used to characterize and discriminate four families of phenolic compounds: two important classes of phenolic acids, hydroxybenzoic and hydroxycinnamic acids, as well as their derivatives and flavonoids.

Method. For this study, 25 standards of phenolic compounds were used (47 standards in total, taking account of the different brands). Repeatability and reproducibility studies were conducted to verify the Raman assignments of gallic acid. Raman characterization with the most significant spectral bands of phenolic compounds was done using spectra ranging from 1,800 to 50 cm⁻¹. A Fisher test was applied to pre-processed Raman spectra (SNV) and 20 Raman scattering signals were used to differentiate each class of phenolic compounds.

Results. Hydroxybenzoic acids, hydroxycinnamic acids, their derivatives and flavonoids have been characterized by Raman spectroscopy; bands have been identified and differentiated within and between groups.

Conclusions. The scattering intensities located around 1,600-1,699, 1,300-1,400 and below 200 cm⁻¹ were responsible for differentiating 100% of phenolic compound families, classes and subclasses.

Keywords. Spectroscopy, phenolic compounds, hydroxybenzoic acids, hydroxycinnamic acids, flavonoids.

Caractérisation et discrimination des composés phénoliques à l'aide de la spectroscopie Raman à transformée de Fourier et des outils chimiométriques

Description du sujet. Les composés phénoliques (CP) sont les métabolites secondaires les plus abondants dans les plantes. Ce travail fait partie d'une étude visant à développer des méthodes rapides de criblage FT-Raman pour identifier et quantifier les classes et/ou types de CPs dans les extraits secs ou les végétaux.

Objectifs. La spectroscopie Raman à transformée de Fourier (FT-Raman) couplée à des outils chimiométriques a été utilisée pour caractériser et distinguer quatre familles de composés phénoliques : deux importantes classes d'acides phénoliques, à savoir des acides hydroxybenzoïques et hydroxycinnamiques, ainsi que leurs dérivés et des flavonoïdes.

Méthode. Pour cette étude, 25 standards de composés phénoliques ont été utilisés (47 standards au total, en tenant compte des différentes marques). Des études de répétabilité et de reproductibilité ont été effectuées pour vérifier les affectations Raman de l'acide gallique. La caractérisation Raman a été réalisée sur base des bandes spectrales les plus significatives dans la région allant de 1 800 à 50 cm⁻¹. Un test Fisher a été appliqué à des spectres Raman pré-traités (SNV) et 20 des signaux Raman ont été utilisés pour distinguer chaque classe de composés phénoliques.

Résultats. Les acides hydroxybenzoïques, les acides hydroxycinnamiques, leurs dérivés et des flavonoïdes ont été caractérisés par spectroscopie Raman ; les bandes ont été identifiées et différenciées au sein mais aussi entre les groupes.

Conclusions. Les intensités Raman situées autour de 1 600-1 699, de 1 300-1 400 et inférieures à 200 cm^{-1} sont responsables de la différenciation de 100 % des familles de composés phénoliques, des classes et sous-classes.

Mots-clés. Spectroscopie, composés phénoliques, acides hydroxybenzoïques, acides hydroxycinnamiques, flavonoïdes.

1. INTRODUCTION

Phenolic compounds (PCs) are the most abundant secondary metabolites in plants. They comprise a wide variety of molecules that have a phenolic structure consisting of a hydroxyl group (-OH) bonded directly to an aromatic hydrocarbon group (Robards et al., 1999; Ignat et al., 2011). Between May 2005 and May 2015, PCs have been cited more than 29,985 times in the literature (scopus.com, title, abstract or keywords, accessed 21 May 2015), illustrating their importance in the scientific world. Despite this importance, there has been limited work on correctly identifying and quantifying them. The current techniques used to determine PCs need trained personnel, are time consuming, and cannot be used for real-time measurements because their application to raw material control remains very limited (Baeten et al., 2015). Rapid, non-destructive and adaptable on-line techniques are needed for the fast characterization of bioactive compounds, especially PCs.

Raman spectroscopy is a branch of vibrational spectroscopy based on shifts in the wavenumber or frequency of an incident exciting monochromatic radiation. The shift results from the inelastic scattering of interaction between the photons and the sample. Raman spectroscopy is usually measured in the 3,600-200 cm^{-1} range. This region corresponds to Raman Stokes scattering bands. This spectroscopic technique is used in chemistry to identify (Schrader et al., 1999; Baranska et al., 2004; Baranska et al., 2006) and characterize substances (Schulz et al., 2005; Paiva-Martins et al., 2011; Zuk et al., 2011) and compounds (Fiuza et al., 2004; Teslova et al., 2007; Corredor et al., 2009; Świsłocka et al., 2012; Machado et al., 2013; Mishra et al., 2013) and to study molecular and crystalline symmetries and identify crystalline polymorphism of compounds (Numata & Tanaka, 2011). The most commonly used Raman spectroscopies are based on two technologies, dispersive Raman and Fourier transform (FT) Raman. Each technology has its advantages and is suited to specific types of analysis. FT-Raman avoids most of the fluorescence perturbation and provides spectra with high frequency precision.

Raman spectroscopy exhibits well-resolved bands of fundamental vibrational transitions and provides a useful amount of information on the molecular structure of compounds. In the case of PCs, spectral features such as the presence or absence of scattering bands, as

well as band scattering positions, have been reported in the literature. Billes et al. (2007) investigated the assignment of the Raman spectra of gallic acid in its crystalline form and the spectral changes due to the presence of water in the structure. Calheiros et al. (2008) studied the influence of the ester alkyl chain (methyl, ethyl, propyl, isopropyl, butyl, octyl and dodecyl groups) on Raman spectral features of caffeic, ferulic and gallic acids. Świsłocka et al. (2013) reported spectral features of three hydroxybenzoic acids (4-hydroxybenzoic, vanillic and syringic acids) and two benzoic derivative (benzoic and 3-methoxybenzoic acids) standards. Eravuchira et al. (2012) investigated the Raman spectra of derivatives of cinnamic acids (3-caffeoylquinic, 4-caffeoylquinic, 5-caffeoylquinic, 3,4-di-*o*-caffeoylquinic, 3,5-di-*o*-caffeoylquinic, 4,5-di-*o*-caffeoylquinic and 3-feruloylquinic acids). In all these studies, vibrational bands were assigned and pointed to characterize these PCs. To date, however, no systematic approach has been developed to differentiate the PCs.

In this study, FT-Raman spectroscopy was used to characterize 25 standards of PCs: six hydroxybenzoic acids (gentisic, protocatechuic, gallic, 4-hydroxybenzoic, vanillic, syringic acids) and four hydroxycinnamic acids (2-hydroxycinnamic, caffeic, ferulic, and sinapic acids), as well as four of their derivatives (catechol, chlorogenic acid, resveratrol and tannic acid) and 11 flavonoids (bavachinin, catechin, daidzein, epicatechin, epicatechin gallate, epigallocatechin, epigallocatechin gallate, genistein, luteolin, quercetin dihydrate and rutin). Various chemometric tools applied to the characteristic Raman spectra were used to exhibit key bands, allowing differentiation between families, classes and subclasses of PCs. This work was part of a study that sought to develop rapid screening FT-Raman methods for identifying and quantifying classes and/or types of PCs in the dry extracts of plant products.

2. MATERIALS AND METHODS

2.1. Chemicals

The chemical standards of the hydroxybenzoic acids, hydroxycinnamic acids, their derivatives and the flavonoids were purchased from Sigma-Aldrich (Steinheim, Germany), Extrasynthèse (Genay, France)

and VWR (Darmstadt, Germany) (**Table 1**). In total, 47 standards (HPLC grade with purity > 95%) were used in the study. The samples were stored at -20 °C and room-conditioned 1 h before the start of the analysis.

2.2. Optimization of measurement conditions, repeatability and reproducibility

Gallic acid was used to optimize the measurement conditions of PCs using FT-Raman:

- two weights (3 and 5 mg) of gallic acid were manually placed and compacted in ten small aluminium ring cups;
- three laser power intensities were used (100, 200 and 400 mW);
- FT-Raman scattering data were collected with a spectral resolution of 1 cm⁻¹ by co-adding 32, 64, 128, 256 and 512 scans.

Once the measurement parameters had been optimized, FT-Raman measurements were taken over 4 days in order to verify the repeatability and reproducibility of this technique for PC determination.

2.3. Raman spectroscopy

FT-Raman spectra were acquired on a Vertex 70–RAM II FT-Raman spectrometer obtained from Bruker (Bruker Optics, Ettlingen, Germany), equipped with an Nd:YAG laser (Yttrium Aluminium Garnet crystal doped with triply ionized Neodymium) with an output at 1,064 nm (or 9,398.5 cm⁻¹) and a liquid-nitrogen cooled germanium detector.

The samples were manually placed and compacted in small aluminium ring cups with a hole that had an inner diameter of 2 mm. Spectra were recorded from 50 to 3,599 cm⁻¹. Each PC was independently and randomly measured in duplicate.

OPUS 6.0 Software (Ettlingen, Germany) was used for the spectral data acquisition.

2.4. Chemometric analysis

Spectral data were smoothed using the Savitzky-Golay algorithm (using a 3 points window and a second order polynomial). Matlab 7.14 (The Mathworks, Natick, MA) was used to develop and apply an algorithm to identify wavenumbers where the Raman scattering intensity was at least 5% of the maximum Raman

Table 1. Phenolic compounds used in this study — *Composés phénoliques utilisés dans cette étude.*

Class	Phenolic compound	Brands
Hydroxybenzoic acids	Gentisic acid	Extrasynthèse, Sigma-Aldrich, VWR
	Protocatechuic acid	Extrasynthèse, Sigma-Aldrich, VWR
	Gallic acid	Extrasynthèse, Sigma-Aldrich, VWR
	4-hydroxybenzoic acid	Extrasynthèse, Sigma-Aldrich
	Vanillic acid	Extrasynthèse, Sigma-Aldrich, VWR
	Syringic acid	Extrasynthèse, Sigma-Aldrich, VWR
Hydroxycinnamic acids	2-hydroxycinnamic acid	Extrasynthèse, Sigma-Aldrich
	Caffeic acid	Extrasynthèse, Sigma-Aldrich, VWR
	Ferulic acid	Extrasynthèse, Sigma-Aldrich
	Sinapic acid	Extrasynthèse, Sigma-Aldrich
Derivatives of hydroxybenzoic and hydrocinnamic acids	Chlorogenic acid	Extrasynthèse, Sigma-Aldrich, VWR
	Ellagic acid	Extrasynthèse, Sigma-Aldrich
	Resveratrol	Extrasynthèse
	Tannic acid	Extrasynthèse, Sigma-Aldrich, VWR*
Flavonoids	Bavachinin	Extrasynthèse
	Catechin	Extrasynthèse
	Daidzein	Extrasynthèse
	Epicatechin	Extrasynthèse
	Epicatechin gallate	Extrasynthèse
	Epigallocatechin	Extrasynthèse
	Epigallocatechin gallate	Extrasynthèse
	Genistein	Extrasynthèse
	Luteolin	Extrasynthèse
	Quercetin dehydrated	Extrasynthèse
	Rutin	Extrasynthèse

*: two different tannic acids purchased from VWR were used in this study — *deux acides tanniques différents achetés chez VWR ont été utilisés dans cette étude.*

scattering intensity. To confirm these Raman bands, second derivative pre-processing was performed using the Savitzky-Golay transformation (second order polynomial; 3 points at right and left).

Separations between families, classes and subclasses of PC were made using Raman data with standard normal variate (SNV) pre-processing. Unscrambler® X 10.3 Software, from CAMO (Computer Aided Modelling, Trondheim, Norway), was used to do classifications.

3. RESULTS AND DISCUSSION

3.1. Optimization of measurement conditions, repeatability and reproducibility

The most commonly used method to quantify total PC content is the colorimetric method using the Folin-Ciocalteu reagent. With this method, calibration curves are usually built using gallic acid for its high stability (Volf et al., 2014), although other chemical standards can be employed, *e.g.* caffeic, chlorogenic and tannic acids. Gallic acid was therefore chosen for the first step of the study.

The measurement conditions were tested:

- amounts of gallic acid in the aluminium ring cup (3 or 5 mg);
- laser power intensity (100, 200 and 400 mW);
- number of scans collected (32, 64, 128, 256 or 512).

As expected, the Raman scattering peaks were the same, irrespective of the measurement conditions (data not shown). With regard to sample quantity, 5 mg were selected as being easier to compact inside the small ring than 3 mg. A laser power intensity of 100 mW gave low Raman scattering intensities, whereas 400 mW could have caused fluorescence damage (Baeten et al., 2001); the intensity was therefore set at 200 mW. The number of scans chosen was 128; below this number (32 and 64 scans) the Raman spectra quality was not good enough to give a clear determination, and working with more than 128 would have required a long measurement time.

Once the FT-Raman conditions had been optimized, precision tests were done. The precision with which the FT-Raman technique is able to characterize PCs was evaluated in terms of repeatability and reproducibility. Repeatability is measurement results under conditions where independent measurement results are obtained with the same method on the identical test items in the same laboratory by the same operator using the same equipment within short intervals of time (ISO 5725, 1994). Reproducibility can be defined as the closeness of agreement between independent results obtained with the same method on identical material but under

different conditions. These precision parameters were evaluated in terms of Raman scattering data (cm^{-1}).

In order to calculate these factors, 10 spectra (each one the mean of 128 scans) of gallic acid were collected over 4 days. Slight differences in Raman intensity and Raman scattering signal shifts were observed in six spectral ranges: 1,260-1,250, 1,100-1,080, 960-950, 710-685, 285-275 and 140-120 cm^{-1} . A second derivative pre-processing on spectral data, however, demonstrated that there were no spectral differences in the Raman scattering data. **Figure 1** presents the original FT-Raman spectra (a) and second derivative FT-Raman spectra (b) obtained from gallic acid.

All interpretations of spectra in our study were based in Socrates (1997).

3.2. Raman characterization of hydroxybenzoic acids

Figure 2 shows the FT-Raman spectra of six hydroxybenzoic acids from different companies in the region of 50-1,800 cm^{-1} . The most important Raman scattering signals observed are summarized in **table 2**, indicating that these PCs present important spectral information in the region studied.

The FT-Raman spectra of hydroxybenzoic acids showed two series of intense spectral bands: the most intense was below 200 cm^{-1} and the second most intense was between 1,715 and 1,590 cm^{-1} . The first one was due to skeletal vibration; this region is also useful for describing lattice vibrations, the main manifestation of the intermolecular forces in crystals. The second one was due to aryl carboxylic acid C=O stretching vibrations (1,715-1,680 cm^{-1}) and C=C stretching vibrations from the aromatic ring (1,625-1,590 cm^{-1}). Some spectral signals were also observed around 1,410-1,310 cm^{-1} , associated mainly with O-H deformation and C-O stretching combination vibrations of phenols. Aromatic =C-H in-plane and out-of-plane deformation vibrations were visible in the 1,290-1,000 cm^{-1} and 965-680 cm^{-1} regions, respectively. The region between 650 and 415 cm^{-1} is more characteristic of aromatic ring vibrations.

3.3. Raman characterization of hydroxycinnamic acids

Figure 3 shows the FT-Raman spectra of four hydroxycinnamic acids from different brands in the region of 50-1,800 cm^{-1} . The most important Raman scattering signals observed are summarized in **table 3**.

The hydroxycinnamic acids studied presented a spectral region below 200 cm^{-1} less intense than the hydroxybenzoic acids. The most intense spectral bands of hydroxycinnamic acids were in the 1,150-1,360 and 1,650-1,590 cm^{-1} regions. Compared with

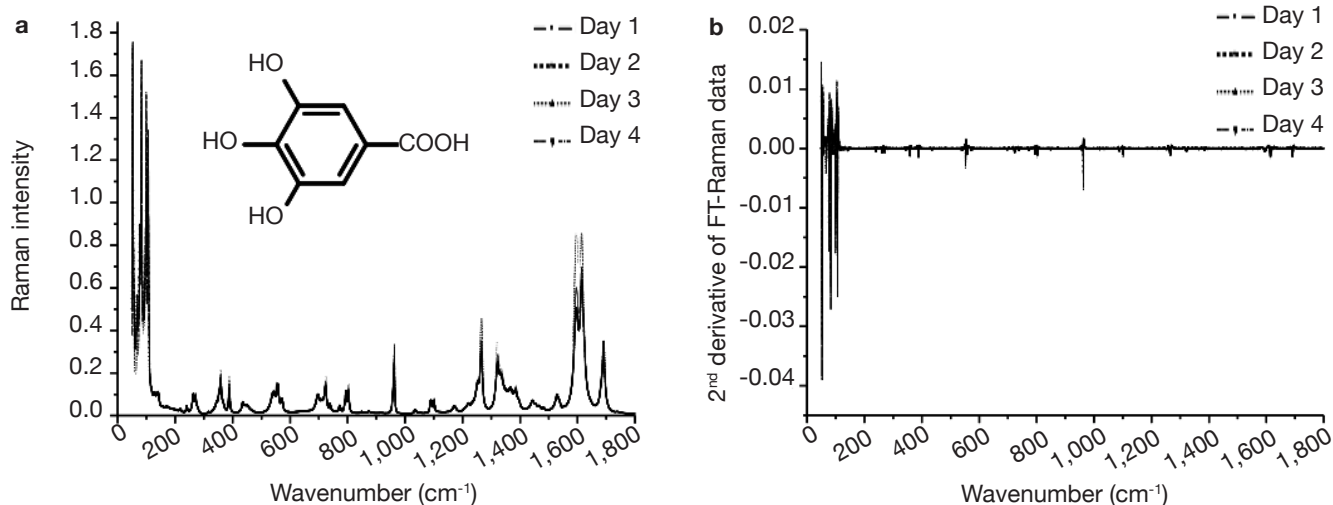


Figure 1. Chemical structures and FT-Raman spectral data of gallic acid. Normal Raman (a) and second derivative FT-Raman data (b) of gallic acid. These figures express average of 10 samples of 3,4,5-trihydroxybenzoic acid measured independently during four days — *Structure chimique et données spectrales FT-Raman d'acide gallique. Données Raman brutes (a) et de dérivée seconde (b) d'acide gallique. Ces chiffres expriment la moyenne de 10 spectres d'acide 3,4,5-trihydroxybenzoïque mesurés indépendamment pendant quatre jours.*

hydroxybenzoic acids, hydroxycinnamic acids have an alkene group between the carboxylic function and the aromatic ring, which results in α,β -unsaturated carboxylic acid theoretically absorbing between 1,715 and 1,680 cm^{-1} . Surprisingly, this band was not visible in our study. The alkene group $\text{C}=\text{C}$ presents bands between 1,640 and 1,610 cm^{-1} due to its conjugation with aryl, but it is also conjugated with $\text{C}=\text{O}$, leading to vibration bands between 1,660 and 1,580 cm^{-1} .

3.4. Raman characterization of derivatives of hydroxybenzoic and hydroxycinnamic acids

Figure 4 presents the FT-Raman spectra of ellagic acid, chlorogenic acid, resveratrol and tannic acid from different brands in the 50-1,800 cm^{-1} region. **Table 4** shows the pointed bands.

Chlorogenic acid and resveratrol presented the two highest peaks in the 1,640-1,600 cm^{-1} spectral region. They can be differentiated by the presence of a small band around 1,690 cm^{-1} corresponding to $\text{C}=\text{O}$ stretching vibration of aryl and α,β -unsaturated ester. For both PCs, the spectral region between 1,000 and 1,400 cm^{-1} was rich in Raman scattering signals. Resveratrol composed of two aromatic rings presented several better resolved and more intense bands in the 1,000-1,400 cm^{-1} region than chlorogenic acid. Tannic acid, which has the most complex structure of all the PCs studied, presented spectral bands that were the least resolved. Ellagic acid deserves some attention. It showed a strong shift as a function of the brand (source company). This shift was confirmed when a second derivative pre-treatment was applied. The

most remarkable differences were at 1,554-1,532, 1,374-1,350, 1,305-1,290, 1,210-1,170, 1,065-1,050, 790-630, 560-320 and below 200 cm^{-1} . It should be remembered that ellagic acid has a center of symmetry; it has a planar compact structure where molecules are interconnected. This might explain the resulting spectrum which was rich in well-resolved bands over the entire spectrum, in addition to bands below 200 cm^{-1} which were numerous and very intense.

3.5. Raman characterization of flavonoids

Flavonoids are molecules with a phenolic benzopyran structure and occur only in plants. They represent a family of PCs. They share a common nucleus consisting of two phenolic rings and an oxygenated heterocycle, and can be divided into classes according to the type of heterocycle involved. In this study, 11 flavonoids from 5 classes (flavanol, flavanone, flavone, flavonol, isoflavone) were investigated.

Figures 5a and **5b** present the FT-Raman spectra of flavonoids in the 50-1,800 cm^{-1} region. **Table 5** shows the pointed signals.

The most intense spectral region was observed below 200 cm^{-1} for almost all the flavonoids. The exceptions were luteolin and rutin. All the flavonoids showed a very important spectral region between 1,570 and 1,700 cm^{-1} . Phenolic compounds without carbonyl function, however, e.g. the flavanols as catechin, epicatechin and epigallocatechin, had only one band around 1,600 cm^{-1} (1,633, 1,617, 1,627 cm^{-1} , respectively) corresponding to the stretching vibrations of aromatic $\text{C}=\text{C}$ groups. The rest of the PCs presented

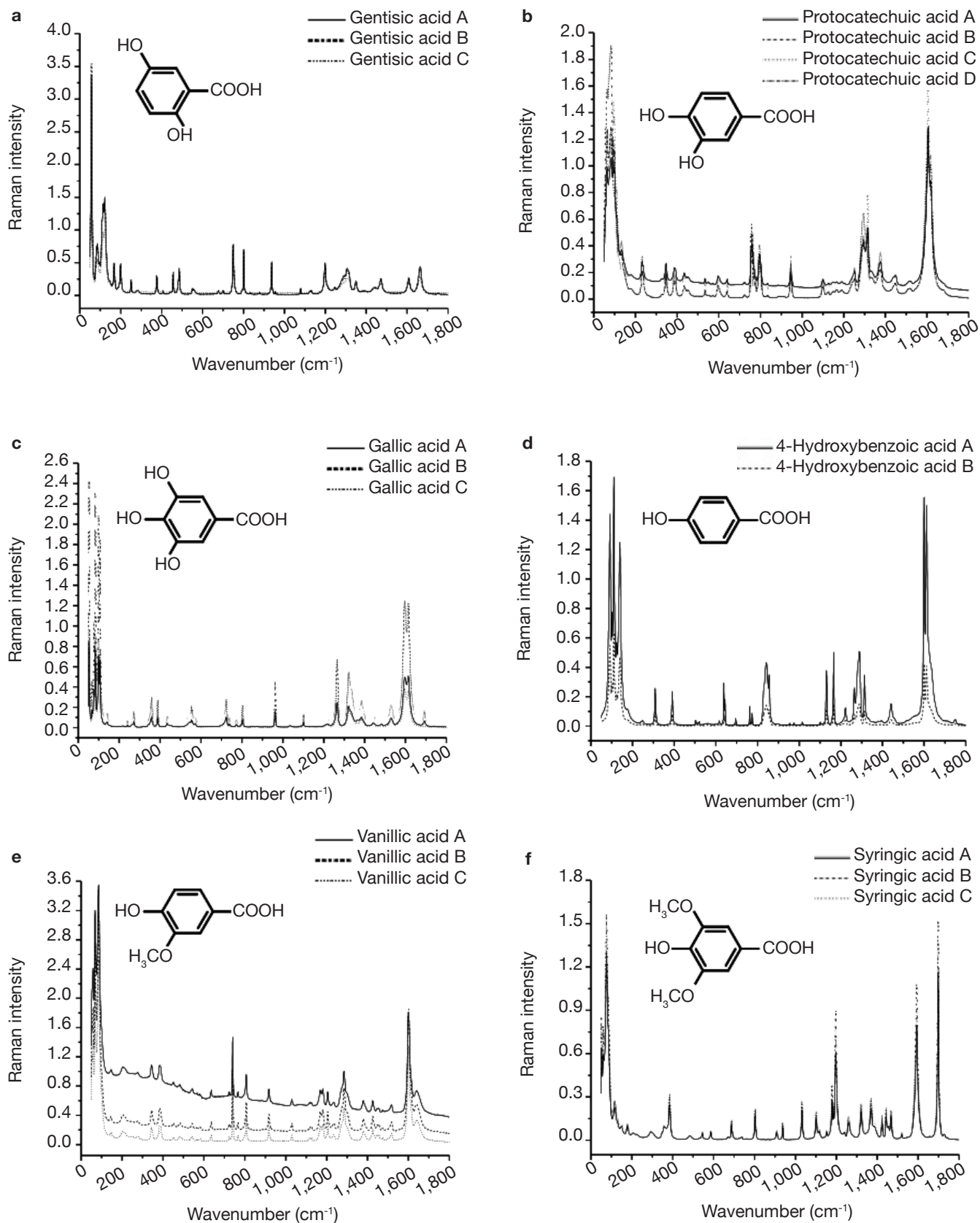


Figure 2. Raman spectra and chemical structures of gentisic (a), protocatechuic (b), gallic (c), 4-hydroxybenzoic (d), vanillic (e) and syringic (f) acids — *Spectres Raman et structures chimiques des acides gentisique (a), protocatéchuïque (b), gallique (c), 4-hydroxybenzoïque (d), vanillique (e) et syringique (f).*

Table 2. The most intense Raman scattering signals (cm^{-1}) of gentisic, protocatechuic, gallic, 4-hydroxybenzoic (4HB), vanillic and syringic acids — *Les signaux Raman les plus intenses (cm^{-1}) des acides gentisique, protocatéchuique, gallique, 4-hydroxybenzoïque (4HB), vanillique et syringique.*

Hydroxybenzoic acids					
Gentisic	Protocatechuic	Gallic	4HB	Vanillic	Syringic
1,663–1,664 _{vw}	1,617–1,618 _{vs}	1,691 _{vw}	1,612 _{vs}	1,602 _{vs}	1,697–1,698 _{vs}
1,315 _{vw}	1,605 _{vs}	1,614 _s	1,600 _{vs}	1,598 _{vs}	1,593 _s
1,306–1,307 _{vw}	1,315 _{vs}	1,596 _s	1,441 _{vw}	1,519–1,520 _{vw}	1,467–1,468 _{vw}
1,290 _{vw}	1,295–1,296 _m	1,530 _{vw}	1,313 _w	1,425 _w	1,443 _{vw}
1,198–1,199 _{vw}	1,251 _w	1,384–1,386 _{vw}	1,289 _w	1,378–1,379 _w	1,370 _w
938 _{vw}	1,095 _s	1,365 _{vw}	1,265 _{vw}	1,284–1,285 _m	1,321 _w
801 _w	946 _w	1,333 _{vw}	1,166 _w	1,205 _w	1,260 _{vw}
750 _w	794–796 _m	1,322 _w	1,131 _w	1,182–1,183 _w	1,197 _s
486 _{vw}	793–794 _m	1,264–1,266 _m	855 _w	1,167–1,168 _w	1,185 _w
456 _{vw}	775–777 _w	963 _w	842 _w	916–918 _w	1,179 _w
375 _{vw}	770 _w	803 _{vw}	763 _{vw}	807–808 _w	1,102 _{vw}
199 _{vw}	758–759 _m	724 _{vw}	644 _{vw}	741 _s	1,032 _w
168 _{vw}	755–754 _m	553 _{vw}	638 _{vw}	724 _{vw}	937 _{vw}
147 _{vw}	534–535 _w	388 _{vw}	391 _{vw}	637 _{vw}	803 _{vw}
122 _s	347–348 _w	359 _{vw}	308 _{vw}	383 _w	688 _{vw}
114 _s	231–233 _m	106 _{vs}	139 _{vs}	343–344 _w	148–149 _{vw}
96 _w	99–100 _{vs}	99 _{vs}	124 _m	86 _{vs}	116–117 _w
86 _w	93–94 _{vs}	83 _{vs}	111 _{vs}	68 _{vs}	76 _{vs}
80 _w	83–85 _{vs}	76 _s	101 _s	58–59 _{vs}	60–61 _s
57 _{vs}	63–64 _{vs}	68 _w	92 _{vs}	52 _{vs}	52 _s
		52 _{vs}			

Each compound from each brand was measured independently twice. These Raman scattering signals represent the average of these measures. The Raman scattering signals were coded with vw (very weak: 5–10%), w (weak: 10–20%), m (medium: 20–30%), s (strong: 30–40%) and vs (very strong: > 40%) in function of their relative Raman intensities concerning the highest Raman intensity — *Chaque composé de chaque marque a été analysé deux fois indépendamment. Ces signaux représentent la moyenne de ces mesures. Les signaux Raman ont été codés avec vw (très faible : 5-10 %), w (faible : 10-20 %), m (moyen : 20-30 %), s (fort : 30-40 %) et vs (très fort : > 40 %) en fonction de leurs intensités Raman relatives à l'intensité la plus élevée.*

several bands in this region and had one band at a higher wavenumber that might have been linked to carbonyl groups of hydroxylated-4H-1-benzopyran-4-one of daidzein, genistein (isoflavones), quercetin hydrate, rutin (flavonols) or luteolin (flavone). In the case of epicatechin gallate and epigallocatechin gallate, the ketone function was related to the ester group, which could explain the highest recorded wavenumbers at 1,683 and 1,692 cm^{-1} , respectively.

For the flavanone, only the PC bavachinin was available. In addition to all the bands associated with

aromatic rings and hydroxyls, the most remarkable spectral bands for this PC were those corresponding to the CH deformation vibration of the methyl $=\text{C}(\text{CH}_3)_2$ group. Bands that were pointed around 1,347 and 1,333 cm^{-1} were of almost equal intensity. There was also a band corresponding to C-C skeletal vibrations of the $=\text{C}(\text{CH}_3)_2$ group.

With regard to the rest of bands observed on the spectra of different flavonoids, they could be associated with aromatic rings and hydroxyl functions, as in the case of the phenolic acids.

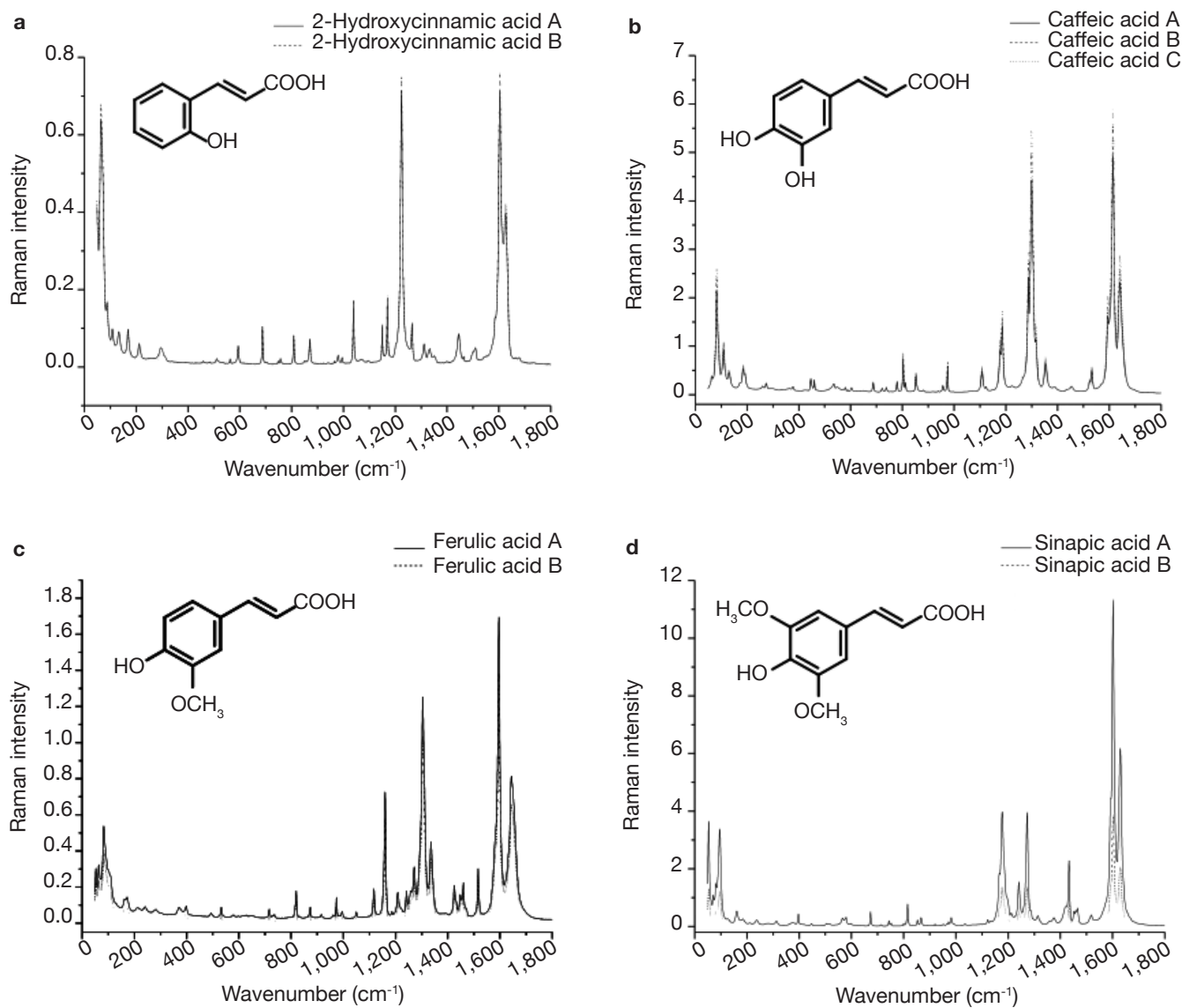


Figure 3. Raman spectra and chemical structures of 2-hydroxycinnamic (a), caffeic (b), ferulic (c) and sinapic (d) acids — *Spectres Raman et structures chimiques des acides 2-hydroxycinnamique (a), caféique (b), férulique (c) et sinapique (d).*

Table 3. The most intense Raman scattering signals (cm^{-1}) of 2-hydroxycinnamic (2HC), caffeic, ferulic and sinapic acids — *Les signaux Raman les plus intenses (cm^{-1}) des acides 2-hydroxycinnamique (2HC), caféique, férulique et sinapique.*

Hydroxycinnamic acids			
2HC	Cafeic	Ferulic	Sinapic
1,626 _{vs}	1,640-1,641 _{vs}	1,631 _{vs}	1,642 _{vs}
1,604 _{vs}	1,625 _m	1,602 _{vs}	1,595-1,596 _{vs}
1,585 _w	1,614-1,615 _{vs}	1,592 _s	1,516 _w
1,445 _w	1,602 _s	1,466 _{vw}	1,472 _{vw}
1,311 _{vw}	1,594 _s	1,459 _{vw}	1,460 _w
1,265 _w	1,533 _{vw}	1,433 _m	1,424-1,425 _w
1,224 _{vs}	1,354 _w	1,273 _s	1,336 _m
1,170 _m	1,317 _m	1,241 _w	1,310 _{vs}
1,150 _w	1,307 _{vs}	1,211 _{vw}	1,304 _{vs}
1,039 _m	1,299 _{vs}	1,201-1,202 _{vw}	1,271 _w
871 _{vw}	1,287 _{vs}	1,196 _{vw}	1,242 _{vw}
808 _w	1,186 _m	1,177 _s	1,209-1,210 _{vw}
687-688 _w	1,177 _w	1,164 _w	1,160 _{vs}
594 _{vw}	975 _w	852 _{vw}	1,117 _w
168-169 _w	803 _w	815 _{vw}	974 _{vw}
108 _w	184 _w	96 _m	820 _{vw}
88 _m	130 _{vw}	81 _w	533 _{vw}
71 _{vs}	109 _w	70-71 _{vw}	82 _s
64 _{vs}	90 _m	63 _{vw}	64 _w
50 _{vs}	82 _{vs}	53-54 _m	52 _w

Each compound from each brand was measured independently twice. These Raman scattering signals represent the average of these measures. The Raman scattering signals were coded with vw (very weak: 5–10%), w (weak: 10–20%), m (medium: 20–30%), s (strong: 30–40%) and vs (very strong: > 40%) in function of their relative Raman intensities concerning the highest Raman intensity — *Chaque composé de chaque marque a été analysé deux fois indépendamment. Ces signaux représentent la moyenne de ces mesures. Les signaux Raman ont été codés avec vw (très faible : 5-10 %), w (faible : 10-20 %), m (moyen : 20-30 %), s (fort : 30-40 %) et vs (très fort : > 40 %) en fonction de leurs intensités Raman relatives à l'intensité Raman la plus élevée.*

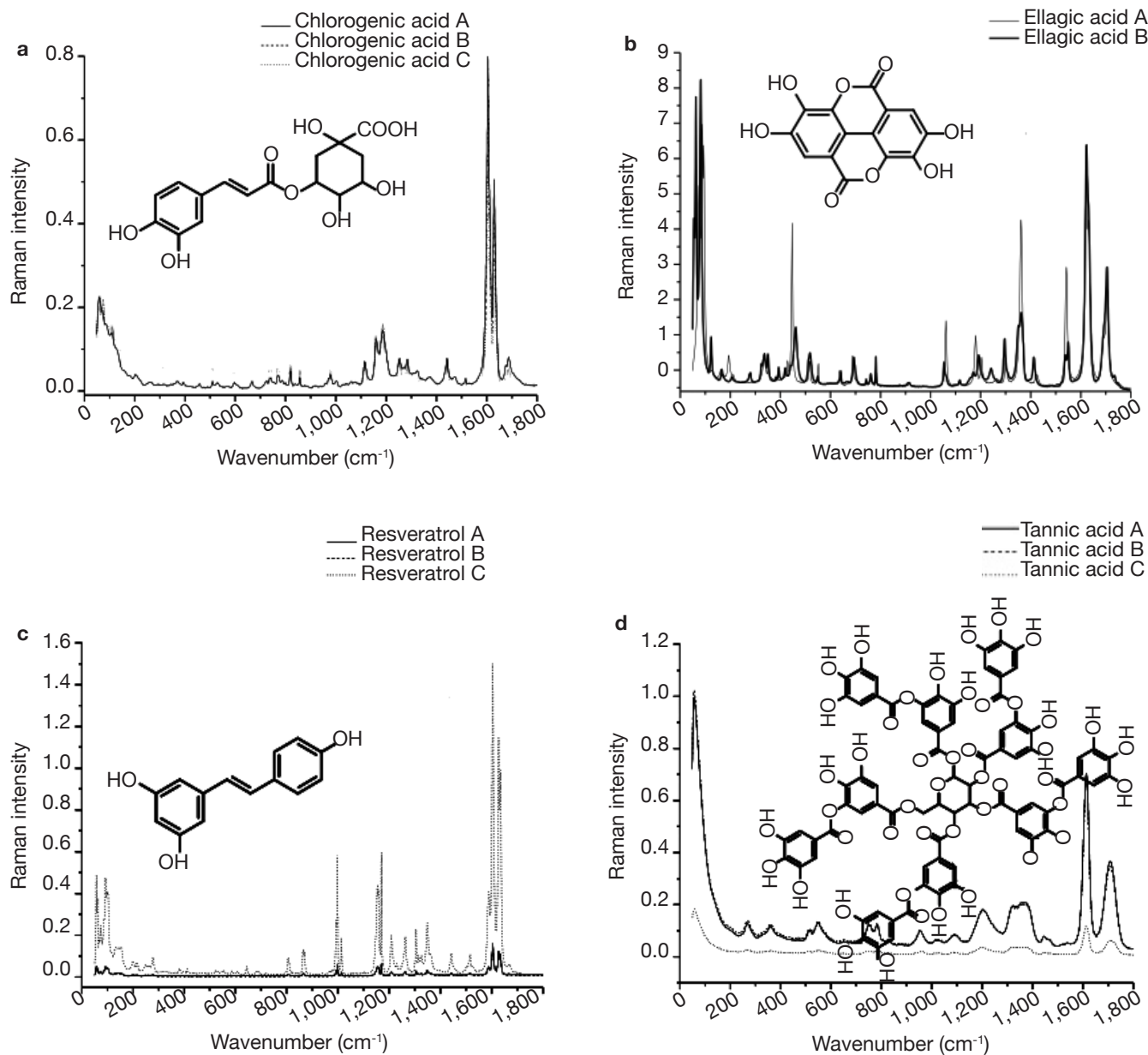


Figure 4. Raman spectra and chemical structures of chlorogenic (a) and ellagic (b) acids, resveratrol (c) and tannic acid (d) — *Spectres Raman et structures chimiques des acides chlorogénique (a) et éllagique (b), du resvératrol (c) et de l'acide tannique (d).*

Table 4. The most intense Raman scattering signals (cm^{-1}) of catechol, chlorogenic acid, resveratrol, and tannic acid — *Les signaux Raman les plus intenses (cm^{-1}) du catéchol, d'acide chlorogénique, du resvératrol et d'acide tannique.*

Derivatives of hydroxybenzoic and hydrocinnamic acids				
Ellagic acid*		Chlorogenic acid	Resveratrol	Tannic acid**
1,707 _s	1,711 _w	1,687 _w	1,634-1,635 _{vs}	1,712 _s
1,633 _{vs}	1,705 _s	1,629 _{vs}	1,627 _{vs}	1,614 _{vs}
1,626 _{vs}	1,697 _m	1,612-1,613 _{vs}	1,604 _{vs}	1,445 _{vw}
1,620 _{vs}	1,640 _w	1,604-1,605 _{vs}	1,589 _m	1,327 _m
1,613 _m	1,633 _{vs}	1,515-1,516 _{vw}	1,348-1,349 _w	1,202 _w
1,543 _{vs}	1,627 _{vs}	1,460 _{vw}	1,304 _w	1,091 _{vw}
1,368 _m	1,622 _{vs}	1,443 ± 1 _w	1,280 _w	960 _w
1,362 _{vs}	1,615 _m	1,285-1,286 _{vw}	1,169-1,170 _{vs}	788 _w
1,356 _s	1,370 _w	1,223 _w	1,158 _m	755 _w
1,180 _w	1,364 _w	1,189 _m	1,153-1,154 _m	547 _w
1,061 _m	1,355 _w	1,186-1,187 _m	997 _s	513 _w
452 _w	1,349 _w	1,157 ± 1 _w	993 _w	363 _w
448 _{vs}	462 _w	1,114 ± 1 _w	103 _m	264 _w
443 _w	123 _w	856-857 _{vw}	98-99 _m	56 _{vs}
99 _{vs}	87 _s	766-767 _{vw}	92 _m	
95 _{vs}	82 _{vs}	76 _m	86 _w	
91 _{vs}	75 _s	73 _m	73 _w	
88 _{vs}	66 _{vs}	61 _m	66 _w	
82 _{vs}	62 _{vs}	54 _m	61 _m	
75 _{vs}	58 _{vs}	52 _m	56-57 _m	

Each compound from each brand was measured independently twice. These Raman scattering signals represent the average of these measures. The Raman scattering signals were coded with vw (very weak: 5–10%), w (weak: 10–20%), m (medium: 20–30%), s (strong: 30–40%) and vs (very strong: > 40%) in function of their relative Raman intensities concerning the highest Raman intensity — *Chaque composé de chaque marque a été analysé deux fois indépendamment. Ces signaux représentent la moyenne de ces mesures. Les signaux ont été codés avec vw (très faible : 5-10 %), w (faible : 10-20 %), m (moyen : 20-30 %), s (fort : 30-40 %) et vs (très fort : > 40 %) en fonction de leurs intensités Raman relatives à l'intensité Raman la plus élevée; *: Raman scattering signals from ellagic acid brand A and B — Signaux Raman correspondant à l'acide éllagique A et B; **: Tannic acid presented just 14 peaks — L'acide tannique présente seulement 14 pics.*

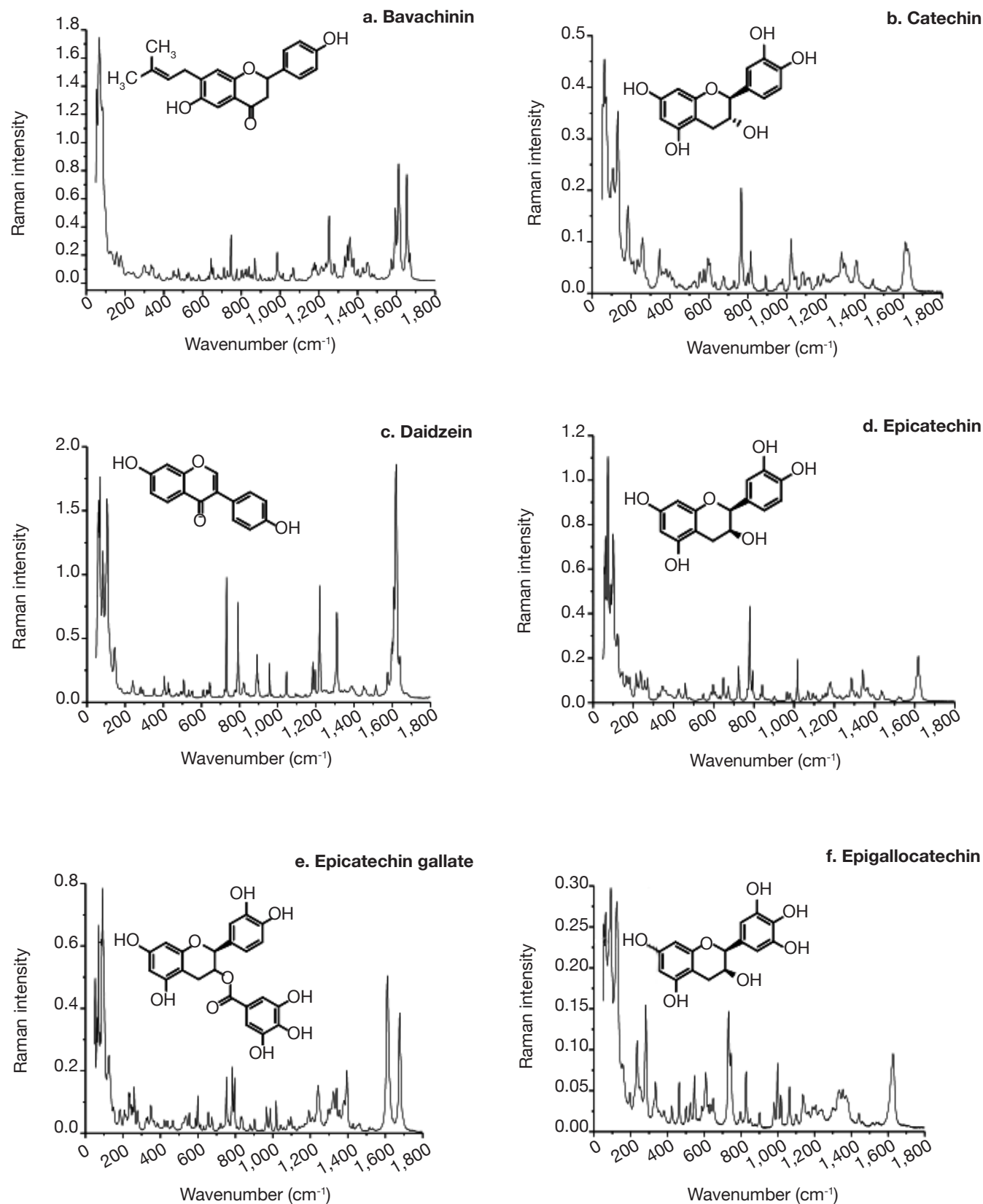


Figure 5a. Raman spectra and chemical structures of bavachinin (a), catechin (b), daidzein (c), epicatechin (d), epicatechin gallate (e), epigallocatechin (f) — *Spectres Raman et structures chimiques de la bavachinine (a), catéchine (b), daïdzéine (c), épicatechine (d), gallate d'épicatechine (e) et épigallocatechine (f).*

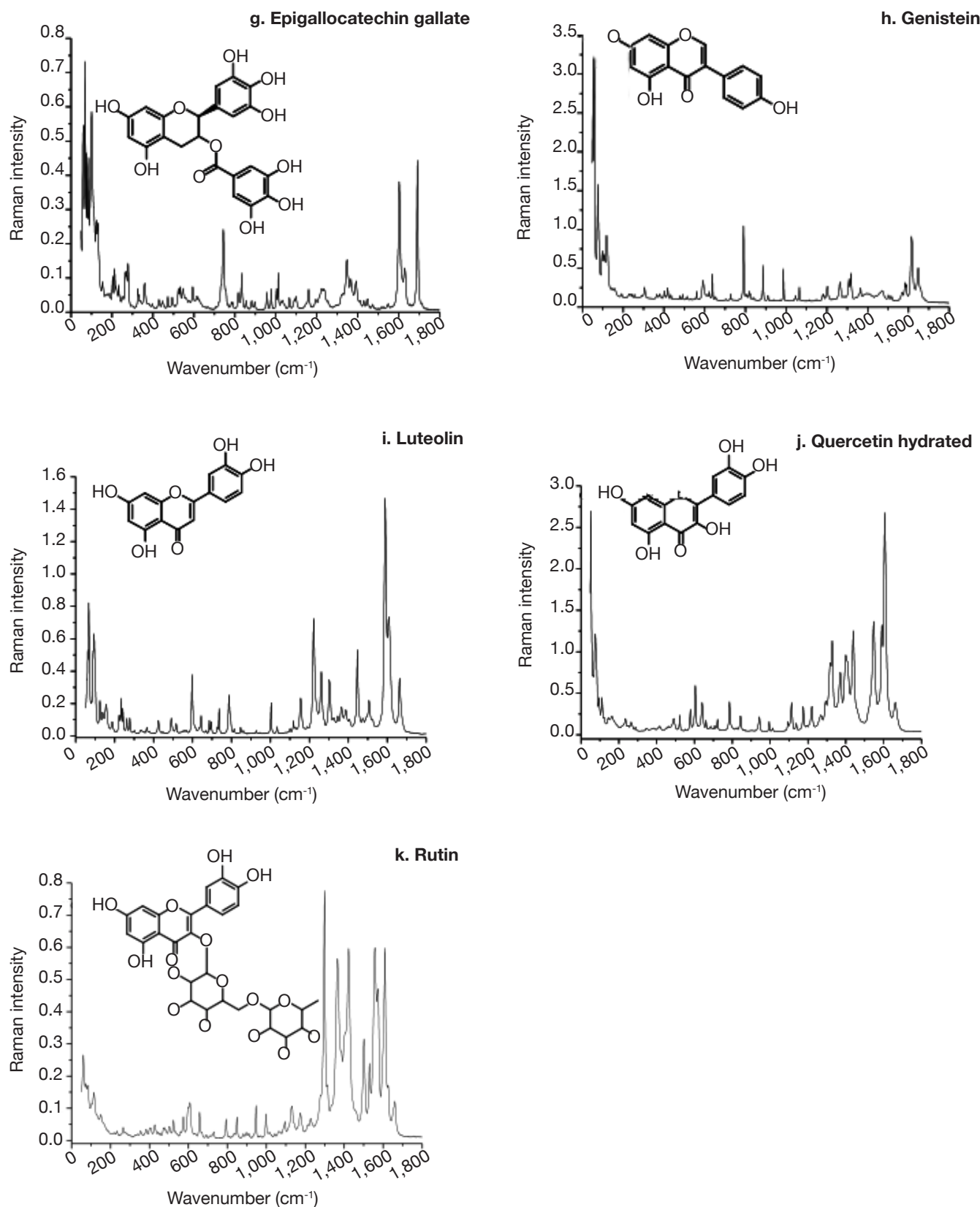


Figure 5b. Raman spectra and chemical structures of epigallocatechin gallate (**g**), genistein (**h**), luteolin (**i**), quercetin hydrated (**j**) and rutin (**k**) — *Spectres Raman et structures chimiques de la bavachinine (a), catéchine (b), daidzéine (c), épicatechine (d), gallate d'épicatéchine (e), épigallocatechine (f), gényistéine (h), lutéoline (i), quercétine hydratée (j) et rutine (k).*

Table 5. The most intense Raman scattering signals (cm^{-1}) of bavachinin (BVC), catechin (CAT), daidzein (DDZ), epicatechin (EPC), epicatechin gallate (EPCG), epigallocatechin (EPG), epigallocatechin gallate (EPGG), genistein (GNT), luteolin (LTL), quercetin dihydrated and rutin (RUT) — *Les signaux Raman les plus intenses (cm^{-1}) de la bavachinine (BVC), la catéchine (CAT), la daïdzéïne (DDZ), l'épicatéchine (EPC), le gallate d'épicatéchine (EPCG), l'épigallocatechine (EPG), le gallate d'épigallocatechine (EPGG), la génistéïne (GNT), la lutéoline (LTL), la quercétine dihydratée et la rutine (RUT).*

Flavonoïds										
BVC	CAT	DDZ	EPC	EPCG	EPG	EPGG	GNT	LTL	QUER	RUT
1,670 _w	1,633 _w	1,640 _m	1,617 _w	1,683 _m	1,627 _s	1,692 _{vs}	1,658 _{vw}	1,663 _m	1,665 _w	1,660 _w
1,653 _{vs}	1,283 _w	1,623 _{vs}	1,342 _w	1,676 _{vs}	1,000 _m	1,603 _{vs}	1,648 _w	1,589 _{vs}	1,606 _{vs}	1,610 _{vs}
1,618 _s	1,024 _m	1,618 _{vs}	1,286 _w	1,614 _{vs}	998 _m	1,012 _w	1,619 _m	1,446 _s	1,590 _{vs}	1,576 _{vs}
1,611 _{vs}	816 _w	1,608 _{vs}	1,017 _w	1,610 _{vs}	745 _s	1,003 _{vw}	1,615 _m	1,301 _m	1,548 _{vs}	1,574 _{vs}
1,594 _s	814 _w	1,599 _m	794 _w	1,609 _{vs}	735 _{vs}	834 _w	1,318 _w	1,223 _{vs}	1,441 _{vs}	1,560 _{vs}
1,347 _w	766 _{vs}	1,310 _s	779 _s	1,396 _m	733 _{vs}	744 _s	1,308 _w	1,004 _w	1,398 _s	1,558 _{vs}
1,333 _w	346 _w	1,220 _{vs}	724 _w	1,341 _w	731 _{vs}	740 _m	1,064 _{vw}	787 _w	1,328 _{vs}	1,556 _{vs}
1,253 _m	262 _m	1,196 _w	648 _w	796 _m	729 _s	279 _w	987 _w	737 _w	1,175 _w	1,553 _{vs}
1,177 _{vw}	134 _{vs}	892 _m	270 _{vw}	784 _m	282 _{vs}	265 _w	887 _w	686 _{vw}	1,114 _w	1,530 _s
985 _w	131 _{vs}	791 _{vs}	256 _{vw}	753 _m	281 _{vs}	233 _w	791 _s	597 _m	995 _{vw}	1,501 _{vs}
839 _{vw}	186 _s	775 _{vw}	235 _w	260 _w	235 _s	213 _w	728 _{vw}	263 _{vw}	785 _w	1,422 _{vs}
823 _{vw}	184 _s	731 _{vs}	214 _w	238 _w	234 _s	203 _w	638 _w	243 _w	722 _{vw}	1,365 _{vs}
776 _{vw}	183 _s	118 _m	181 _w	125 _s	124 _{vs}	131 _s	561 _{vw}	234 _w	662 _{vw}	1,299 _{vs}
747 _w	182 _s	106 _{vs}	121 _m	100 _{vs}	119 _{vs}	124 _s	491 _{vw}	124 _w	605 _m	998 _w
728 _{vw}	92 _{vs}	84 _{vs}	100 _{vs}	92 _{vs}	92 _{vs}	100 _{vs}	419 _{vw}	95 _{vs}	579 _w	947 _w
710 _{vw}	76 _{vs}	74 _{vs}	88 _{vs}	86 _{vs}	84 _{vs}	86 _{vs}	123 _m	89 _s	522 _{vw}	849 _{vw}
653 _{vw}	70 _{vs}	68 _{vs}	73 _{vs}	78 _{vs}	76 _{vs}	77 _{vs}	117 _m	72 _s	111 _w	793 _{vw}
85 _{vs}	63 _{vs}	62 _{vs}	65 _{vs}	69 _{vs}	64 _{vs}	68 _{vs}	80 _s	67 _{vs}	83 _s	657 _w
66 _{vs}	60 _{vs}	58 _{vs}	60 _{vs}	61 _{vs}	61 _{vs}	60 _{vs}	76 _{vs}	62 _s	76 _{vs}	523 _{vw}
53 _{vs}	54 _{vs}	53 _{vs}	56 _{vs}	52 _{vs}	53 _{vs}	58 _{vs}	58 _{vs}	57 _m	72 _{vs}	84 _m
							52 _{vs}		52 _{vs}	60 _s
							50 _{vs}			

Each compound was measured independently twice. These Raman scattering signals represent the average of these measures. The wavenumbers were coded with vw (very weak: 5–10%), w (weak: 10–20%), m (medium: 20–30%), s (strong: 30–40%) and vs (very strong: > 40%) in function of their relative Raman intensities concerning the highest Raman intensity — *Chaque composé a été analysé deux fois indépendamment. Ces signaux de diffusion Raman représentent la moyenne de ces mesures. Les nombres d'ondes ont été codés avec vw (très faible : 5-10 %), w (faible : 10-20 %), m (moyen : 20-30 %), s (fort : 30-40 %) et vs (très fort : > 40 %) en fonction de leurs intensités Raman relatives à l'intensité Raman la plus élevée.*

4. DIFFERENTIATION OF PHENOLIC COMPOUNDS WITHIN AND BETWEEN GROUPS

Principal component analysis (PCA) of SNV pre-treated Raman scattering signals was used to differentiate PCs within each class. For the hydroxybenzoic acids, six classes were formed and the most important Raman scattering signals that differentiated these PCs were: 1,697, 1,612, 1,600, 1,594, 1,592 and 1,198 cm^{-1} , as well as the region between 140 and 50 cm^{-1} . For

the hydroxycinnamic acids, all the PCs were well differentiated and four groups were formed. The most important Raman scattering signals responsible for this differentiation were: 1,642, 1,630, 1,614, 1,602, 1,596, 1,305-1,270, 1,224, 1,178, 1,160 and 64 cm^{-1} . The derivatives of hydroxybenzoic and hydroxycinnamic acids (chlorogenic and ellagic acids, resveratrol and tannic acid) were very well separated. The most important Raman scattering signals responsible for this differentiation were: 1,635, 1,631, 1,628, 1,626, 1,611, 1,606, 1,604, 1,348, 1,170,

997, 447, 88, 61 and 57 cm^{-1} . In the flavonoid family, three groups were clearly separated. In the first there were two flavonols, quercetin and rutin, and a flavone, luteolin; these PCs are very close chemically. In the second group there were two isoflavones, dadzein and genistein, and a prenylflavone, bavachinin. The third group consisted entirely of flavanols: catechin, epicatechin, epicatechin gallate, epigallocatechin and epigallocatechin gallate. The most important Raman scattering signals responsible for this differentiation were: 1,616, 1,608, 1,557, 1,423, 1,298, 1,222, 791 and below 130 cm^{-1} . Differentiation between classes was also done. Here, the Fisher ratio was used to select the 20 most important Raman scattering signals that allowed hydroxybenzoic (HBA) and hydroxycinnamic acids (HCA) to be differentiated, as well as their derivatives (DEV) and flavonoids (FLAV). The differentiation combinations were: HBA *versus* HCA; HBA *versus* DEV; HBA *versus* FLAV; HCA *versus*

DEV; HCA *versus* FLAV; and DEV *versus* FLAV. **Table 6** shows the Raman scattering signals that allowed these combinations to be differentiated. The 1,600-1,699 cm^{-1} and 50-199 cm^{-1} spectral ranges presented 19 and 25 peaks, respectively, that were used to discriminate the PCs. Peaks around 1,600-1,699 cm^{-1} were due to stretching vibrations of the C=C and C=O groups and those below 200 cm^{-1} were due to skeletal vibration. Another important spectral range was from 1,300 to 1,399 cm^{-1} , which had 10 Raman scattering signals. These signals were due to stretching of the CH groups and the OH bending vibrations.

Acknowledgements

The authors are grateful to the Académie de Recherche et d'Enseignement supérieur (ARES) (Belgium), Conselho Nacional de Desenvolvimento Científico e Tecnológico

Table 6. Raman scattering signals (cm^{-1}) used to discriminate groups of phenolic compounds — *Signaux Raman (cm^{-1}) utilisés pour discriminer les groupes de composés phénoliques.*

HBA vs HCA	HBA vs DEV	HBA vs FLAV	HCA vs DEV	HCA vs FLAV	DEV vs FLAV	HBA vs HCA, DEV, FLAV	HCA vs HBA, DEV, FLAV	DEV vs HBA, HCA and FLAV	FLAV vs HBA, HCA, DEV
1,697	1,706	1,697	1,705	1,692	1,706	1,697	1,697	1,705	1,697
1,642	1,697	1,676	1,642	1,642	1,634	1,628	1,642	1,692	1,628
1,630	1,634	1,619	1,634	1,630	1,628	1,615	1,631	1,653	1,614
1,614	1,628	1,609	1,627	1,615	1,604	1,604	1,615	1,628	1,603
1,603	1,604	1,554	1,615	1,602	1,589	1,598	1,602	1,604	1,556
1,595	1,598	1,421	1,605	1,366	1,557	1,362	1,595	1,594	1,398
1,300	1,361	1,365	1,600	1,300	1,423	1,300	1,300	1,362	1,364
1,224	1,315	1,316	1,595	1,287	1,398	1,289	1,287	1,299	1,304
1,198	1,289	1,288	1,300	1,273	1,299	1,224	1,273	1,288	1,287
1,186	1,265	1,221	1,287	1,224	1,221	1,170	1,224	1,224	1,221
1,178	1,170	1,198	1,273	1,186	1,170	1,158	1,186	1,170	1,198
1,160	997	802	1,223	1,178	1,155	997	1,160	1,158	1,179
749	802	779	1,178	1,160	1,043	802	776	997	1,158
139	776	731	1,154	791	997	750	123	776	791
122	139	233	997	731	776	741	111	123	731
111	122	140	776	124	125	139	106	106	124
100	106	126	117	101	101	122	100	100	100
83	83	111	94	74	81	106	83	83	83
68	76	83	80	68	69	100	76	69	69
57	55	69	57	58	55	52	57	57	59

HBA: hydroxybenzoic acids — *acides hydroxybenzoïques*; HCA: hydroxycinnamic acids — *acides hydroxycinnamiques*; DEV: derivatives of hydroxybenzoic and hydroxycinnamic acids — *dérivés d'acides hydroxybenzoïques et hydroxycinnamiques*; FLAV: flavonoids — *flavonoïdes*; These Raman scattering signals were selected using the Fisher test at $\alpha = 0.05$ — *Ces signaux Raman ont été sélectionnés en utilisant le test de Fisher à $\alpha = 0,05$.*

(CNPq) and Fundação Amazônia de Amparo a Estudos e Pesquisas do Pará (FAPESPA) (Brazil) for the financial support received to carry out this work. The authors would also like to thank very much Quentin Arnould and Sandrine Mauro from Walloon Agricultural Research Centre (CRA-W) for their technical assistance.

Bibliography

- Baeten V., Dardenne P. & Aparicio R., 2001. Interpretation of Fourier transform Raman spectra of the unsaponifiable matter in a selection of edible oils. *J. Agric. Food Chem.*, **49**, 5098-5107.
- Baeten V. et al., 2015. Vibrational spectroscopy methods for the rapid control of agro-food products. In: Nollet L.M.L. & Toldra F., eds. *Handbook of food analysis*. 3rd ed. Vol. 2. Boca Raton, FL, USA: CRC Press, 591-614.
- Baranska M. et al., 2004. Identification of secondary metabolites in medicinal and spice plants by NIR-FT-Raman microspectroscopic mapping. *Analyst*, **129**, 926-930.
- Baranska M., Schulz H., Joubert E. & Manley M., 2006. *In situ* flavonoid analysis by FT-Raman spectroscopy: identification, distribution, and quantification of aspalathin in green rooibos (*Aspalathus linearis*). *Anal. Chem.*, **78**, 7716-7721.
- Billes F., Mohammed-Ziegler I., Mikosch H. & Tyihák E., 2007. Vibrational spectroscopy of resveratrol. *Spectrochim. Acta Part A*, **68**, 669-679.
- Calheiros R. et al., 2008. Antioxidant phenolic esters with potential anticancer activity: a Raman spectroscopy study. *J. Raman Spectrosc.*, **39**, 95-107.
- Corredor C. et al., 2009. Raman and surface-enhanced Raman spectra of chrysin, apigenin, and luteolin. *Vib. Spectrosc.*, **49**, 190-195.
- Eravuchira P.J. et al., 2012. Raman spectroscopic characterization of different regioisomers of monoacyl and diacyl chlorogenic acid. *Vib. Spectrosc.*, **61**, 10-16.
- Fiuza S.M. et al., 2004. Phenolic acid derivatives with potential anticancer properties – A structure-activity relationship study. Part I: methyl, propyl and octyl esters of caffeic and gallic acids. *Bioorg. Med. Chem.*, **12**, 3581-3589.
- Ignat I., Volf I. & Popa V.I., 2011. A critical review of methods for characterisation of polyphenolic compounds in fruits and vegetables. *Food Chem.*, **126**, 1821-1835.
- ISO 5725, 1994. *Accuracy (trueness and precision) of measurement methods and result – Part 1: general principles and definitions*. Roma: FAO.
- Machado N.F.L., Batista de Carvalho L.A.E., Otero J.C. & Marques M.P.M., 2013. A conformational study of hydroxyflavones by vibrational spectroscopy coupled to DFT calculations. *Spectrochim. Acta Mol. Biomol. Spectrosc.*, **109**, 116-124.
- Mishra S. et al., 2013. Vibrational spectroscopy and density functional theory analysis of 3-O-caffeoylquinic acid. *Spectrochim. Acta Mol. Biomol. Spectrosc.*, **104**, 358-367.
- Numata Y. & Tanaka H., 2011. Quantitative analysis of quercetin using Raman spectroscopy. *Food Chem.*, **126**, 751-755.
- Paiva-Martins F., Rodrigues V., Calheiros R. & Marques M.P.M., 2010. Characterization of antioxidant olive oil biophenols by spectroscopic methods. *J. Sci. Food Agric.*, **91**, 309-314.
- Robards K. et al., 1999. Phenolic compounds and their role in oxidative processes in fruits. *Food Chem.*, **66**, 401-436.
- Schrader B., Klump H.H., Schenzel K. & Schulz H., 1999. Non-destructive NIR FT Raman analysis of plants. *J. Mol. Struct.*, **509**, 201-212.
- Schulz H. et al., 2005. Characterisation of essential oil plants from Turkey by IR and Raman spectroscopy. *Vib. Spectrosc.*, **39**, 249-256.
- Socrates G., 1997. *Infrared characteristic group frequencies. Tables and Charts*. 2nd ed. John Wiley & Sons.
- Świsłocka R., Kowczyk-Sadowy M., Kalinowska M. & Lewandowski W., 2012. Spectroscopic (FT-IR, FT-Raman, ¹H and ¹³C NMR) and theoretical studies of p-coumaric acid and alkali metal p-coumarates. *Spectroscopy*, **27**, 35-48.
- Świsłocka R., Regulska E., Samsonowicz M. & Lewandowski W., 2013. Experimental and theoretical study on benzoic acid derivatives. *J. Raman Spectrosc.*, **1044**, 181-187.
- Teslova T. et al., 2007. Raman and surface-enhanced Raman spectra of flavone and several hydroxy derivatives. *J. Raman Spectrosc.*, **38**, 802-818.
- Volf I., Ignat I., Neamtu M. & Popa V.I., 2014. Thermal stability, antioxidant activity, and photo-oxidation of natural polyphenols. *Chem. Pap.*, **68**(1), 121-129.
- Zuk M. et al., 2011. IR and Raman studies of oil and seedcake extracts from natural and genetically modified flax seeds. *Spectrochim. Acta Part A*, **78**(3), 1080-1089.

(24 ref.)