Postprint of Roca, M., Ríos, J.J. & Pérez-Gálvez, A. Phytochem Rev (2017). https://doi.org/10.1007/s11101-017-9543-z

Mass Spectrometry: The Indispensable Tool for Plant Metabolomics of Colourless Chlorophyll Catabolites

María Roca[1], José J. Ríos[2] & Antonio Pérez-Gálvez[1]*

[1]*Food Phytochemistry Department, Instituto de la Grasa (CSIC), Campus Universitario, Building 46, 41013, Sevilla, Spain*

[2]*Laboratory of Mass Spectrometry, Instituto de la Grasa (CSIC), Campus Universitario, Building 46, 41013, Sevilla, Spain*

*author for correspondence at Instituto de la Grasa (CSIC), Campus Universitario, Building 46, 41013, Sevilla, Spain. Email: aperez@cica.es. Phone. +34954611550.

Abstract

Senescence and ripening of plant tissues engage the pheophorbide *a* oxygenase pathway, reducing the chlorophyll content to inactive chlorophyll catabolite products, termed phyllobilins. These products are open-macrocycle derivatives, but present different structural features related to species-dependent enzyme activity. This review encompasses a brief outline of the chlorophyll catabolism pathway, a detailed description of the structural motifs of known phyllobilins, giving details of how mass spectrometry provides hints for the characterization of phyllobilins. The structural approach for the identification of phyllobilins requires several spectroscopic methodologies to reach a complete structural identification, including UV-visible spectroscopy, circular dichroism, nuclear magnetic resonance and mass spectrometry. Among these techniques, mass spectrometry presents several advantages for showing the structural features of phyllobilins, through acquisition of accurate mass, elemental composition, and detection of product ions, which provide valuable structural information. The combination of mass spectra with data-managing and *in silico* prediction tools greatly enhances the comprehensive building of the phyllobilin structure, and the resolving of the intriguing puzzle of enzymatic and chemical reactions involved in chlorophyll catabolism. Indeed, some strategies based on structural constraints that phyllobilins present, with recent developments in software prediction tools are proposed to foster the unravelling of phyllobilin structures.

Keywords: Chlorophyll catabolism, mass spectrometry, phyllobilins, structure elucidation, tetrapyrrole

Abbreviations: PaO, pheophorbide *a* oxygenase; NCC, non-fluorescent chlorophyll catabolite; RCCR, red chlorophyll catabolite reductase; FCC, fluorescent chlorophyll catabolite; DFCC, dioxobilin-type fluorescent chlorophyll catabolite; DNCC, dioxobilin-type non-fluorescent chlorophyll catabolite; DESI, desorption electrospray ionization.

Introduction

The fate of chlorophyll during the continuous chlorophyll turnover, or through senescence of leaves and ripening of fruits, remained puzzled until the identification and structural elucidation of the first non-green chlorophyll catabolite, the *Hv*-NCC1 (Kräutler et al. 1991) a compound that provided elements for establishing the chlorophyll breakdown pathway. Since then, the accumulation of different chlorophyll catabolites has attracted scientific interest from a range of diverse perspectives.

The transition of green tissues to senescent ones occurs for several reasons and means the establishing of different metabolic processes. The decline and dismantling of the photosynthetic apparatus produce free chlorophyll molecules, which would lead to uncontrolled photo-oxidative damage in a tissue that still needs suitable living conditions (Hörtensteiner 2004). Consequently, the metabolic machinery of plant tissues has provided a controlled degradation route for the release and accumulation of photo-inactive chlorophyll catabolites. The route is known as the PaO (Pheophorbide *a* Oxygenase) pathway, which controls the structural rearrangement of chlorophylls into their catabolites. The significance of this route stands out as it processes an estimated total amount of 10^9 tons of chlorophylls per year, yielding considerable amounts of chlorophyll metabolic products accumulated in senescent tissues (Kräutler et al. 1991). The first main part of the PaO pathway takes place in the chloroplast and involves a family of enzymes that transform the tetrapyrrolic macrocycle into an open and reduced fluorescent structure (Hörtensteiner and Kräutler 2011; Kräutler and Hörtensteiner 2014). This catabolite is released into the stroma to experience further structural rearrangements, which are species-specific by means of different enzymatic processes, some of them still not well understood. Finally, the modified fluorescent catabolite is translocated into the vacuole and non-enzymatically isomerized to the non-fluorescent chlorophyll catabolite (NCC).

The structural approach concerning the identification of the chlorophyll breakdown products that are accumulated in senescent tissues, either leaves or fruits, has been developed by means of their different spectroscopic features including the UV-visible, circular dichroism, nuclear magnetic resonance, and mass spectrometry (MS^n) spectra, all of them necessary to definitively ascertain the structure, nature of chemical substituents attached to functionalized groups, and spatial configuration. The structural constraints that all the chlorophyll catabolites seem to follow according to the structures identified so far, relieve the acquisition of structural information of each compound. The core structure is inherited from the pheophorbide *a* precursor and a limited group of functions is introduced in the peripheral positions and at well-defined locations, that is, specific groups are attached to a specific place and not to the other ones (see Table **1** and Fig. **1**). Additionally, spatial configuration presents restrictions in the chiral centres of the chlorophyll degradation derivatives (Kräutler 2014).

Consequently, interpretation of spectroscopic data to get the final elucidated structure is alleviated by these boundaries that even help to design strategies for the acquisition and analysis of data, either when those spectroscopic data lead to known structures or to unknown ones. Although the use of a single analytical method is not enough to completely elucidate the chemical structure of molecules, it is possible to selectively detect compounds from a chemical family in complex biological media and to obtain their elemental composition with high resolution time of flight (TOF) mass spectrometers fitted with electrospray (ESI)/atmospheric pressure chemical ionization (APCI)-sources in combination with high-performance liquid chromatography. Thus, the combination of accurate mass measurements and dedicated acquisition and data-managing software with *in silico* prediction database software tools allows MS^n experiments to be performed in a single run, and enables to go deeper in the structural characterization of those compounds. MS has been a fundamental technique for the structural analysis of chlorophyll breakdown products since they were

discovered (Müller et al. 2014), as well as for a comprehensive building of the reaction mechanisms involved in the known enzymatic steps of chlorophyll catabolism (Hörtensteiner et al. 1998).

It is important to highlight that the numbering system is different in chlorophylls than in open-macrocycle chlorophyll catabolites. As it is shown in Fig. **1**, RCCs, FCCs, DFCCs, NCCs and DNCCs follow the numbering system of linear tetrapyrroles (Kräutler 2014), and this atom labelling rule is used in this work, while the atom numbering of porphyrins is applied for the chlorophyll precursors when required. The open chlorophyll catabolites are grouped under the term phyllobilins, reminiscent of the linear structure of bilins. The current nomenclature of the open-macrocycle chlorophyll catabolites is an inherited system that due to the increasing identification of new chlorophyll catabolites is non-functional and probably will require an update. The model includes the initials of the Latin name of the species in which the compound has been identified, followed by the initials of the type of opened chlorophyll catabolites and finally a number reflecting the polarity in the HPLC chromatogram. Consequently, the same compound presents different names depending of the number of plant species in which it has been identified, as displayed in Table **1**.

The literature survey details data regarding accurate mass and elemental composition of chlorophyll degradation compounds, and common product ions that come from fragmentation of protonated ions, which provide information for structural characterization of this group of linear tetrapyrroles.

Structures and metabolism of chlorophyll degradation

All the chlorophyll breakdown products identified up to now (Pérez-Gálvez and Roca 2017) derive from pheophorbide *a*, except *At*-NCC3 (Müller et al. 2006). The catabolic pathway starts in the chloroplast but continues through the cytosol and ends in the vacuole (Kräutler 2014). Pheophorbide *a* is oxygenolytically opened at the ‘Northern’ *meso*-position

(C4-5, porphyrin numbering) by pheophorbide *a* oxygenase (PaO) enzyme (Hörstensteiner et al. 1995) delivering the red chlorophyll catabolite (RCC), the first intermediate catabolite. RCC is rapidly metabolized and consequently it does not accumulate in the cells. Following this, RCC is reduced stereo specifically by RCC-reductase enzyme (RCCR) in the C15-16 (Fig. **1**) yielding the primary fluorescent chlorophyll catabolite (*p*FCC). Each plant species presents one type of RCCR activity; there is RCCR-type I that forms *p*FCC and RCCR type-II of which *epi*-*p*FCC originates, and both compounds are epimers at C16 (Pružinska et al. 2007). FCCs are currently defined as 1-formyl-19-oxobilins to be differentiated from DFCCs, the 1,19-dioxibilins (Kräutler 2014). D*p*FCCs are formed by deformylation at C1, with this reaction being catalysed by the cytochrome P-450, CYP89A9 (Christ et al. 2013). Once the route reaches this point, *p*FCC and D*p*FCC experience the same series of reactions, introducing new functional groups at specific positions, but although different types of FCCs have been identified, only one DFCC has been characterized up to now (Süssenbacher et al. 2015). The first modification is the hydroxylation at $C3^2$, yielding secondary FCCs (*s*FCCs) (Moser et al. 2012), a reaction catalysed by the hydrolase TIC55 (Hauenstein et al. 2016) taking place at the chloroplast. Both FCCs (*p*FCC and *s*FCC) are exported actively to the cytosol by an unknown transporter (Matile et al. 1992), where further functional groups are introduced in the structure creating the modified FCCs (*m*FCC). Only two *m*FCCs (*At*-FCC1 and *At*-FCC2) have been identified (Pružinska et al. 2005) due to the fleeting nature of these chlorophyll catabolites, which are rapidly transformed. The de-esterification at $O8^4$ creates *At*-FCC2, a process carried out by the enzyme MES16 (Christ et al. 2012), while *At*-FCC1 presents an additional hydroxylation at C3 besides the de-esterification at $O8^4$. As it will be described later, all the modifications identified in the final chlorophyll catabolites occur at *m*FCC structure in the cytosol, but the identification of most *m*FCCs remains elusive except for the two compounds noted above. An interesting modification from *m*FCCs consists in the

esterification with different functional groups in the propionic acid at C12, yielding the hypermodified FCCs (*hm*FCCs). The importance of such modification lies in the fact that the propionic acid should be kept unbound to allow further metabolism of the *m*FCCs. Consequently, *hm*FCCs are not transformed to NCCs and are easily recognized by their characteristic UV spectrum, allowing the detection of eight different compounds (Table **1**). The fluorescent chlorophyll derivatives are transported to the vacuole through active transport mediated by an ATP binding cassette (ABC) transporter (Hinder et al. 1996) not yet identified. The acidic pH of the organelle induces a non-enzymatic isomerization yileding non-fluorescent chlorophyll catabolites NCCs/DNCCs through a reaction mechanism that requires a free propionic acid at C12 (Oberhuber et al. 2003).

Table **1** shows the different NCCs/DNCCs identified up to date, all of them with specific functional groups at fixed positions. The first one is the vinyl group at C18 that can be hydroxylated or not, the second position is the methyl group at $O8^4$ that can be de-esterified, and finally the hydroxylation at $C3^2$ that allows further esterification with O-β-glucopyranosyl, O-β-malonyl or O-β-(6´-O-malonyl)-glucopyranosyl groups. The number of phyllobilins esterified with glucopyranosyl has recently increased, including occasionally compounds with a different position for the esterification than the usual $C3^2$ one. This is the case of the *Pd*-NCC32, a phyllobilin that accumulates in the senescent leaves of the plum tree (*Prunus domestica* L.) and presents one glucopyranosyl group at the $O18^3$ position, and another one at $C3^2$ (Erhart et al. 2016). The senescent leaves of the wych elm tree (*Ulmus glabra* Huds.) contains a NCC, the *Ug*-NCC53, with one glucopyranosyl unit attached both to the $C3^2$ position and to $O12^4$.

MS techniques applied for analysis of colorless chlorophyll catabolites

Analysis by MS of porphyrins in general, and particularly of phyllobilins, required the introduction of direct-insertion probes in the design of MS equipment, including first the fast-

atom bombardment (FAB) source, and later APCI, ESI and matrix-assisted laser desorption/ionization (MALDI). Among these sources, ESI has prevailed in the MS analysis of porphyrins. It is easily coupled to liquid chromatography for *on-line* separation, requires minute amounts of samples to obtain high-quality mass spectra, and allows acquisition of mass features of the compounds (accurate mass and elemental composition), as well as the helpful structural information obtained from tandem MS/MS fragmentation experiments. These analytical capabilities are greatly improved when the mass analyser performs the acquisition of exact mass measurements either from the protonated ion and their product ions, so that reliable information about molecular composition and significant structural arrangements drive to unequivocal identification and characterization of the main structural determinants.

The constitution of chlorophyll breakdown products started with the experimental research of Philippe Matile and Bernhard Kräutler, who applied combinatorial spectrophotometric measurements to ascertain the structure of *Hv*-NCC1, the first characterized chlorophyll breakdown product (Kräutler et al. 1991). In addition to NMR spectroscopic data, they performed exact mass determinations and obtained the FAB-MS data of the target compound, isolated by semi-preparative HPLC, in a matrix of 3-nitrobenzyl alcohol (NOBA) with cesium bombardment, detecting some product ions arising from the parent ion. Particularly, the loss of ring D unit became the first characteristic probe for NCCs in MS. The subsequent identification of NCCs in further plant tissues produced new characteristic FAB-MS data. Thus, the characterization of the chlorophyll breakdown product in senescent rape leaves (*Brassica napus*), the *Bn*-NCC1, included MS data acquired with FAB in positive mode, spectra where the base peak appears at $[M+K]^+$, while the loss of ring D and some other product ions, which were not characterized, are observed at lower mass regions (Mühlecker et al. 1993).

Several of the advantages that MS offers in deriving the concentrations and fluxes of metabolites were applied very soon to follow *in vivo* the advances of the biochemical pathway that controls the catabolism of chlorophyll. Thus, Curty et al. (1995) demonstrated that a monooxygenase enzyme catalyses the introduction of formyl oxygen ($O1^2$) from dioxygen, while the lactamic oxygen atom (019^1) derives from water. Data to support this conclusion were obtained from FAB-MS measurements of the products obtained in different incubation experiments *in vivo* under ^{18}O-enriched oxygen atmosphere (Curty et al. 1995). Such experimental approaches could be achieved only with the high sensitivity of MS and the possibility of screening different metabolites in minute amounts.

As new plant species were analysed, FAB source continued to be the standard configuration for acquisition of MS spectra of NCCs once the compounds were isolated and purified by semi-preparative HPLC. In addition to the NOBA some other alternative matrices were used, including glycerol or glycerol/1-thioglycerol (1:1). The acquisition of the mass spectra still required the isolation of the target compound(s) by semi-preparative HPLC, and the MS information in addition to UV-visible, CD and NMR spectroscopic data finally yielded the structural constitution. Nevertheless, the list of NCC structures started to increase, and they became about a dozen different compounds at the turn of the century (Berghold et al. 2004). The experimental approach in the initial works consisted in the same analytical routine: extraction from the senescent tissue with a mixture of organic/buffer solvent, then liquid chromatography and screening of UV-visible spectrum of each chromatographic peak to find those ones that fit with the characteristic UV-visible spectrum of an NCC compound, then isolation of the pure compound(s) and finally acquisition of NMR, FAB-MS and CD spectra, as well as other experiments required to confirm the stereochemical pattern (Oberhuber et al. 2001). The structural characteristics of the NCCs consist in the basic linear tetrapyrrole skeleton inherited from the pheophorbide *a* precursor, with different residues at

the three key positions indicated above, $C3^2$, $O8^4$ and C18, which are arranged following definite combinations (Table **1** and Fig. **1**) (Kräutler 2014; Pérez-Gálvez and Roca 2017). The controlled structural arrangements are consequences of the activity of the early described enzymatic machinery that performs such changes, starting with the basic *p*FCC structure (Fig. **1**).

The work of Losey and Engel (2001) meant a significant advance in the identification and characterization of phyllobilins in plant species. They isolated a non-fluorescent chlorophyll catabolite from *Hordeum vulgare* L. that presents an oxidative deformylation of the *p*FCC, so that the oxygen arrangements at the open macrocycle are like those in urobilinogen. This was the first identified compound of the second line of bilin-type chlorophyll catabolites, the type-II phyllobilins or dioxobilin-type NCCs (DNCCs), a line that slowly received the incorporation of several new structures. But a second improvement of that work was the application of ESI in the MS analyses of the dioxobilin identified in *Hordeum vulgare* L. (Losey and Engel 2001). The information obtained from the ESI-MS spectra of the DNCC was basically the same as it would have attained from FAB-MS, that is, exact mass measurement and elemental composition of the compound. Moreover, the MS experiments were performed with the isolated compounds, that is, the high capabilities of the hyphenation of HPLC with MS were still not brought.

It was from the isolation of new NCC structures of *Spinacia oleracea* L. and *Nicotiana rustica* when the identification of novel characteristic fragmentation patterns in the ESI/FAB-MS was improved. Thus, in addition to the already noted fragmentation of ring D, those NCCs with a methylated function at $O8^4$ display a prominent signal corresponding to the loss of 32 Da (CH_3OH neutral loss) (Berghold et al. 2003; 2004), while decarboxylation proceeds when the carboxylic acid group is un-esterified (Berhold et al. 2006). The detection and identification of a yellow chlorophyll catabolite (*Cj*-YCC) arising from the dehydrogenation

of *Cj*-NCC1 (Fig. **2**) introduced the loss of ring A as a new fragmentation route from ESI-MS in phyllobilins (Moser et al. 2008). The presence of other available functional groups located at the side chains of the linear tetrapyrrolic skeleton are also assigned with the product ions derived from the corresponding fragmentation process. Thus, the functionalization of the hydroxyl group at $C3^2$ with a sugar moiety yields the product ion with a $C_6H_{10}O_5/C_6H_{12}O_5$ loss from the protonated ion, observed in the MS spectra of *Nr*-NCC2, *At*-NCC1, *Zm*-NCC1 and *Zm*-NCC2 and even in the MS spectra of those *hm*FCCs containing a sugar residue at that position (Moser et al. 2009). The MS spectra of these compounds were acquired in a liquid secondary ion mass spectrometer using glycerin as the matrix and internal standard.

Hence, the map of structural assignments in MS spectra of phyllobilins enlarged as result of the growing number of chlorophyll catabolites detected in plant species. And the technical developments and analytical approaches introduced in MS technology were applied for acquisition of MS data and rapid identification of phyllobilins. Thus, the MALDI technique coupled to a time-of-flight (TOF) mass analyser, using 2,5-dihydroxybenzoic acid as matrix, was applied to the acquisition of MS in the identification of phyllobilins from the senescent leaves of the peace lily (*Spathiphyllum wallissii*) (Kräutler et al. 2010). Additionally, the ESI-MS/MS data contributed to characterizing the residues esterifying the propionic acid tail of the *hm*FCC detected in senescent leaves of the peace lily. Another ionization technique known as desorption electrospray ionization MS (DESI-MS) allowed direct analysis of NCCs from leaf tissues within seconds (Müller et al. 2011). Moreover, tandem mass spectrometry and reactive-DESI experiments contributed to increasing the sensitivity and selectivity of the target compound(s) by means of *in situ* derivatization of the oxygenated functions and acquisition of the MS spectra. This strategy was applied to the acquisition of structural information of the underivatized chlorophyll catabolite(s). With the advent of MALDI and DESI as ionization techniques with other technical improvements in

the mass analysers, and the already consistent knowledge of the MS behaviour of phyllobilins, the capabilities that MS offers started to be fully exploited. Indeed, MS data were the main (and single) reliable experimental proof in the identification of several phyllobilins. Some examples are the *Tc*-YCC described in senescent leaves of *Tilia cordata* tree (Scherl et al. 2012), certain NCCs present in ripe peels of *Musa acuminata* (Moser et al. 2012), and several DNCCs observed in *Arabidopsis thaliana* Heynh. wild-type (Christ et al. 2013) and in *Arabidopsis thaliana methylesterase16* mutant (Süssenbacher et al. 2014). Also, several hydroxylated analogues to known NCCs have been identified solely based on their UV and MS spectra (Süssenbacher et al. 2015).

The scarce amounts that plant tissues accumulate of these compounds made their isolation for acquisition of NMR spectroscopic data unfeasible, and therefore the identification should rely on UV-visible and MS spectra in combination with tandem MS/MS spectrometry. Consequently, this made necessary to take most of the capabilities of MS, including the hyphenation of HPLC with MS and the use of algorithms and evaluation of experimental data with advanced software tools (Müller et al. 2014; Pérez-Gálvez and Roca 2017; Ríos et al. 2015; Chen et al. 2015a, b). This was the main strategy applied for the identification of phyllobilins in ripe fruits of *Cydonia oblonga* Mill. tree (Ríos et al. 2014a). The MS spectra of *Co*-phyllobilins were acquired in high-resolution mode to measure the accurate mass. The experimental results were compared with an *in-house* mass database containing the elemental composition and monoisotopic masses of already identified phyllobilins. The software performs the comparison automatically and generates a report list with those experimental ion chromatograms values that are present in the database and with mass error values below a threshold (usually <3 ppm). A second data screen is the evaluation of the experimental and theoretical isotopic patterns. The software performs the calculation of the match factor based on deviations of the signal intensities and isotopic profile yielding a

numerical value (mSigma value). Only the positive hits that fulfil mSigma value below a threshold (usually <50) are included in the final report list. Therefore, two criteria (mass error and isotope pattern factor) delimitates the identification of NCCs/DNCCs with this strategy in addition to UV-visible characteristics. Even more, MS^2-based reactions contribute to ascertaining the functionalization(s) present in the chlorophyll catabolite. Finally, the injection of plant tissues extracts containing already characterized NCCs/DNCCs provides additional evidences to definitively assign each compound of the HPLC-MS profile in the target tissue to some of the existing NCCs/DNCCs. Thus, two of the NCCs/DNCCs described in *Cydonia oblonga* Mill., *Co*-NCC1 and *Co*-DNCC2, were identified as structural isomers equivalent to *Zm*-NCC1/*Tc*-NCC1 and *Hv*-DNCC1, respectively. Comparison of the spectroscopic characteristics of chlorophyll catabolites in senescent florets of broccoli (*Brassica oleracea* var. *italica*) was also the strategy for screening the structural equivalences of type I *Bo*-phyllobilins and type II *Bo*-phyllobilin with those described in *Brassica napus* and *Arabidopsis thaliana* Heynh., respectively (Roiser et al. 2015). Full structural characterization would require isolation of the compounds to perform 2D-NMR spectroscopy and further experiments to describe stereochemical conformation.

This experimental approach directly applies for the known components, but with a few adjustments it greatly boosts identification of the unknown ones. First, the database can include an elemental composition of new NCC structures, but considering that only those substituents/functions already detailed in the literature can enter in the linear tetrapyrrolic skeleton and at the same carbon atom positions, a few number of new combinations is possible. Some of the examples of *new* NCC structures that follow the structural constrains described for NCCs are depicted in Fig **3**. Moreover, once these NCC structures are generated, the corresponding DNCC alternatives are figured out and included in the *in-house* database. Therefore, the software performs the screening for these *partly* unknown

chlorophyll catabolites, at least not described so far. With this strategy, two *Co*-phyllobilin structures, *Co*-NCC2 and *Co*-DNCC1 were identified in the senescent peels of quince fruits (*Cydonia oblonga* Mill.) (Ríos et al. 2014a), two new chlorophyll catabolites were described in loquat fruits, *Ej*-NCC2 and *Ej*-DNCC1 (Ríos et al. 2014b) and one structure in peels of lemon fruits (*Citrus limon* L.), the *Cl*-NCC3 (Ríos et al. 2015).

The query arises when the MS spectral features of an *unknown* with the typical UV-visible profile of NCCs/DNCCs do not fit with any elemental composition and monoisotopic masses included in the database. Then, it is necessary to come back to the routine analytical methodology, i.e. isolation of the compound and acquisition of the spectroscopic characteristics including NMR, MS and other experimental contributions. However, MS and tandem MS/MS are the pillars to quickly obtaining a first notion of the main structural arrangements of the *unknown* chlorophyll catabolite. Thus, the careful analysis of the MS behaviour, including MS^2-based reactions in the gas phase, of the observed unknown phyllobilins in addition to the knowledge of mass spectrometric profiles delineated previously with the known chlorophyll catabolites, definitively contribute to the structural elucidation. This knowledge is the single approach to allow tentative detection of chlorophyll catabolites when the UV-visible spectrum is unclear, due to their accumulation in slight amounts or when they co-elute with other components. The available knowledge regarding MS and MS/MS behaviour of chlorophyll catabolites is precisely described in the next section, and it is the heart for building mass spectrometry libraries aimed at the identification and quantification of phyllobilins in plant tissues (Christ et al. 2016). The resolution of the database may break solely with the information regarding already identified structures, or may include mass spectrometric features of unknowns as stated above.

Characteristics of mass spectra of colourless chlorophyll catabolites

MS behaviour of phyllobilins should follow similar patterns as described for other linear tetrapyrroles like bilirubin and its derivatives, and phycobilins. Three main reaction pathways yield the characteristic product ions in linear tetrapyrroles: fragmentation of side chains, internuclear cleavage and formation of tricyclic product ions (Jackson et al. 1967). The MS will be dominated by some of the product ions arising from those cleavages, independently of the nature of the substituents located at the side chains. The analysis of the product ions arising from MS of phyllobilins described in the following sections has been ordered considering the reaction pathways noted above, aimed to be a reference for future structural elucidation of phyllobilins. Table **2** contains the significant losses from protonated molecular ions and the structural hint they provide, and Fig. **4** depicts the sites for observed fragmentation processes and possible numerical losses in NCCs.

Product ions from side chain fragmentations

The first product ion that may arise from fragmentation of side chains is that formed after the loss of 32 Da (methanol) from the protonated molecular ion. This fact points to one structural arrangement, that is the methylated form of the carboxyl function located at the $C8^2$ position (R_4 in Fig. **4**) and consequently that the phyllobilin presents an intact methoxycarbonyl group as the immediate FCC precursor does. The product ion resulting from this fragmentation presents a ketene function that subsequently may yield secondary product ions arising from the interaction of that ketene function with the vicinal propionic residue. The recently published compendiums of type-I phyllobilin structures (Erhart et al. 2016; Pérez-Gálvez and Roca 2017) indicate that most of the structures retain the methoxycarbonyl group originally present in the pheophorbide *a* precursor and in subsequent intermediates of the phyllobilins pathway, while few plant species present activity of MES16, the methylesterase enzyme responsible for demethylation at the $C8^2$ group (Christ et al. 2012), so that the 32 Da loss is very common. The cleavage of the β-keto ester group (Fig. **4**) may also

take place yielding a 60 Da loss product ion (Table **2**), and the same reaction mechanism described for the same loss in 13^2-hydroxy-chlorophyll derivatives applies (Canjura and Schwartz 1991; Chen et al. 2015a, b), although for those parent derivatives the fragmentation of the β-keto ester group is frequently in sequence with an additional loss of 18 Da. Only Bale et al. (2010) describe the loss of 60 Da from RCC and FCC structures isolated from laboratory culture of the alga *Chlorella protothecoides* assuming a hydrogen transfer during ionization to yield such fragmentation. This loss is particularly favoured in those chlorophyll catabolites where an extended conjugated double bond system remains, that is, in RCC and to a lesser extent in FCCs, while the NCC structures are prone to yield a 32 Da loss, i.e., MeOH from the β-keto ester group, as they lack the original pheophorbide *a* conjugated double bond system. Indeed, product ions corresponding to lower losses (18 Da and 28 Da, i.e., H_2O and CO, respectively) are more recurrent in the NCC structures (Bale et al. 2010). Some of these characteristic fragmentations are observed in the tandem mass spectra of phyllobilins included in the Appendix.

Only two NCCs structures do not present a hydroxylated function at $C3^2$, *At*-NCC3 and *So*-NCC5 and their corresponding structural isomers present in other plant species. Consequently, it is common to detect product ions indicating the presence of that hydroxylated function at $C3^2$, which may be bonded to a sugar moiety with an extra malonyl residue, or only to a malonyl chain (Fig. **4**). Indeed, an additional loss of sugar may indicate the presence of a second attachment to the 18^1,18^2-dihydroxyethyl group at ring D described in senescent leaves of the plum tree (*Prunus domestica* L.) (Table **1**) (Erhart et al. 2016). Caution must be taken as such product ions could be produced in combination with a H_2O loss (Scherl 2012; Ríos et al. 2014a, b).

The propionic chain at the C12 position may exhibit an ester function with different moieties like methyl or daucic acid as shown for extracts from peels of ripening bananas

(*Musa acuminata*) (Moser et al. 2012), either for *Mc*-*hm*FCCs or *Mc*-NCCs structures, and digalactosylglyceryl or the glucopyranosyl (3,4-dihydroxyphenyl)-ethyl ester groups as described for senescent banana leaves (Banala et al. 2010; Vergeiner et al. 2013). Although the product ions arising from cleavage of such ester residues have not been described (R_5 in Fig. **4**), they should appear under the convenient MS experimental conditions and help to successfully establish nature of peripheral residues of the chlorophyll catabolite structure.

Product ions from cleavage at the *meso*-positions

The most significant product ions for structural elucidation of chlorophyll catabolites are those arising from cleavage at the *meso*-positions as they provide valuable information about the chemical modifications produced in the original rings A and D of the *p*FCC precursor. Particularly, it is common to find the product ion produced after fragmentation of ring D at the C15-16 bond, yielding a neutral monopyrrole unit (Fig. **4**). The analysis of such fragmentation indicates whether ring D presents the original 18^1,18^2-vinyl arrangement or if it has been modified to the 18^1,18^2-dihydroxyethyl version, and whether this function carries an additional sugar moiety as described for the senescent leaves of the plum tree (*Prunus domestica* L.) (Erhart et al. 2016). The lack of cyclic structure and conjugation are the driving factors that promote the loss of this monopyrrole unit in linear chlorophyll catabolites by means of a proton transfer mechanism that may also originate from the C14-15 bond. Some of these characteristic fragmentations are observed in the tandem mass spectra of phyllobilins included in the Appendix. The second *meso*-position available for fragmentation is at the C4-5 bond producing the loss of ring A as a neutral monopyrrole unit that would provide additional information about the chemical constitution of the substituents bonded to this ring (Fig. **4**). The proton transfer mechanism also applies for this fragmentation and such process may also start at the C5-6 bond. Consecutive fragmentation of both rings has been described for NCC precursors although additional evidences from MS^n studies were not available (Bale

et al. 2010). Finally, the cleavage at the C10-11 bond would yield product ions consisting of each half of the NCC structure. The analysis of the chemical constitution of these product ions introduces valuable data for the indirect structural assignment of those units (modifications at rings A, D, and the propionic chain ester group if present), and this configuration would not have been fully confirmed by the product ions coming from direct fragmentation of them.

Conclusions

MS has successfully provided relevant information for the structural elucidation of phyllobilins, which were considered for long time elusive products of the chlorophyll breakdown pathway. The continuous developments both in MS equipment and experimental strategies work together to ascertain the molecular building that phyllobilins follow, revealing the existence of three sub-groups of chlorophyll degradation products, *hm*FCC, NCCs and DNCCs, with different structural characteristics. Even MS is a fundamental contributor to unmasking the intriguing battery of enzymatic processes and chemical reactions involved in the pheophorbide *a* oxygenase route. MS allows the acquisition of significant data in minute amount samples, which could be chemically derivatized in order to achieve immediate information of molecular characteristics. Indeed, the incorporation of modern predictive software tools to the data analysis process speeds up the interpretation of results and increases the certainty level. Finally, MS is conveniently coupled to *in silico* platforms containing plenty of mass spectra information about both already known phyllobilins and unravelled phyllobilin structures, increasing the research gates in this field.

Appendix

Figures corresponding to the tandem mass spectra of some representative phyllobilins.

Fig. 1. MS^2 of $[M+H]^+$ at *m/z* = 807.3447 Da (*Zm*-NCC2 and its structural equivalents in Table 1). Some of the structural features arise from the characteristic fragmentations described in Table 2: 32 Da, methylation at $O8^4$ (at *m/z* = 775 Da $[M+H\text{-}MeOH]^+$); 123 Da, ring D presents the 18^1,18^2-vinyl arrangement (at *m/z* = 683 Da [M+H- ring D]$^+$); 155 Da, ring A is hydroxylated at $C3^2$; 285 Da, presence of ring D-β-glucopyranoyl.

Fig. 2. MS^2 of $[M+H]^+$ at *m/z* = 731.29237 Da (*Ej*-NCC2 in Table 1). Some of the structural features arise from the characteristic fragmentations described in Table 2: 88 Da, The structure presents a malonyl group and it is hydroxylated at $C3^2$ (at *m/z* = 643 Da [M+H-malonyl]$^+$); 209 Da, ring D presents the 18^1,18^2-vinyl arrangement and it is hydroxylated at $C3^2$ (at *m/z* = 522 Da [M+H- ring D-malonyl]$^+$).

Fig. 3. MS^2 of $[M+H]^+$ at *m/z* = 667.2974 Da (UCC and its structural equivalents in Table 1). Some of the structural features arise from the characteristic fragmentations described in Table 2: 18 Da, presence of hydroxyl group (at *m/z* = 649 Da $[M+H\text{-}H_2O]^+$); 32 Da, structure is methylated at $O8^4$ (at *m/z* = 649 Da $[M+H\text{-}MeOH]^+$); 157 Da, ring D presents the 18^1,18^2-dihydroxyethyl arrangement (at *m/z* = 510 Da [M+H-ring D]$^+$).

Acknowledgment

This work was supported by the Comisión Interministerial de Ciencia y Tecnología (CICYT-EU, Spanish and European Government, AGL 2015-63890-R). All the authors contributed equally to the performance and writing of this review.

References

Bale NJ, Llewellyn CA, Airs R (2010) Atmospheric pressure chemical ionisation liquid chromatography/mass spectrometry of type II chlorophyll-a transformation products: Diagnostic fragmentation patterns. Org Geochem 41:473-481.

Banala S, Moser S, Müller T et al (2010) Hypermodified chlorophyll catabolites: source of blue luminescence in senescent leaves. Angew Chem Int Ed 49:5174-5177.

Berghold J, Breuker K, Oberhuber M et al (2002) Chlorophyll breakdown in spinach: on the structure of five nonfluorescent chlorophyll catabolites. Photosynth Res 74:109-119.

Berghold J, Eichmüller C, Hörtensteiner S et al (2004) Chlorophyll breakdown in tobacco: on the structure of two nonfluorescent chlorophyll catabolites. Chem Biodivers 1:657-668.

Berghold J, Müller T, Ulrich M, et al (2006) Chlorophyll breakdown in maize: on the structure of two nonfluorescent chlorophyll catabolites. Monatsh Chem 37:751-753.

Canjura FL, Schwartz SJ (1991) Separation of chlorophyll compounds and their polar derivatives by high-performance liquid chromatography. J Agric Food Chem 39:1102-1105.

Chen K, Ríos JJ, Pérez-Gálvez, A et al (2015a) Development of an accurate and high-throughput methodology for structural comprehension of chlorophyll derivatives. (I) Phytylated derivatives. J Chromatogr A 1406:99-108.

Chen K, Ríos JJ, Roca M et al (2015b) Development of an accurate and high-throughput methodology for structural comprehension of chlorophyll derivatives. (II) Dephytylated derivatives. J Chromatogr A 1412:90-99.

Christ B, Hauenstein M, Hörtensteiner S (2016) A liquid chromatography-mass spectrometry platform for the analysis of phyllobilins, the major degradation products of chlorophyll in Arabidopsis thaliana. Plant J 88:505-518.

Christ B, Schelbert S, Aubry S, et al (2012) MES16, a member of the methylesterase protein family, specifically demethylates fluorescent chlorophyll catabolites during chlorophyll breakdown in Arabidopsis. Plant Physiol 158:628–641.

Christ B, Süssenbacher I, Moser S, et al (2013) Cytochrome P450 CYP89A9 is involved in the formation of major chlorophyll catabolites during leaf senescence in Arabidopsis. Plant Cell 25:1868-1880.

Curty C, Engel N (1996) Detection, isolation and structure elucidation of a chlorophyll a catabolite from autumnal senescent leaves of Cercidiphyllum japonicum. Phytochem 42:1531-1536.

Curty C, Engel N, Gossauer A (1995) Evidence for a monooxygenase-catalyzed primary process in the catabolism of chlorophyll. FEBS Lett 364:41-44.

Djapic N, Pavlovic M (2008) Chlorophyll catabolite from Parrotia Persica autumnal leaves. Rev Chim 59:878-882.

Djapic N, Pavlovic M (2009) Chlorophyll biodegradation products from Hamamelis Virginiana autumnal leaves. IJQR 3:1-8.

Erhart T, Mittelberger C, Vergeiner C, et al (2016) Chlorophyll catabolites in senescent leaves of the plum tree (Prunus domestica). Chem Biodiversity 13:1441-1453.

Hauenstein M, Christ B, Das A, et al (2016) A role for TIC55 as hydrolase of phyllobilins, the products of chlorophyll breakdown during plant senescence. Plant Cell 28:2510-2527.

Hinder B, Schellenberg M, Rodoni S, et al (1996) How plants dispose of chlorophyll catabolites: directly energized uptake of tetrapyrrolic breakdown products into isolated vacuoles. J Biol Chem 271:27233-27236.

Hörtensteiner S (2004) The loss of green color during chlorophyll degradation - a prerequisite to prevent cell death? Planta 219:191-194.

Hörtensteiner S, Kräutler B (2011) Chlorophyll breakdown in higher plants. Biochim Biophys Acta 1807:977-988.

Hörtensteiner S, Vicentini F, Matile P (1995) Chlorophyll breakdown in senescent cotyledons of rape, Brassica napus L.: enzymatic cleavage of phaeophorbide a in vitro. New Phytol 129:237-246.

Hörtensteiner S, Wüthrich KL, Matile P, et al (1998) The key step in chlorophyll breakdown in higher plants: cleavage of pheophorbide a macrocycle by a monooxygenase. J Biol Chem 273:15335-15339.

Jackson AH, Kenner GW, Budzikiewicz H et al (1967) Pyrroles and related compounds – X: Mass spectrometry in structural and stereochemical problems – XC Mass spectra of linear di=, tri- and tetrapyrrolic compounds. Tetrahedron 23:603-632.

Kräutler B, (2014) Phyllobilins – the abundant bilin-type tetrapyrrolic catabolites of the green plant pigment chlorophyll. Chem Soc Rev 43:6227-6238.

Kräutler B, Banala S, Moser S et al (2010) A novel blue fluorescent chlorophyll catabolite accumulates in senescent leaves of the peace lily and indicates a split path of chlorophyll breakwon. FEBS Lett 584:4215-4221.

Kräutler B, Hörtensteiner S (2014) Chlorophyll breakdown: chemistry, biochemistry and biology. In: Ferreira GC, Kadish KM, Smith K, Guilard R (eds) Handbook of Porphyrin Science - Chlorophyll, Photosynthesis and Bio-inspired Energy, vol 719. World Scientific Publishing: Singapore, pp117-185.

Kräutler B, Jaun B, Bortlik K et al (1991) On the enigma of chlorophyll degradation: the constitution of a secoporphinoid catabolite. Angew Chem Int Ed 10:1315-1318.

Losey FG, Engel N (2001) Isolation and characterization of a urobilinogenoidic chlorophyll catatolite from Hordeum vulgare L. J Biol Chem 276:8643-8647.

Matile P, Schellenberg M, Peisker C (1992) Production and release of a chlorophyll catabolite in isolated senescent chloroplasts. Planta 187:230–235.

Moser S, Müller T, Holzinger A et al (2009) Fluorescent chlorophyll catabolites in bananas light up blue halos of cell death. Proc Natl Acad Sci USA 106:15538-15543.

Moser S, Müller T, Holzinger A, et al (2012) Structures of chlorophyll catabolites in bananas (Musa acuminata) reveal a split path of chlorophyll breakdown in a ripening fruit. Chem Eur J 18:10873-10885.

Moser S, Ulrich T, Müller T et al (2008) A yellow chlorophyll catabolite is a pigment of the fall colours. Photochem Photobiol Sci 7:1577-1581.

Mühlecker W, Kräutler B (1996) Breakdown of chlorophyll: constitution of nonfluorescing chlorophyll-catabolites from senescent cotyledons of the dicot rape. Plant Physiol Biochem 34:61-75.

Mühlecker W, Kräutler B, Ginsburg S, et al (1993) Breakdown of chlorophyll: a tetrapyrrolic chlorophyll catabolite from senescent rape leaves. Helv Chim Acta 76:2976-2980.

Müller T, Moser S, Ongania KH et al (2006) A divergent path of chlorophyll breakdown in the model plant Arabidopsis thaliana. Chem Biochem 7:40–42.

Müller T, Oradu S, Ifa DR et al (2011) Direct plant tissue analysis and imprint imaging by desorption electrospray ionization mass spectrometry. Anal Chem 83:5754-5761.

Müller T, Rafelsberger M, Vergeiner C et al (2011) A dioxobilane as product of a divergent path of chlorophyll breakdown in Norway maple. Angew Chem Int Ed 50:10724-10727.

Müller T, Ulrich M, Ongania KH et al (2007) Colorless tetrapyrrolic chlorophyll catabolites found in ripening fruit are effective antioxidants Angew Chem Int Ed 46:8699-8702.

Müller T, Vergeiner S, Kräutler B (2014) Structure elucidation of chlorophyll catabolites (phyllobilins) by ESI-mass spectrometry-pseudo molecular ions and fragmentation analysis of a nonfluorescent chlorophyll catabolite (NCC). Int J Mass Spectrom 365-366:48-55.

Oberhuber M, Berghold J, Breuker K et al (2003) Breakdown of chlorophyll: A nonenzymatic reaction accounts for the formation of the colorless “nonfluorescent” chlorophyll catabolites. Proc Natl Acad Sci USA 100:6910–6915.

Oberhuber M, Berghold J, Mühlecker W et al (2001) Chlorophyll breakdown – on a nonfluorescent chlorophyll catabolite from spinach. Helv Chim Acta 84:2615-2627.

Pérez-Gálvez A, Roca M (2017) Phyllobilins: a new group of bioactive compounds. In: Rahman A (ed) Studies in Natural Products Chemistry, vol 52. Elsevier Science BV, Amsterdam, pp159-191.

Pružinska A, Anders I, Aubry S, et al (2007) In vivo participation of red chlorophyll catabolite reductase in chlorophyll breakdown. Plant Cell 19:369–387.

Pružinska A, Tanner G, Aubry S, et al (2005) Chlorophyll breakdown in senescent Arabidopsis leaves. Characterization of chlorophyll catabolites and of chlorophyll catabolic enzymes involved in the degreening reaction. Plant Physiol 139:52-63.

Ríos JJ, Pérez-Gálvez A, Roca M (2014a) Non-fluorescent chlorophyll catabolites in quince fruits. Food Res Int 65:255-262.

Ríos JJ, Roca M, Pérez-Gálvez A (2014b) Non-fluorescent chlorophyll catabolites in loquat fruits (Eriobotrya japonica Lindl.). J Agric Food Chem 62:10576-10584.

Ríos JJ, Roca M, Pérez-Gálvez A (2015) Systematic HPLC/ESI-high resolution-qTOF-MS methodology for metabolomics studies in nonfluorescent chlorophyll catabolites pathway. J Anal Methods Chem 2015:1-10.

Roca M, Ríos JJ, Chahuaris A, Pérez-Gálvez A. Non-fluorescenta and yellow chlorophyl catabolites in Japanese plum fruits (*Prunus salicina*, Lindl.). Food Res Int 100:332-338.

Roiser MH, Müller T, Kräutler B (2015) Colorless chlorophyll catabolites in senescent florets of broccoli (Brassica oleracea var. italica). J Agric Food Chem 63:1385-1392.

Scherl M, Müller T, Kräutler B (2012) Chlorophyll catabolites in senescent leaves of the lime tree (Tilia cordata). Chem Biodivers 9:2605-2617.

Scherl M, Müller T, Kreutz CR et al (2016) Chlorophyll catabolites in fall leaves of the wych elm tree present a novel glycosylation motif. Chem Eur J 22:9498-9503.

Süssenbacher I, Christ B, Hörtensteiner S, et al (2014) Hydroxymethylated phyllobilins: a puzzling new feature of the dioxobilin branch of chlorophyll breakdown. Chem Eur J 20:87-92.

Süssenbacher I, Hörtensteiner S, Kräutler B (2015) A dioxobilin-type fluorescent chlorophyll catabolite as a transient early intermediate of the dioxobilin-branch of chlorophyll breakdown in Arabidopsis thaliana. Angew Chem Int Ed 54:1-6.

Süssenbacher S, Kreutz CR, Christ B et al (2015) Hydroxymethylated dioxobilins in senescent Arabidopsis thaliana leaves: sign of a puzzling biosynthetic intermezzo of chlorophyll breakdown. Chem Eur J 21:11664-11670.

Vergeiner C, Banala S, Kräutler B (2013) Chlorophyll breakdown in senescent banana leaves: catabolism reprogrammed for biosynthesis of persistent blue fluorescent tetrapyrroles. Chem Eur J 19:12294-12305.

Figure legends

Fig. 1. Structure of products of the pheophorbide a oxygenase pathway (RCC and pFCC) and of type-I and type-II phyllobilins.

Fig. 2. Non-fluorescent chlorophyll catabolite 1 and its related yellow chlorophyll catabolite 1 isolated from senescent leaves of *Cercidiphyllum japonicum* tree, showing the dehydrogenation at the C15-16 bond as the single structural difference between both catabolites.

Fig. 3. Two examples of non-fluorescent chlorophyll catabolites structures not described so far, but fulfilling the structural constrains of type-I phyllobilins described in literature.

Fig. 4. Location of fragmentation processes described in MS/MS analysis of type-I and type-II phyllobilins and the corresponding numerical losses (Da) from the protonated ion $[M+H]^+$ that are relevant for structural identification.