Journal of Chromatography A 1717 (2024) 464656

Contents lists available at ScienceDirect

Journal of Chromatography A

journal homepage: www.elsevier.com/locate/chroma

Structural elucidation of derivatives of polyfunctional metabolites after methyl chloroformate derivatization by high-resolution mass spectrometry gas chromatography. Application to microbiota metabolites

M. Bajo-Fernández, G. Montero, V. Alonso-Herranz, C. Barbas, F. Rey-Stolle, A. García *

Centro de Metabolómica y Bioanálisis (CEMBIO), Facultad de Farmacia, Universidad San Pablo-CEU, CEU Universities, Urbanización Montepríncipe, 28660, Boadilla del Monte, Spain

A R T I C L E I N F O

Keywords:
Microbiome
High-resolution mass spectrometry
Alkyl derivatization
Gas chromatography
Chloroformates

A B S T R A C T

Metabolomics has become an essential discipline in the study of microbiome, emerging gas chromatography coupled to mass spectrometry as the most mature, robust, and reproducible analytical technique. Silylation is the most widely used chemical derivatization strategy, although it has some limitations. In this regard, alkylation by alkyl chloroformate offers some advantages, such as a rapid reaction, milder conditions, better reproducibility, and the generation of more stable derivatives. However, commercial spectral libraries do not include many of the alkyl derivatives, mainly for polyfunctional metabolites, which can form multiple derivatives. That introduces a huge bias in untargeted metabolomics leading to common errors such as duplicates, unknowns, misidentifications, wrong assignations, and incomplete results from which non-reliable findings and conclusions will be retrieved. For this reason, the purpose of this study is to overcome these shortcomings and to expand the knowledge of metabolites in general and especially those closely related to the gut microbiota through the thorough study of the reactivity of the different functional groups in real matrix derivatized by methyl chloroformate, a common representative alkylation reagent. To this end, a systematic workflow has been developed based on exhaustive structural elucidation, along with computational simulation, and taking advantage of the high sensitivity and high-resolution gas chromatography-mass spectrometry. Several empirical rules have been established according to chemically different entities (free fatty acids, amino acids, polyols, sugars, amines, and polyfunctional groups, etc.) to predict the number of derivatives formed from a single metabolite, as well as their elution order and structure. In this work, some methyl chloroformate derivatives not previously reported as well as the mechanisms to explain them are given. Extremely important is the interconversion of *E*- and *Z*- geometric isomers of unsaturated dicarboxylic acids (case of fumaric-maleic and case of citraconic-mesaconic acids), or the formation of cycled derivatives for amino acids, as well as common metabolites, as in the case of serine and cysteine, and many others.

1. Introduction

Metabolomics involves the characterization of the phenotype of an organism and its complex metabolic processes. In untargeted metabolomics, both known and unknown metabolites are extensively investigated in a global approach.

Among the analytical techniques employed in metabolomics, gas chromatography coupled to mass spectrometry (GC–MS) is not as universally applicable as liquid chromatography-mass spectrometry (LC-MS). Nonetheless, this analytical technique possesses unique features, including excellent chromatographic efficiency, resolution, and the highest retention time (RT) repeatability [1]. Concerning MS detection, high-resolution detectors (time-of-flight, orbitrap) give superior characteristics from the accurate mass *m/z* (mass-to-charge) fragments and that enhances enormously the selectivity and the sensitivity of the analysis [2].

The microbiota or microbial community is defined as the set of microorganisms that coexist in a specific habitat, usually establishing a symbiotic commensal relationship with a multicellular living organism, such as humans [3]. On the other hand, the term microbiome, originally inferred by Hegstrand and Hine in 1986 [4,5], refers to the microbiological environment as the collection of microorganisms and their

* Corresponding author.
E-mail address: antogar@ceu.es (A. García).

https://doi.org/10.1016/j.chroma.2024.464656
Received 18 August 2023; Received in revised form 6 December 2023; Accepted 11 January 2024
Available online 17 January 2024
0021-9673/© 2024 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

"theatre of activity," as defined by Whipps *et al.* in 1988 [6]. This includes their collective genomic information, surrounding environmental conditions, metabolic activities, and ecological functions.

The study of gut microbiota-derived metabolites has become increasingly significant. Some of these bioactive molecules can interact with receptors in entero-endocrine cells, while others may enter systemic circulation [7,8]. Consequently, these metabolites play a crucial role in influencing human health and disease.

Fecal samples can provide valuable insights into the state of the gut microbiota, given their proximity to the intestinal mucosa [9]. The gut microbiota is involved in the production of such important metabolites as short-chain fatty acids (SCFAs) other free fatty acids, amino acids, microbial metabolites, and tricarboxylic acid (TCA) cycle intermediates. Changes in SCFAs along the colon have been associated with altered physiological states [10], such as an increased prevalence of both benign and malignant tumors. Butyrate has been identified as a critical metabolite, suggesting an inhibition of tumor cell proliferation. Acetate and propionate also induce apoptosis, albeit to a lesser extent [11]. Furthermore, in a study performed in mice focused on colorectal cancer, fecal samples exhibited the most increased altered metabolites compared to plasma and tumor tissue [12]. Therefore, metabolic profiling of feces may offer valuable insight into the disease pathology and reveal potential diagnostic biomarkers [13].

In recent years, considerable advances have been achieved in the development of MS-based analytical methods aimed at unveiling the fecal metabolome, mainly through hyphenation to high-resolution analytical techniques such as GC, LC, and capillary electrophoresis (CE) as reviewed elsewhere [14]. Despite this considerable progress, there is no single analytical approach suitable for detecting, identifying, and quantifying all the possible metabolites that may be present.

Derivatization is required before GC analysis to reduce analyte polarity and increase thermal stability and volatility, improving the reliability of the findings. There are two main strategies in untargeted metabolomics for obtaining volatile derivatives: oxymation-silylation or alkyl chloroformate derivatization. Silylation, which is more widespread, involves replacing the active hydrogens with trialkyl silyl groups (usually trimethylsilyl or TMS), while methoxymation protects ketones from keto-enol tautomerization and decarboxylation of α-ketoacids, and inhibits the cyclization of reducing sugars. However, these coupled reactions require anhydrous media and heating, so samples must be evaporated up to total dryness, resulting in significant SCFA loss.

Alkyl derivatization was proposed by Husek in the early '90s [13,15] and followed by other groups [16,17]. Carboxylic, hydroxyl, and amino compounds could be derivatized to their alkyl esters or *O/N*-alkoxycarbonyl ethers respectively as depicted in Supplementary data Fig. S1. Additionally, these derivatives are more stable and reproducible than those obtained after oxymation-silylation.

On the other hand, alkyl chloroformate derivatization offers an instantaneous reaction without the need for heating or anhydrous media, reducing sample preparation time (one hour *versus* six hours) along with lower reagent costs, and allowing easy separation of derivatives by a robotic workstation.

Alkyl chloroformates have been used in metabolomics, specifically methyl chloroformate (MCF), for the quantitative analysis of metabolites containing amino and carboxylic groups with excellent characteristics. In this regard, Kvitvang *et al.* [18] developed a method, in which 67 metabolites, such as amino acids and organic acids, were included. Other groups have described methods employing ethyl chloroformate (ECF) [19], and even isobutyl chloroformate (BCF) [20]; Qiu *et al.* [19] developed and optimized a method based on ECF, using 25 representative compounds of different chemical classes, such as amino acids, organic acids, and amines. This method was tested in rat urine samples, selecting 14 standards among the 25 compounds tested. Afterward, the method was applied to the urine of a precancerous colon rat model, enabling the annotation of 200–250 metabolites.

Considering alkyl chloroformate derivatization, there are still some challenges to be addressed. To date, the identification of the derivatives after the chemical reaction is arduous because most of them are still not included in the spectral libraries commercially available. In addition, multiple derivatives formed from polyfunctional metabolites have not been properly elucidated yet. From the currently available literature, only the retention index (n-alkane scale) and the nominal values of their characteristic m/z fragments allow for distinguishing between major and minor peaks. However, this distinction lacks adequate elucidation [17,21]. In consequence, relevant information from key-pathway metabolites could be overlooked and underestimated, which likely may lead to misidentifications and uncontrolled duplicates. In short, there is a lack of information currently about the use of alkyl chloroformate for untargeted metabolomics. Therefore, a huge effort should be made to reliably identify and annotate the compounds in the GC–MS profile in untargeted approaches.

In a previous tutorial from our group [2], high-resolution GC–MS was highlighted due to its high capabilities for the identification of metabolites after an oxymation-silylation reaction. Similarly, we propose here a systematic workflow for metabolite profiling and the structural elucidation of the polyfunctional metabolite derivatives after the chemical reaction with MCF and the application to feces. Nevertheless, the scope extends beyond this matrix, as numerous essential metabolites in various matrices, such as proteinogenic amino acids and TCA cycle metabolites, have been investigated.

This workflow encompasses structural elucidation tools and the strategy for the identification of novel compounds based on the accurate mass determination of the *m/z* fragments, and enabling the application of high-quality algorithms, facilitating the determination of molecular ions via isotope distribution-based molecular formula, along with fragmentation rules, spectral similarity searches, and *de novo* structural elucidation.

Consequently, the main objective of this research is to increase the reliability and extend the knowledge of the microbiome by untargeted metabolomics after MCF derivatization because of the increased identification capabilities of high resolution MS coupled to GC with a thorough approach. These aims were accomplished through the reliable identification of multiple derivatives and the expansion of the commercial databases generating an accurate mass library of all the studied compounds.

2. Materials and methods

2.1. Chemicals

The derivatization reagents including MCF, pyridine, and sodium bicarbonate were purchased from Sigma-Aldrich (Sigma, Merck, Darmstadt, Germany), chloroform from Romil (Cambridge, UK), methanol of MS quality from Fisher (168 Third Avenue, 02,451, Waltham (MA)), and sodium hydroxide from VWR (100 Matsonford Rd, 19,087, Radnor (PA)). All were of analytical grade.

All standard compounds (purity >99.5 %) such as acetic, propionic, butyric, valeric, isovaleric, benzoic, α-hydroxybutyric, α-hydroxyisobutyric, α-hydroxyisovaleric, β-hydroxybutyric, lactic, glycolic, glyceric, methylmalonic, succinic, citraconic, fumaric, maleic, α-hydroxyglutaric, malic, tartaric, citric, isocitric, *cis*-aconitic, *trans*-aconitic, pyruvic, acetoacetic, levulinic, oxaloacetic, α-ketoglutaric acids, 20 α-amino acids, β-alanine and ornithine, free medium (lauric acid) and long chain fatty acids (oleic, linoleic and arachidonic acids) along with the internal standards 4-methylvaleric acid, d31-palmitic acid and tricosane, were from Sigma-Aldrich. The stock solutions of all reference standards were prepared independently in ultrapure water or methanol with a concentration of 1000 mg·L^{-1}. Working standard solutions were prepared by dilution with ultrapure water and/or methanol. Ultrapure water was prepared by the Milli-Q system (Millipore, Merck, Darmstadt, Germany).

A commercial mixture of n-alkanes (C8-C40) from Merck (40,147-U

Supelco) was analyzed every day as a reference for the Retention Index (RI).

2.2. Sample, blanks, and standard pre-treatment

2.2.1. Feces

Sample preparation was carried out following a modified method described elsewhere [17], as illustrated in Supplementary data Fig. S2. In brief, faecal samples from CEMBIO volunteers were previously lyophilized in Labogene Scanvac CoolSafe (1 g each). Then, lyophilized feces were mixed to prepare a pool of homogenized samples. 10 mg of powder were weighed in a 0.5 mL-Eppendorf tube. The samples were vortex-mixed for 2 min with 300 μL of 1 M NaOH solution in Milli-Q water and 10 μL of 1000 mg·L^{-1} 4-methylvaleric acid (the first internal standard, IS1) in water and centrifuged for 20 min at 16,000 × *g* at 4 °C. 200 μL of the supernatant were transferred into a GC vial and 200 μL of 50 mg·L^{-1} d31-palmitic acid (the second internal standard, IS2) in cold methanol was added to the residue in the 0.5 mL-Eppendorf tube to perform a second extraction, followed by centrifugation as in the previous step. Next, 167 μL of the supernatant was transferred into the same vial where the first supernatant was collected and mixed. Afterward, 20 μL of MCF and 34 μL of pyridine were added, and vials were vortex-mixed for 30 s. Another 20 μL of MCF was added, and vials were vortex-mixed again for 30 s. Successively, 400 μL of 20-mg·L^{-1} tricosane solution in chloroform (the third internal standard, IS3) was added and vials were vortex-mixed for 30 s. Next, 400 μL of 50 mM $NaHCO_3$ solution in Milli-Q Water was added to every vial, which was subsequently centrifuged at 2000 × *g* for 20 min at 4 °C. Finally, 350 μL of the bottom chloroformic phase was transferred into a GC vial without insert containing 100 mg of anhydrous sodium sulphate.

2.2.2. Standards

All the standards were prepared independently and analyzed at least twice to confirm the results. 25 μL of the solution containing 300 mg·L^{-1} of each standard were vortex-mixed for 2 min with 300 μL of 1 M NaOH aqueous solution and 10 μL of 1000 mg·L^{-1} 4-methylvaleric acid (IS1) in cold methanol. The rest of the preparation was the same as for the feces sample.

2.2.3. Spiked samples

Spiked feces samples were prepared following the same protocol as the sample, except for a step of standard addition (25 μL of the solution containing 300 mg·L^{-1} of each standard) along with 300 μL of 1 M NaOH aqueous solution and 10 μL of 1000 mg·L^{-1} 4-methylvaleric acid (IS1) in cold methanol.

2.2.4. Blanks

Blanks were prepared every day (at least 2 blanks per batch) following the same protocol as for the faecal sample, but substituting the lyophilized feces with Milli-Q water.

2.3. GC-QTOF-MS analysis

Standards, feces samples, spiked feces samples and blanks were analyzed by GC-QTOF-MS (Agilent 7890B gas chromatography coupled with a 7250 Accurate Mass Q-TOF time-of-flight mass spectrometer). 1 μL of each derivatized standard/sample/blank was injected in splitless mode into a DB5-MS column (30 m length, 0.250 mm i.d., 0.25 mm film thickness, 95 % dimethyl/5 % diphenyl polysiloxane) with a pre-column (10 m) integrated PN 122–5532 G (Agilent Technologies) with helium as the carrier gas at a constant flow rate of 1.0 mL·min^{-1}. The solvent delay time was set to 2.9 min. The optimized oven temperature program was the following: the initial temperature was set to 30 °C and held for 1 min, 10 °C/min rate up to 120 °C, then 20 °C/min rate up to 320 °C, and held for 2 min. The total time of analysis was 22 min. The temperatures in the injector, transfer line, ion source, and quadrupole were set to 270, 270, 250, and 150 °C, respectively. Electron impact ionization (70 eV) at the m/z range of 38−650 was used. The acquisition rate was 10 spectra·s^{-1}. Mass calibration was carried out after every injection. System suitability was verified by multiple injections of the same internal standard solution (tricosane, IS3). The retention time locking tool RTL was used to fix the RT of the IS3 at 18.8 min.

2.4. Data processing and software

Raw data from GC-QTOF-MS analysis was checked by MassHunter Qualitative Analysis 10.0 (Agilent Technologies). Spectral Library National Institute of Standards and Technology 2017 (NIST 17), and the in-house CEMBIO's Personal Compound Data Library (PCDL) were used for the identification. For the analysis of the standards, NIST MS Search v2.2, and its tool MS Interpreter were used. To predict the possible fragmentations of each derivative in the EI source (70 eV) the ACD/Labs tool MS Fragmenter (Ontario, Canada) was employed.

Raw data from the analysis of the pool of feces samples were exported in .D format to MassHunter Quantitative Analysis Unknowns Analysis 10.0 (Agilent Technologies), and subjected to deconvolution and spectral library search. This library search was based on both spectral similarity and RT. The database used for this identification was CEMBIO's PCDL. In case of no match, the spectral search was performed with MassHunter Qualitative Analysis 10.0 in NIST 17, after subtracting the background. Then, the identification proposed was checked according to its RI in the *n*-alkane scale retention time.

All the chromatographic information related to assigned derivatives and their spectrum was included in the in-house CEMBIO's PCDL using the program MassHunter Workstation Qualitative Analysis 10.0 and PCDL Manager B.08.00.

3. Results and discussion

3.1. Optimization of the oven program

The published method was enhanced to broaden SCFA coverage by enabling the analysis of acetic acid, previously excluded [17]. Initiating the analysis at 30 °C with a low heating rate allows the detection of this metabolite. This optimization extended the total analysis time by about 5 min.

3.2. Systematic workflow

The workflow followed (Fig. 1) aims to process the extensive information generated in the GC-QTOF-MS analysis. In the experimental phase, the feces pool, pure standards, spiked samples, and blanks were analyzed in the same sequence. Representative metabolites in feces were used as standards for the systematic workflow.

Firstly, the standards were classified according to the complexity of the functional group:

- (I) Monocarboxylic acids.
- (II) Dicarboxylic acids.
- (III) Tricarboxylic acids.
- (IV) Keto acids.
- (V) Polyols.
- (VI) Amino acids (AAs): aliphatic, acid, amide, basic, hydroxyl and sulfhydryl, and miscellaneous.

Secondly, the chromatograms of the blanks and the pure standard, sample, and spiked sample were compared, looking for the derivative peak(s). When only one peak derivative was detected, the null hypothesis established that the expected derivative is the one described in the literature [17,22,23]. For structural elucidation, the experimental spectrum and the RT/Kovats RI were first compared to the information included at CEMBIO's PCDL, and if no match was found, a second search

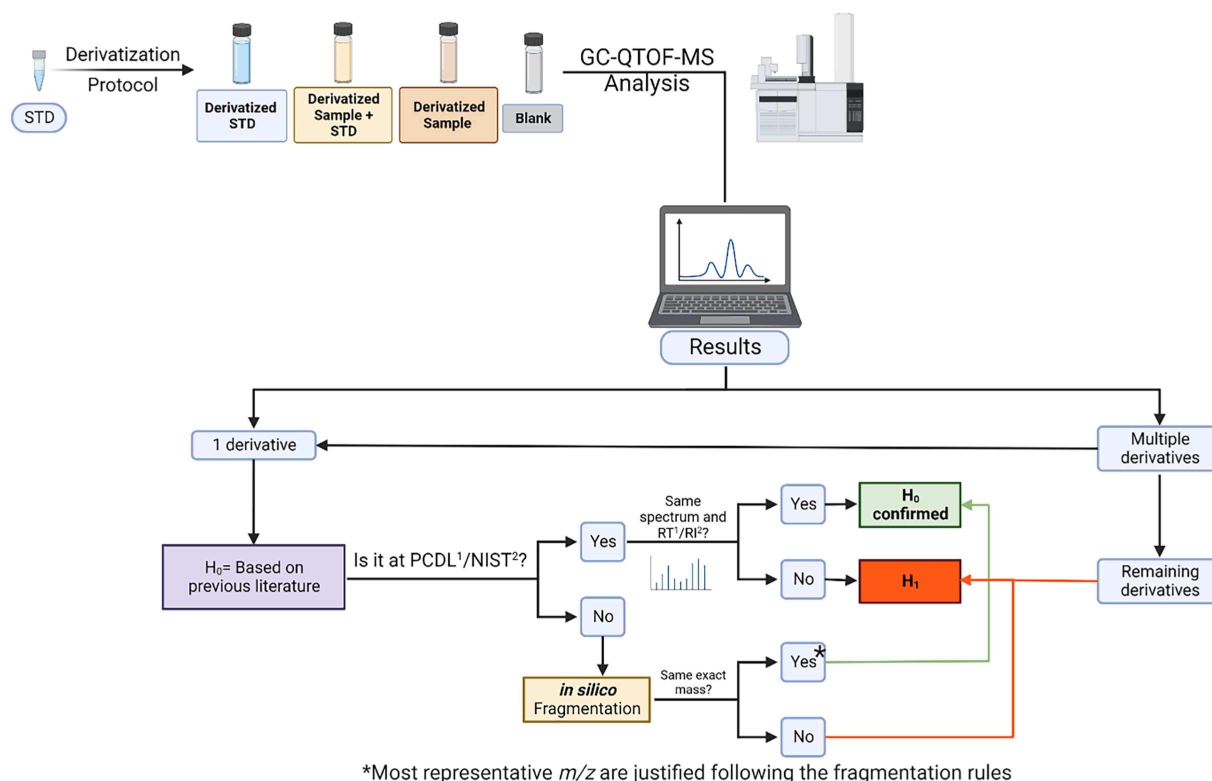

Fig. 1. Workflow established according to the presence of one or multiple derivatives. H0 or initial hypothesis when the expected derivative is described in previous literature. H1 or alternative hypothesis, which includes partial derivatizations, double derivatizations of the same functional group, and other possible reactions. The biological sample analyzed is feces.

was performed in NIST 17. If there was a match between the experimental and the theoretical information, the null hypothesis would be confirmed. On the contrary, a computational simulation of the fragmentation pattern was performed, giving *m/z* with exact mass. If the most representative fragments agreed, the null hypothesis was accepted. The most representative compounds of each group were further validated following the established GC–MS fragmentation rules [24–29].

Nonetheless, if the mass spectrum did not match either PCDL or the NIST 17 published spectrum, or the fragmentation pattern generated *in silico*, the null hypothesis was rejected, and an alternative hypothesis was formulated. This alternative hypothesis should include native functional groups, double derivatives of the same functional group, or other possible chemical reactions in the derivatizing matrix. At this point, the new hypothesis followed the same workflow as the initial hypothesis until the correct derivative was confirmed.

Whenever multiple derivatives were found for the same compound, one of them was proposed as the initial hypothesis, following the general reaction of MCF derivatization described elsewhere [13,16,17,22,23,30,31] (Supplementary data Fig S1). The rest of the derivatives should be studied as the alternative hypothesis.

The structural elucidation of the multiple derivatives of metabolites with polyfunctional groups such as hydroxy acids and amino acids is a real breakthrough for future metabolomics analysis of complex samples. To date, most of these were not previously described.

In the following sections, the most representative compounds of each group will be detailed, including the number of derivatives and the chemical structure of each one of them, indicating the main and secondary derivatives, the fragmentation spectrum, and the structure of the most representative *m/z* obtained following the fragmentation rules.

3.3. Monocarboxylic acids (I)

3.3.1. Saturated and unsaturated without any additional functional groups

The compounds analyzed include acetic, propionic, butyric, valeric, isovaleric, benzoic, and lauric acids, as well as some unsaturated fatty acids, such as oleic (18:1ω9), linoleic (18:2) and arachidonic acids (20:4ω6). Several of these compounds are SCFAs, which are crucial microbiota-derived metabolites playing significant roles in the gut-brain axis [32].

The result obtained in all the cases was a single derivative corresponding to the methyl ester of the acid (Supplementary data Fig. S1), as expected from the literature [17,22,23].

3.3.2. Hydroxy acids

The compounds analyzed were glycolic, lactic, α-hydroxybutyric, α-hydroxyisobutyric, α-hydroxyisovaleric and β-hydroxybutyric acids. Some of them are ketone bodies used as energy production systems in situations of fasting or inefficient glucose utilization, such as in type 1 diabetes [33].

The representative compound of choice was β-hydroxybutyric acid, yielding two derivatives (RIs 842 and 1146, respectively). The expected derivative, the *O*-methoxycarbonyl (hydroxyl group derivatized) and the methyl ester of the acid, coincided with the secondary derivative, eluting at a longer RT (RI 1146). The structure was confirmed through spectrum analysis and fragmentation rules (Fig. 2A). Regarding the main derivative (RI 842), it was confirmed that the structure corresponded to β-hydroxybutyric acid with only the carboxylic group derivatized (Fig. 2B).

Hence, two main types of fragmentation can be distinguished. The heterolytic cleavage or charged-inductive cleavage (*i*), which is induced by the charge generated in the radical-cation. This causes the movement of the pair of electrons of the σ C—C bond, thus splitting the molecule into a neutral radical (which is not detected in MS) and the cation of

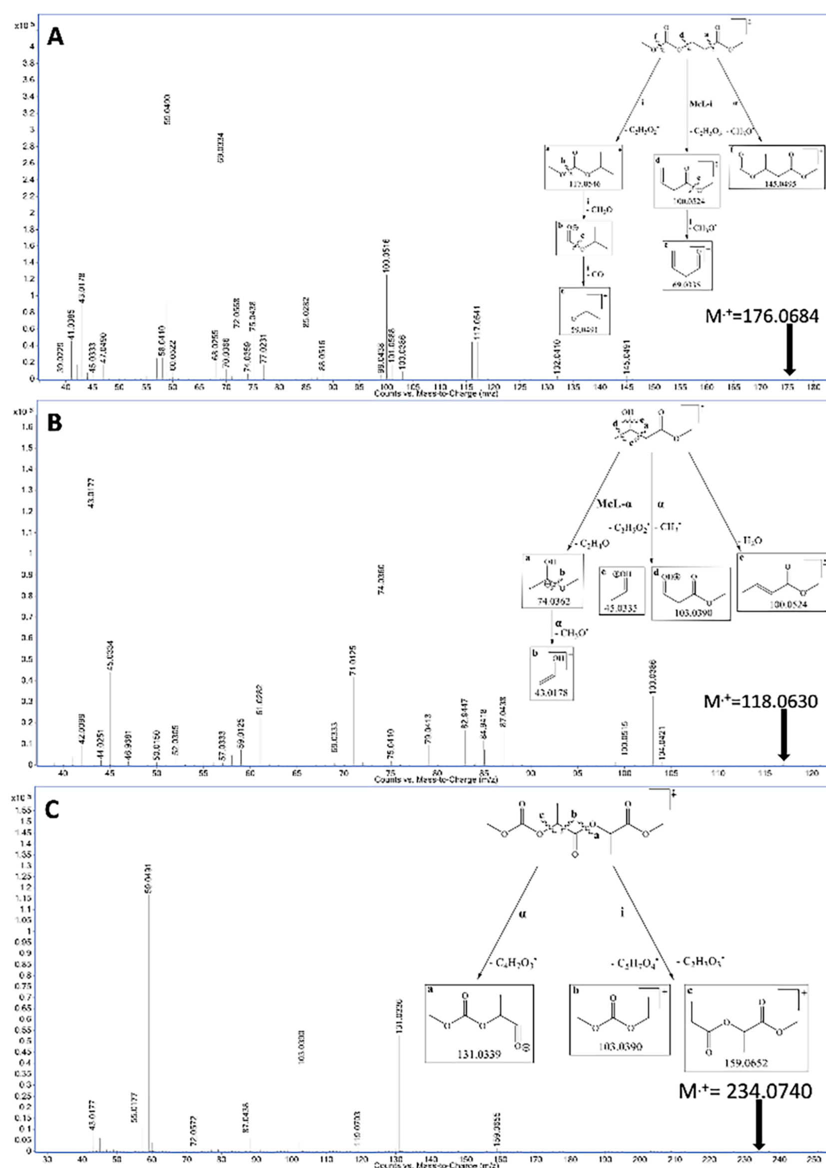

Fig. 2. Mass spectrum and proposed fragmentation of the derivatives of monocarboxylic hydroxy acids: β-hydroxybutyric acid and lactic acid. **A.** Secondary derivative of β-hydroxybutyric acid (RI 1146). **B.** Main derivative of β-hydroxybutyric acid (RI 842). **C.** Lactic acid dimer derivative. α: homolytic cleavage; i: heterolytic cleavage; McL-α: McLafferty rearrangement followed by homolytic cleavage; McL-i: McLafferty rearrangement followed by heterolytic cleavage.

mass 117.0546 (Supplementary data Fig. S3A).

The second type of fragmentation is the homolytic cleavage or radical-inductive cleavage (*α*). After the molecular ion is generated, the electron forming the radical is the one that induces the cleavage, leading to the formation of a cation of mass 145.0495 and a neutral radical corresponding to the methoxy group, which is not detected by the MS (Supplementary data Fig. S3B).

Lastly, multiple cleavages followed by rearrangements, namely McLafferty rearrangements (*McL*) [24], can also occur. These rearrangements involve the transfer of hydrogen from an element 4 atoms away from the carbonyl oxygen, generating a dystonic cation radical, meaning that the radical and that of the cation are spatially separated (Supplementary data Fig. S3C). Once the rearrangement has occurred, the molecule can be fragmented by any other of the mechanisms described above.

The presence of the hydroxyl group in the molecule allows the loss of water, resulting in the formation of an unsaturated fragment and conjugated, producing the fragment of mass 100.0515 (Fig. 2B).

In summary, β-hydroxy acids result in the formation of two derivatives. The main derivative corresponds to the methyl ester of the acid and the non-derivatized hydroxyl group, while the secondary derivative presents both functional groups derivatized, as theoretically expected [16]. In the case of α-hydroxy acids, the main product corresponds to the fully derivatized molecule, while the secondary product is partially derivatized.

Based on these results, it can be concluded that the position of the hydroxyl group to the carboxylic acid directly influences the derivatization process. Petr Husek (alkyl chloroformate derivatization pioneer) has already studied how the derivatization conditions affect the hydroxy acids, concluding that the hydroxyl groups beyond the α-position do not derivatize, and in the case of α-hydroxy acids, side products are formed due to the reactivity of the hydroxyl group [22,23,30,34]. This work qualifies this information by explaining that the hydroxyl group derivatizes but forms secondary products.

Finally, it should be noted that for C2, and C3 hydroxy acids, such as glycolic, lactic, and glyceric acid, the main derivative corresponds to the expected (methyl ester of the acid and *O*-methylcarbonyl of the hydroxyl group). Surprisingly, the first derivative (hydroxyl group not

derivatized) observed in the previous compounds was not detected. A possible explanation could be that this other derivative is more volatile and co-elutes with the solvent, not being detected in the chromatogram in this study, which was also previously observed by other authors [23]. Moreover, Villas-Bôas *et al.* reported lower MCF derivatization efficiency for lactic acid compared to silylation [16].

Besides, in the case of glycolic and lactic acids, a second derivative was detected (RI 1395 and 1415, respectively) not previously reported. As these derivatives eluted after the fully-derivatized hydroxyl group, the formation of a dimer was proposed. In this regard, the tendency to form inter-ester oligomers has been described for α-hydroxy acids, being considered as a “self-alcoholysis”, caused by the interaction of the activated carboxylic group of one molecule with the activated 2-OH of another molecule [22,23,35]. This possible structure was confirmed by studying the mass spectrum and following the fragmentation rules as depicted in Fig. 2C.

3.4. *Dicarboxylic acids (II)*

3.4.1. *Without additional functional groups*

The metabolites analyzed in this group included succinic and methylmalonic acids. These two metabolites are very important in the metabolism, as succinic acid is a key intermediate in the TCA cycle, and it is involved in the macrophages’ metabolism as well as in the synthesis of amino acids, and heme group [36]. In the case of methylmalonic acid, it is important as a biomarker indicative of vitamin B12 deficiency and methylmalonic aciduria, an autosomal recessive disease caused by a dysfunction in the catabolism of valine, leucine, isoleucine, and methionine [37]. In both instances, a single derivative was formed, which corresponds to the dimethyl ester of the acid [22,23].

3.4.2. *Unsaturated dicarboxylic acids*

The compounds analyzed in this group were fumaric and maleic acids, two *cis–trans* isomers which have different physicochemical

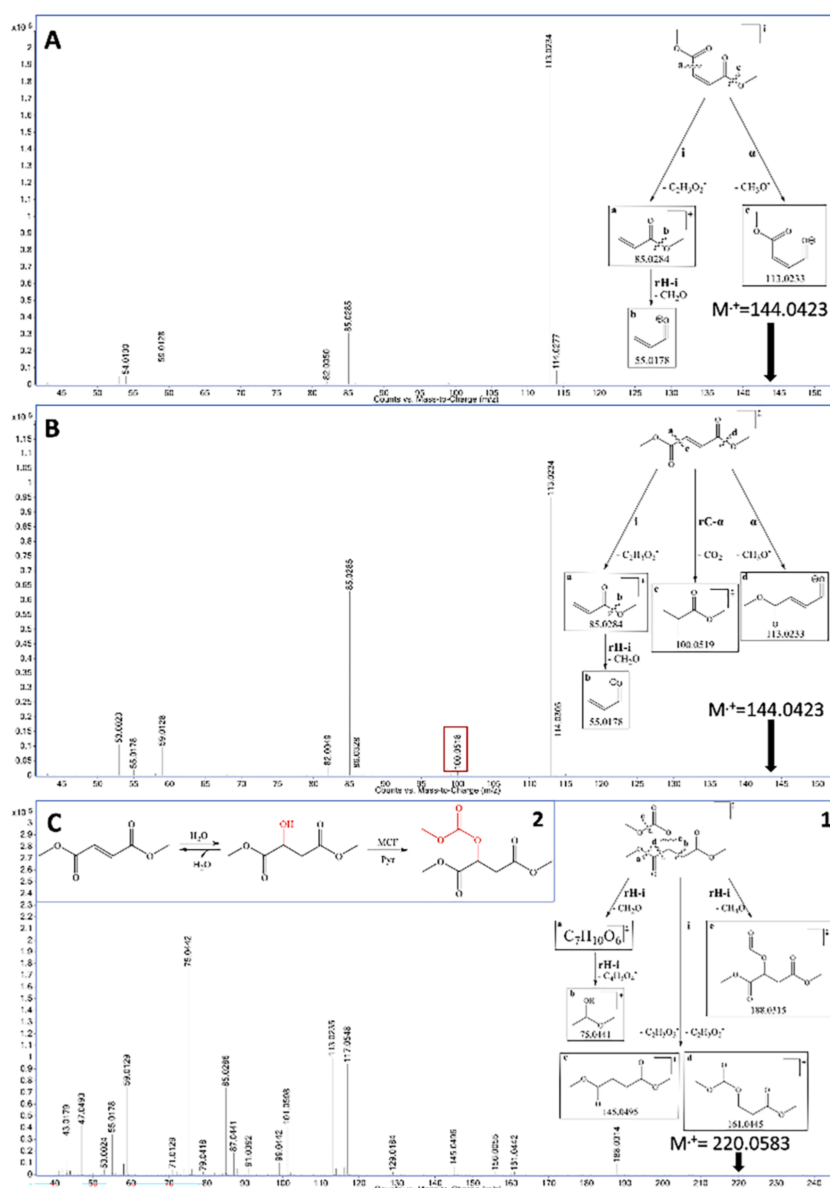

Fig. 3. Mass spectrum and proposed fragmentation for derivatives of fumaric acid. **A.** Main derivative (*cis* configuration) (RI 1013). **B.** Secondary derivative (*trans* configuration) (RI 1019). **C.** Secondary derivative (RI 1406); **C.1.** Mass spectrum and proposed fragmentation; **C.2.** Proposed mechanism to explain the presence of *O*-methoxycarbonyl derivative. α: homolytic cleavage; i: heterolytic cleavage; rH-i: hydrogen rearrangement followed by heterolytic cleavage; rC- α: carbon rearrangement followed by homolytic cleavage.

properties. This phenomenon is due to the presence of bipolar vectors in fumaric acid, which cancel each other because of their symmetry, resulting in a much more hydrophobic acid compared to maleic acid. Maleic acid presents greater polarity due to the position of the carboxylic functional groups concerning the unsaturation. Another compound analyzed in this group was citraconic acid.

The representative compound of this group is fumaric acid. This compound results in three derivatives (RIs 1013, 1019, and 1406). The first two derivatives, almost co-elute, and present very similar fragmentation spectra (Fig. 3A, 3B). The hypothesis proposed based on this observation was that the two derivatives were the two isomers, maleic acid and fumaric acid. The presence of the 100.0519 m/z was decisive,

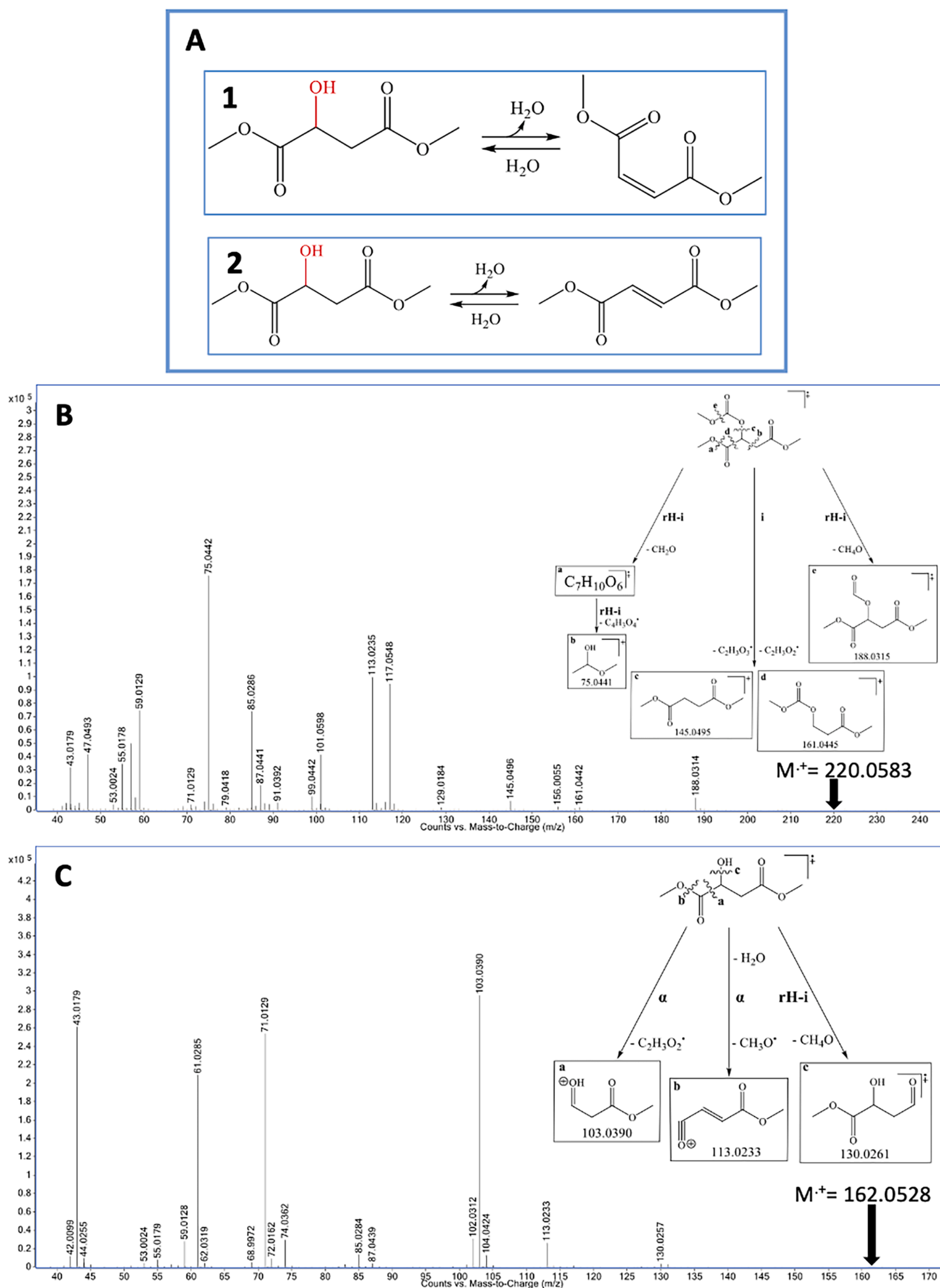

Fig. 4. Mass spectrum and the proposed fragmentation of derivatives of malic acid. **A.** Proposed mechanism to explain the presence of common derivatives with maleic/fumaric acids; **A.1.** *Cis* configuration; **A.2.** *Trans* configuration. **C.** Main derivative (RI 1406). **D.** Secondary derivative (RI 1143). α: Homolytic cleavage; i: Heterolytic cleavage; rH-i: Hydrogen rearrangement followed by heterolytic cleavage.

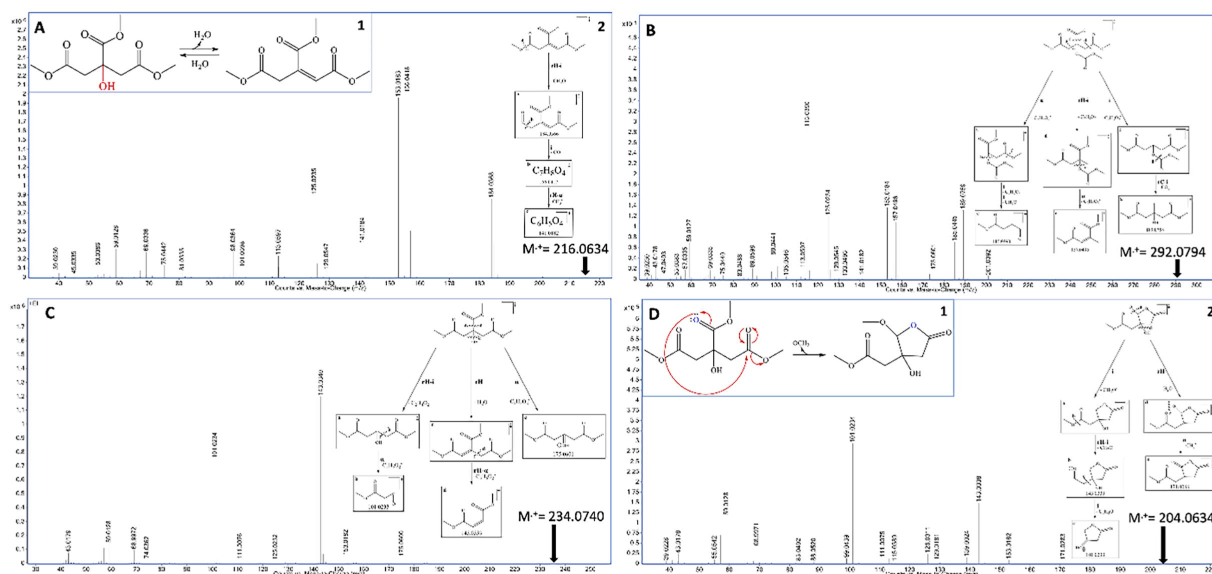

Fig. 5. Mass spectrum and the proposed fragmentation of derivatives of citric acid. **A.** Secondary derivative (RI 1464); **A.1.** Proposed mechanism to explain the presence of common derivative with *cis/trans*-aconitic acid (*cis* configuration)**; A.2.** Mass spectrum and proposed fragmentation. **B.** Secondary derivative (RI 1675). **C.** Main derivative (RI 1496). **D.** Secondary derivative (RI 1408). **D.1.** Proposed reaction of internal cyclization**; D.2.** Mass spectrum and proposed fragmentation. α: homolytic cleavage; i: heterolytic cleavage; rH: hydrogen rearrangement; rH-α: hydrogen rearrangement followed by homolytic cleavage; rH-i: hydrogen rearrangement followed by heterolytic cleavage.

since this ion is only present in the spectrum of fumaric acid (*trans* configuration), corresponding to the second derivative (RI 1019). The presence of the double derivative in the fumaric acid standard was surprising, as the purity of the compound used was >99 %, in addition to the fact that the main derivative coincides with maleic acid (Supplementary data Fig. S4A). To confirm this result, the analysis was repeated with the standard compound acquired from a different supplier, obtaining the same proportions of fumaric and maleic acids. Moreover, the maleic acid standard (>99 % purity) was analyzed, and the same result was obtained. The conclusion drawn was the partial interconversion between the two isomers, which has never been reported previously. This same phenomenon was observed when analyzing aconitic and citraconic acids.

One explanation considered was the transformation of fumaric acid to maleic acid through an anhydride intermediate. However, this is energetically disfavored due to carboxylate relative positions in fumaric acid (Supplementary data Fig. S5A) [38], and the lack of heat during this protocol.

We hypothesize that chloride generated in the media enhances this transformation. It is known that chloride induces the generation of a carbanion, eliminating the unsaturation of the acid temporarily, and indistinctly producing both geometric isomers [39]. The sources of the chloride in the applied protocol are the derivatization reagent, MCF, as each functional group when reacting with MCF released a chloride equivalent to the reaction medium, and chloroform, which decomposes to phosgene ($COCl_2$) and HCl in the presence of oxygen or sunlight. At the same time, HCl dissociates into its chloride anions caused by the basic medium necessary for the derivative reaction to take place due to its strong acid nature. Hence, it is proposed that, once the compound is derivatized, chloride acts as a catalyst in the transformation of the two derivatives, as it is shown in Supplementary data Fig. S5B. This hypothesis was reinforced by the formation of a double α, β-unsaturated carbonyl, which tends to react through conjugated nucleophilic addition by the presence of an electron-attracting group (oxygen), facilitating the movement of the unsaturation electrons. This transformation does not occur when chloride is absent in the media. It was confirmed when both standard solutions were analyzed by an orthogonal technique such as capillary electrophoresis coupled to high-resolution mass spectrometry (CE-TOF-MS) with negative ESI ionization mode (data not shown).

The third and minor derivative was the most retained (RI 1406). After comparing the mass spectrum with the one from PCDL and NIST and proposing the structure based on the fragmentation rules (Fig. 3C.1), it was concluded that the derivative comes from the formation of 2-hydroxybutanedioic acid (malic acid). This derivative corresponds to the dimethyl ester of the butanedioic acid, which also contained the group *O*-methoxycarbonyl in position 2. Furthermore, the RI and spectrum of this derivative matched the main malic acid derivative. This third derivative was absent in maleic acid, being only observed when analyzing fumaric acid.

We propose that the formation of this common derivative with malic acid as shown in Fig. 3C.2 is based on the specific properties of α, β-unsaturated carbonyls that facilitate hydration. Subsequently, the hydroxyl group is derivatized by MCF to the *O*-methylcarbonyl derivative.

3.4.3. Hydroxy dicarboxylic acids

The compounds analyzed in this group include α-hydroxyglutaric, malic, and tartaric acids. α-Hydroxyglutaric acid is an intermediate in the TCA cycle and is related to organic aciduria. Additionally, the enantiomer D- α-hydroxyglutaric acid is considered an oncometabolite, and a biomarker of tumors such as glioma [40]. Moreover, malic acid is an essential intermediate in energy metabolism and an effective agent for inactivating common food pathogens due to its impact on pH. Tartaric acid, found in grapes and bananas, serves as an acidity regulator, antioxidant, flavor enhancer, and sequestrant in the food sector [41].

The representative compound of this group was malic acid, yielding four derivatives (RIs:1013, 1019, 1143, and 1406). The first two and the last derivatives coincided in RIs and mass spectrum with 2-butenedioic acid products, which were described previously (Fig. 3) (Section 3.4.2). Thus, the first and second peaks correspond to maleic acid and fumaric acid derivatives. The explanation is the same as for the presence of maleic acid in the latter: as 2-butenedoic acid could be hydrated, the dehydration of malic acid derivative could also occur (Fig. 4A.1, A.2). The main peak was the fully derivatized compound as reported in the literature (Fig. 4B).

The derivative not described above (RI 1143), was searched against

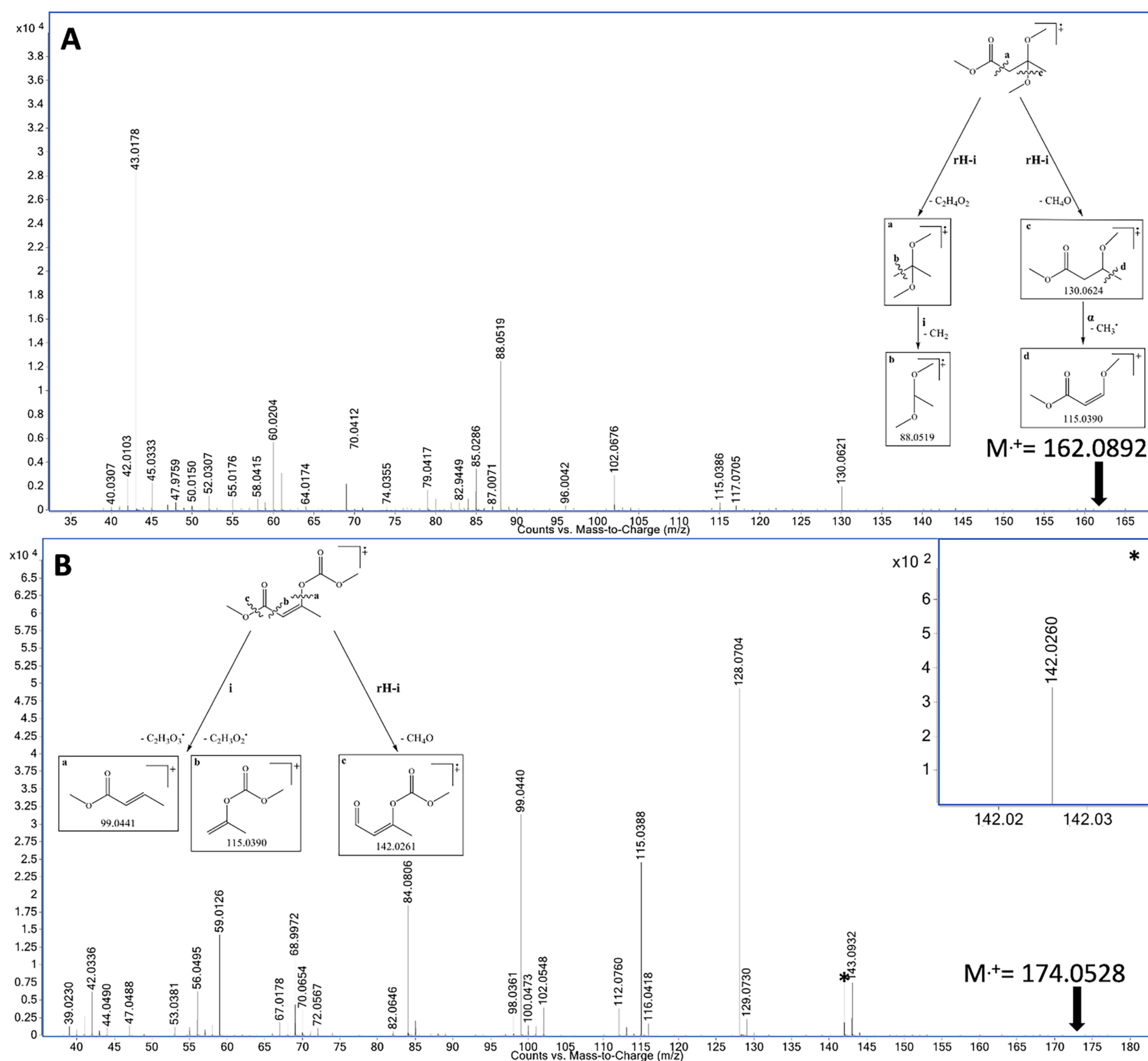

Fig. 6. Mass spectrum and the proposed fragmentation of the secondary derivatives of acetoacetic acid. **A.** Secondary derivative (RI 929). **B.** Secondary derivative (RI 1184). i: heterolytic cleavage; α: homolytic cleavage; rH-i: hydrogen rearrangement followed by heterolytic cleavage.

the PCDL and NIST spectrum library and it was found to be the dimethyl ester of malic acid, with the native hydroxyl group. This phenomenon was previously reported [42] and, eventually, can be justified following the fragmentation rules (Fig. 4C).

When analyzing tartaric acid (dihydroxy dicarboxylic acid) three peaks were detected (RIs: 1427, 1466, and 1683), the latter being the main derivative that corresponds to the fully derivatized metabolite (Supplementary data Fig. S6A). Furthermore, the two remaining derivatives were assessed from the findings previously obtained: one of the two products corresponded to the partially derivatized (RI 1466), and the less retained (RI 1427) to the dehydrated derivative (Supplementary data Fig. S6B, S6C).

3.5. Tricarboxylic acids (III)

3.5.1. Unsaturated

The molecules analyzed in this group include *cis*- and *trans*- aconitic acids. Although both participate in the metabolism, *cis*-aconitic acid is of great importance due to its role in the TCA cycle, as it is necessary for the isomerization between citric acid and isocitric acid. Alterations in the aconitase, the enzyme responsible for its synthesis, have been related to several types of cancers, such as gastric and prostate cancer [43].

In the case of these acids, both chromatograms showed two peaks (RIs: 1463 and 1477), being the second one less intense. From the previous results obtained with maleic acid and fumaric acid, the hypothesis proposed was that both peaks corresponded to the trimethyl esters of the geometric isomers, which was further confirmed (Supplementary data Fig. S7). Since both have the same spectrum, to determine which derivative (*cis* or *trans*) eluted first, Kovaḱs index was the key, leading to the conclusion that the first observed peak belonged to the *cis* isomer, which was the main derivative in both cases (Supplementary data Fig. S4B).

3.5.2. Hydroxy tricarboxylic acids

In this group, the compounds included were citric and isocitric acids. Both acids are involved in important metabolic pathways, such as the TCA cycle, where citrate synthase is responsible for the condensation of oxaloacetate with acetyl-CoA to form citrate, being converted to *cis*-aconitate by the enzyme aconitase. This intermediate is further converted to isocitrate through a reaction of dehydration followed by rehydration. Therefore, these two metabolites are essential for the energy supply [44].

The analysis of both compounds resulted in four derivatives, one of the products (RI 1464, second to elute) being shared. This derivative

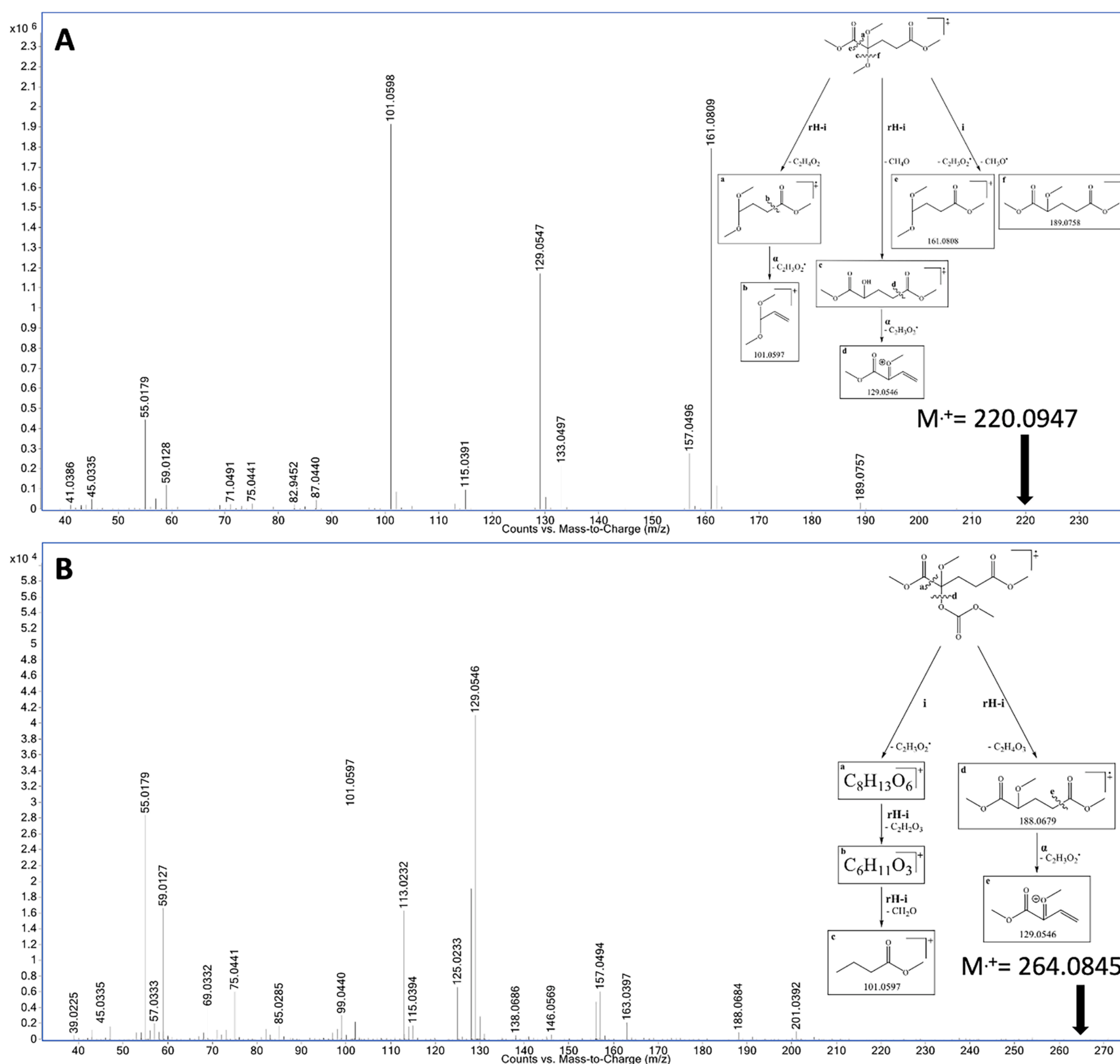

Fig. 7. Mass spectrum and the proposed fragmentation of the secondary derivatives of α-ketoglutaric acid. **A.** Secondary derivative (RI 1424). **B.** Secondary derivative (RI 1556).i: heterolytic cleavage; α: homolytic cleavage; rH-i: hydrogen rearrangement followed by heterolytic cleavage.

corresponded to the *cis*-aconitic acid product, which is explained by the water loss of citric/isocitric acids (Fig. 5A.1). Husek *et al.* described this derivative in a previous study for citric acid, although no further description of the mechanism was detailed [30].

The representative compound chosen in this group was citric acid. The last compound to elute (RI 1675) corresponded to the expected derivative from the literature: the fourfold derivative, with all the functional groups derivatized (trimethyl ester and *O*-methoxycarbonyl) (Fig. 5B). Furthermore, the results obtained in the group of dicarboxylic hydroxy acids were applied, and the main derivative (RI 1496) was identified as the partially-derivatized compound (Fig. 5C).

Lastly, for the first derivative, RI 1408, an intramolecular cyclization was proposed giving the lactone, an inner ester (Fig. 5D.1). The *in silico* fragmentation of the cycled form matches with the experimental mass spectrum, which was further justified with the fragmentation rules (Fig. 5D.2). The intra-cyclization of tricarboxylic acids has never been reported before as a possible product of alkyl derivatization.

3.6. Ketoacids (IV)

This group was described separately from the monocarboxylic and dicarboxylic acids (Sections 3.3 and 3.4) due to the distinct reactivity with MCF.

3.6.1. Keto monocarboxylic acids

In this group, the compounds analyzed include pyruvic, acetoacetic, and levulinic (4-oxopentanoic) acids. Pyruvic and acetoacetic are ketone bodies, and therefore can be used to diagnose diabetes mellitus. Pyruvic acid is a key intermediate in several metabolic pathways, gluconeogenesis, TCA cycle, and glycolysis, among others, and acetoacetic acid in the synthesis of cholesterol, fatty acids, and complex lipids [45]. Levulinic acid is reported as a secondary metabolite that can act as a defense or signaling molecule.

Within this group, acetoacetic acid was chosen to be the representative compound, although controversial comments in the literature were found. In previous studies, the degradation of acetoacetic acid in the system was described [23]. However, when analyzing acetoacetic acid here, three derivatives were observed (RIs: 854, 929, and 1184).

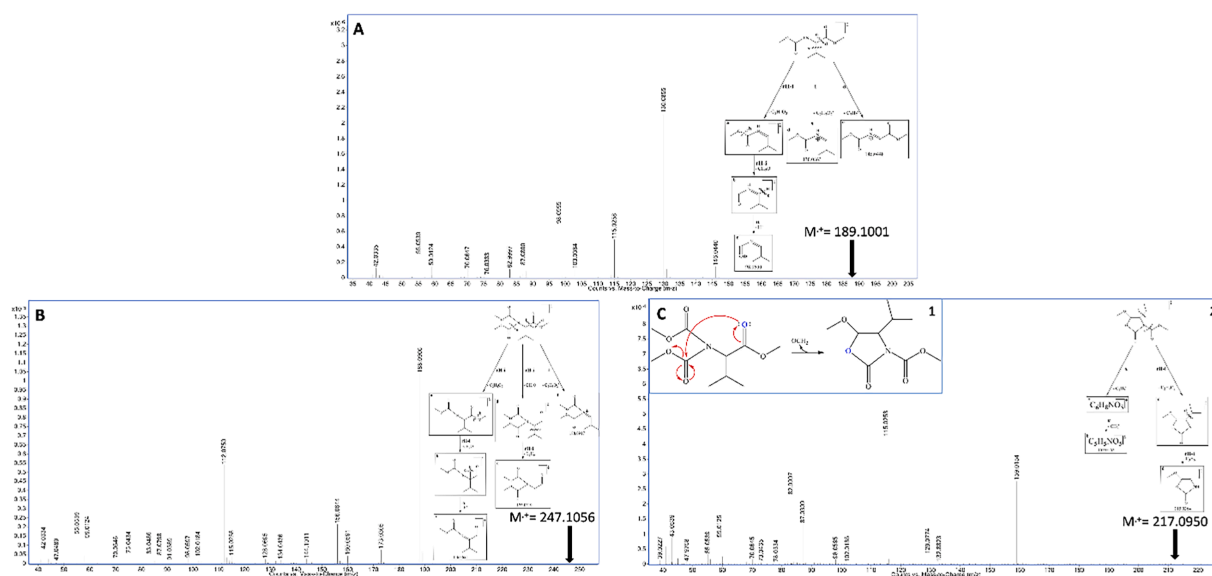

Fig. 8. Mass spectrum and the proposed fragmentation of the derivatives of valine. **A.** Main derivative (RI 1291). **B.** Secondary derivative (RI 1480). **C.** Secondary derivative (RI 1405); C**1.** Proposed reaction of intramolecular cyclization**; C2.** Mass spectrum and proposed fragmentation. i: heterolytic cleavage; α: homolytic cleavage; rH-i: hydrogen rearrangement followed by heterolytic cleavage.

The main derivative (RI 852) corresponded to the expected product: acetoacetic acid, methyl ester. This is also the result obtained for levulinic acid, which only formed one derivative (RI 974).

Regarding the second derivative, a possible acetalization was proposed, which was further confirmed (Fig. 6A). The formation of this acetal derivative has never been reported previously.

For the remaining derivative (RI 1184) the possibility of the derivative coming from the tautomeric enol form of acetoacetic acid was proposed, although there were no reports about this product in keto acids. The structure was justified and further confirmed through the possible cleavages that the compound could undergo (Fig. 6B).

In the case of pyruvic acid, the acetal derivative was also formed and constitutes the main and the first product (RI 928). However, additional derivatives (RIs 1069, 1102) were detected. The formation of multiple side-products was previously described by Husek *et al.*, although no information about the structure was reported [30]. The same author suggested a possible formation of oligomers by the interaction of the activated carboxyl with the enolic hydroxyl [23]. Thus, an aldol condensation was proposed to explain the presence of multiple derivatives. Nevertheless, the third derivative (RI 1102) corresponded to the citraconic acid product. Therefore, besides aldol condensation, the resulting molecule must have undergone a decarbonylation (Supplementary data Fig. S8A), resulting in the same derivative as citraconic acid.

Surprisingly, the mass spectrum and RT/RI of the second derivative (RI when compared to those available at PCDL) coincide with those of the methyl succinic acid derivative. After searching the literature and studying possible reactions that may have happened, hydration of the third derivative was proposed (RI 1069).

3.6.2. *Keto dicarboxylic acid*

In this group, the compounds analyzed were oxaloacetic and α-ketoglutaric acids. Oxaloacetic acid is an intermediate metabolite that participates in different metabolic pathways, such as gluconeogenesis, urea cycle, and TCA cycle, in which α-ketoglutaric acid plays a very important role, determining the overall rate of this process. Furthermore, α-ketoglutaric acid constitutes a source of glutamate and glutamine that stimulates protein synthesis, inhibits protein degradation in muscles, and participates in the differentiation of TH1 to Treg (regulatory T-cell) [46]. The ingestion of supplements of this compound can increase circulation plasma levels of hormones, such as insulin, growth hormone, and insulin-like growth factor-1 [47].

In the case of oxaloacetic acid, only one derivative was observed, which coincided with the first derivative of pyruvic acid (RI 928). This observation led to the conclusion that oxaloacetic acid underwent a decarboxylation giving pyruvic acid, typical of β-ketoacids, followed by the acetalization reaction.

Regarding α-ketoglutaric acid, five derivatives were detected. The lack of oxymation at the beginning of the protocol enables the transformation of the keto to the enol form. Based on this premise, the presence of the two tautomeric forms was proposed, as the second and third derivatives, both eluting very close (RIs: 1261 and 1295). The fragmentation spectrum was very similar in terms of m/z and relative intensities. The main derivative corresponds to the enol form (third derivative), which elutes after the keto form (second derivative). In a previous report [30], the formation of two tautomeric forms when analyzing α-ketoglutaric acids was observed, although no further information was given.

Based on the results obtained in the keto monocarboxylic acids, the decarbonylation of the keto form was proposed to explain the first derivative (RI 1029), which resulted in the same structure as the derivative of succinic acid (Supplementary data Fig. S8B).

The fourth product (RI 1424) was identified as the acetyl derivative, as occurred in the case of acetoacetic acid and pyruvic acid (Fig. 7A). Furthermore, the possibility of the formation of a hemiacetal and the derivatization of the hydroxyl group was suggested for the las derivative to elute (RI 1556). This structure was confirmed by *in silico* fragmentation and further justified by the fragmentation rules (Fig. 7B).

All these findings for carboxylic acids are summarized in Supplementary data Table S1.

3.7. *Polyols and monosacharides (V)*

Within this group, metabolites with several hydroxyl groups, such as arabitol (C5) and myo-inositol (C6), were analyzed as representatives. Some of them are biomarkers of microorganism overgrowth, such as arabitol [48], or participate in multiple cell signaling pathways, as is the case of myo-inositol.

The same result was obtained for polyols: no discriminant peak was found when comparing the standard to the blanks. This is a major

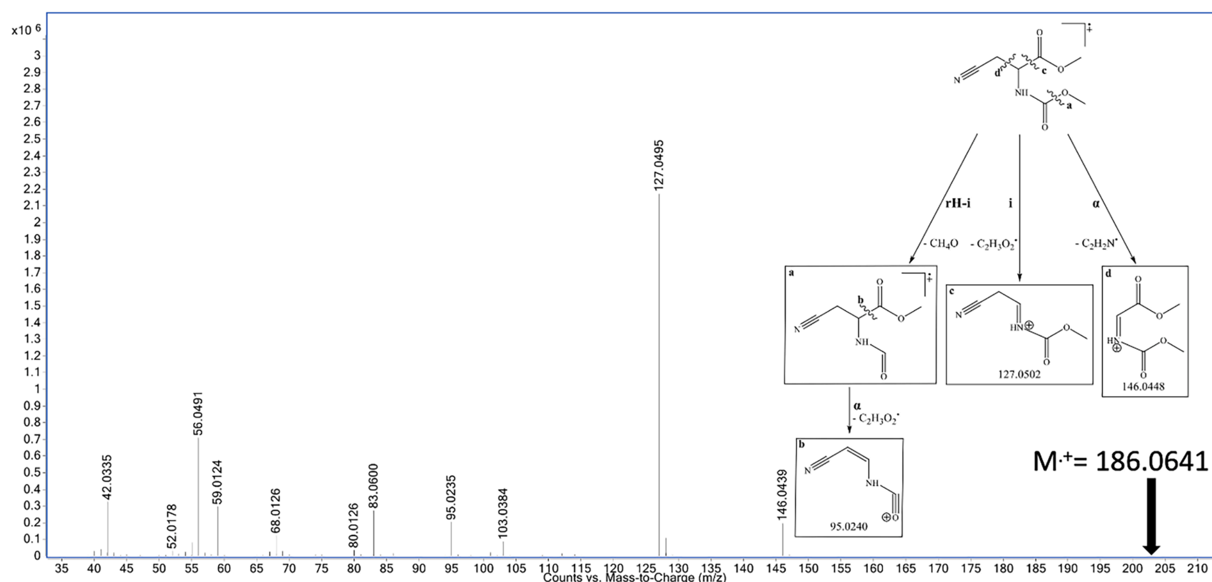

Fig. 9. Mass spectrum and proposed fragmentation of the main derivative of asparagine (RI 1422). α: homolytic cleavage; i: heteroltic cleavage; rH-i: hydrogen rearrangement followed by heterolytic cleavage.

limitation for untargeted metabolomic studies applying MCF derivatization. One possible explanation consists of a partial derivatization of hydroxyl groups, as seen in previous results (Sections 3.3.2 and 3.4.3). In liquid-liquid extraction (LLE), two phases are formed, a chloroformic and an aqueous phase, and only the former is analyzed. Therefore, it could be possible that the partially derivatized compound with multiple native hydroxyls was not extracted in the organic phase.

To corroborate this hypothesis, derivatization by methoxymation followed by silylation (method described elsewhere [2]) was performed using the aqueous phase separated after the MCF alkylation. In the case of myo-inositol, the silylated derivative was obtained, proving that the compound was not extracted in the chloroformic phase and likely not derivatized by MCF derivatization (data not shown).

For glucose, several derivatives were obtained (data not shown). It suggests that the hydroxyl groups were likely partially derivatized using the MCF method. However, this partial derivatization was insufficient to either transfer the compound to the chloroformic phase or ensure adequate thermal stability, leading to degradation in the gas chromatograph injector port. This limitation in analyzing polyols was previously reported by other authors [16,49,50], where the same hypothesis was proposed but there were no further experiments to confirm it.

Hence, it is concluded that alkyl chloroformate derivatization is not recommended for either the study of carbohydrates or polyols.

3.8. Amino acids (VI)

The study of amino acids is crucial for understanding host-microbiota interactions. Intestinal bacteria are known to affect free amino acid concentration in the gastrointestinal tract, thereby impacting the bioavailability of these compounds. Moreover, amino acids play a vital role in the synthesis of the organic acids mentioned earlier in this study, establishing a link between microbiota metabolism and host amino acids and organic acids' homeostatic state [51]. Since amino acids are widely analyzed by GC–MS, it is especially important to study in detail the derivatives obtained in the MCF derivatization process.

Amino acids were classified as indicated in Section 3.2, into four groups based on the functional groups of the side chain: aliphatic, acid, amide, basic, hydroxyl, sulfhydryl, and miscellany (including those that present special structural characteristics).

3.8.1. Aliphatic amino acids

The compounds analyzed in this group include valine, leucine, and isoleucine. Although these molecules have only two functional groups that could be derivatized, three products were observed for each of the compounds.

Among the three amino acids, valine was chosen to be the representative compound. The main derivative was observed at RI 1291, corresponding to the one described in the literature, the methyl ester of the acid and the *N*-methoxycarbonyl (amine group derivatized) (Fig. 8A).

Since the nitrogen of the derivatized amine has an acidic hydrogen in a highly basic medium, a double derivatization of the amine group was proposed for one of the remaining side products, as previously described elsewhere [52]. Although the molecular ion was absent in the spectrum, the structure was confirmed simulating the fragmentation *in silico* and justifying the most representative *m/z* (Fig. 8B). This derivative was the most retained (RI 1480).

Hence, for the derivative eluting at RI 1405, the hypothesis proposed consisted of the intramolecular cyclization, once the amine group was double derivatized, releasing a methoxy group and giving a cycled carbamate (Fig. 8.C.1). This proposal was confirmed by studying the mass spectrum and applying the fragmentation rules (Fig. 8.C.2). This reaction has not been previously described in the literature.

Regarding leucine and isoleucine, it should be noted that the main derivatives eluted at RI 1373 and 1386, respectively, and they were perfectly separated and identified (Supplementary data Fig. S4C), showing the high efficiency of the chromatographic method. The separation of these two isomers was described also in other studies [17,21]. Moreover, leucine did not present the cycled derivative.

3.8.2. Acid amino acids

In this group, glutamic acid and aspartic acid were analyzed. As reported in the literature, only one derivative could be expected, the compound with the two carboxylic groups transformed into methyl esters, and the amine group into *N*-methoxycarbonyl. This was the case for aspartic acid, giving a single derivative which eluted at RI 1490 (Supplementary data Fig. S9A).

In the case of glutamic acid, the derivative expected from the literature was observed at RI 1608. However, two secondary peaks were also found. Regarding the second derivative to elute (RI 1586), an intramolecular cyclization was proposed. Husek *et al.* previously described

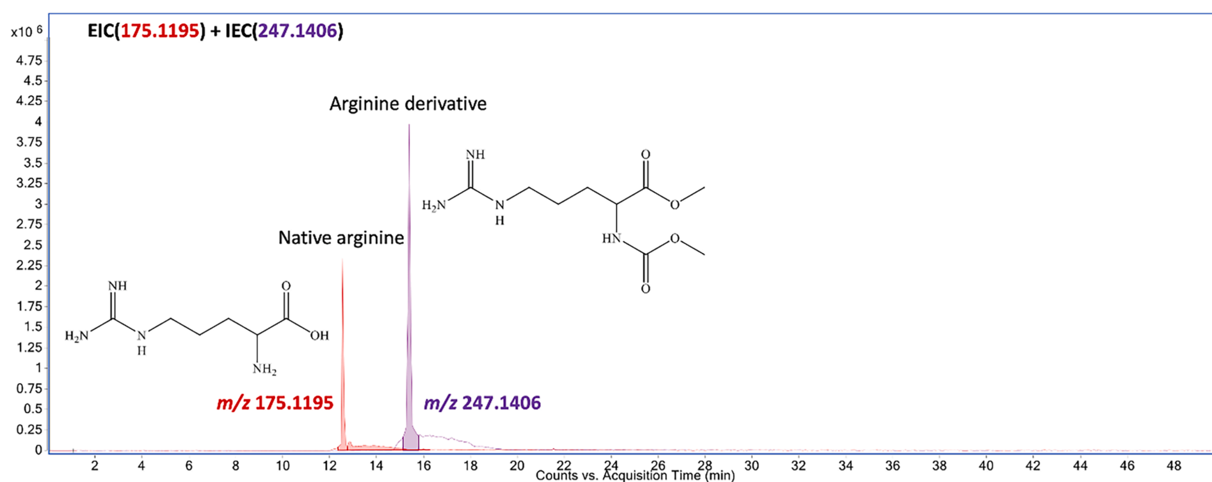

Fig. 10. Electropherogram of the aqueous phase spiked with arginine. EIC for *m/z* 175.1195 (arginine) and for *m/z* 247.1406 (arginine derivatized).

Fig. 11. Mass spectrum and proposed fragmentation of the derivatives of cysteine. **A.** Main derivative (RI 1710). **B.** Secondary derivative shared with serine (RI 1162). α: homolytic cleavage; β: homolytic cleavage in β position; rH-i: hydrogen rearrangement followed by heterolytic cleavage.

the cyclization of the main derivative, although they concluded that basifying the medium suppressed the secondary derivative, contrary to what we observed [13,23]. As with the threefold derivatized valine, the glutamic acid underwent an intramolecular cyclization (Supplementary data Fig. S9B).

Moreover, the third peak corresponded to pyroglutamic acid, derivatized in the carboxylic group only (RI 1418). The formation of this derivative when analyzing glutamic acid has been also observed in previous studies [22].

3.8.3. Amide amino acids

In this group glutamine and asparagine were included. In both cases, the expected derivative was not observed and, instead, dehydrations seemed to occur, as reported previously [22,23,52,53]. This resulted in the conversion of the amide group into the cyano group. Even though this reaction was described, in the databases consulted it was not considered, leading to the misidentification of the metabolite. These derivatives eluted at RIs 1422 and 1560, for asparagine and glutamine, respectively, corresponding to the main derivative in the case of asparagine and the secondary derivative for glutamine (Supplementary data Fig. S10; Fig. 9).

For the secondary derivative of asparagine and the main derivative of glutamine, the possibility of the derivatization of the amide group was discarded. The typical stability of the functional group is based on electron resonance between nitrogen and oxygen, which significantly reduces the nucleophilicity of the group, thus preventing the attack on the MCF. Nevertheless, spontaneous hydrolysis of the amide group is probably giving the corresponding acid, which usually happens in the human intestine, as previously described [54]. This deamination occurs via a cyclical intermediate, which can result not only in the aspartic acid and glutamic acid, but also in the isoglutamic and isoaspartic acids (Supplementary data Fig. S11). Hence, the second derivative shows the same RI and spectrum as the corresponding aspartic and glutamic acid derivatives. This finding should be considered in metabolomic analysis based on this method, as it could lead to an over-quantification of aspartic and glutamic acid in the detriment of asparagine and glutamine.

Only for glutamine, the previously described secondary derivatives for glutamic acid were observed (Section 3.8.2): pyroglutamic acid and cycled glutamic acid derivatives.

3.8.4. Basic amino acids

This group of amino acids includes lysine, arginine, histidine, and ornithine. The analysis of lysine and histidine yielded a single product, where both amine groups and the carboxylic group were derivatized. The derivatization of the imidazole moiety in histidine was described in previous studies [13,30]. In the case of ornithine, besides the expected derivative, another product was detected. Husek *et al.* described a possible intramolecular cyclization of this amino acid, as occurred with glutamic acid (Section 3.8.2) [23]. This cycled formed was confirmed by the fragmentation rules (Supplementary data Fig. S12).

3.8.4.1. Arginine. A case study. Interestingly, the chromatogram of arginine did not show any additional peak compared to the blanks. The impossibility of detecting arginine applying MCF derivatization was previously reported in different studies [16,22,52,55], although the specific mechanism was not detailed. Nevertheless, Kaspar *et al.* and Husek *et al.* [30,56], among other groups, suggested that the arginine derivative could not be analyzed due to the thermal instability of the compound, or due to the adsorption of the derivative on the column through its guanidine group [13,57,58].

The most plausible hypothesis is related to the isoelectric point of the guanidine group, with pKa_3:12.48, thus with a positive charge at pH below and close to this value. This would not only prevent the side chain from being derivatized, but also it would give higher polarity to arginine. Consequently, arginine remains in the aqueous phase rather than being extracted into the organic phase, as observed with polyols. Due to this behavior, arginine would go undetected when analyzing the chloroformic phase.

To confirm this hypothesis, the aqueous phase was collected (instead of the chloroformic phase) and analyzed along with a standard solution of arginine (50 mg·L^{-1}) by CE-TOF-MS (method described elsewhere [59]). Our results validated the hypothesis because the derivative of arginine (methyl ester and *N*- methoxycarbonyl) with the free guanidine group was confirmed by its accurate mass, whereas native arginine was not detected (Fig. 10). Hence, this study unveils facts against the myths explaining that the absence of arginine detection is not attributable to thermal instability but rather to its high basicity, rendering the compound highly polar and unsuitable for extraction by LLE in the chloroformic phase, as observed with polyols.

3.8.5. Hydroxyl and sulfhydryl amino acids

In this group, serine and cysteine were analyzed, both chemically related structures, as well as other hydroxyl and sulfhydryl amino acids, such as threonine and methionine.

In the analysis, both serine and cysteine gave two derivatives. The main derivative eluted at the largest RT, at 1555 (serine) and 1710 (cysteine). This example is shown in Fig. 11A.

It should be noted that the thiol group was derivatized as the hydroxyl group due to its similar chemical properties [13,60,61]. Although serine presents a β-hydroxy acid structure, it did not follow the same rules, as the main derivatives kept the hydroxyl group non-derivatized, and contrary to what was previously reported in other studies [13,30,34,52,57]. Thus, it is concluded that the presence of another functional group in the α position, either carboxylic acid or amine, enhances the derivatization of hydroxyl groups.

Interestingly, the secondary derivative for both amino acids has the same RI (1162) and mass spectrum. Since these derivatives eluted before the main product, the possible loss of the only differential functional group was proposed. While serine underwent dehydration, cysteine underwent desulfurization. The study of mass spectrum using the fragmentation rules justified this structure (Fig. 11B). Moreover, the molecular ion of this derivative was present in the spectrum, further validating our hypothesis. This common derivative of serine and cysteine has never been reported in the literature.

Concerning threonine, two derivatives were generated, representing the partially derivatized product (RI 1390) as the primary derivative, and the fully-derivatized product (RI 1552) as the secondary derivative. This differs from serine, where the fully derivatized amino acid predominated, presenting a more coherent pattern than previously reported for hydroxyl amino acids [13,30,34,52,57].

In the case of methionine, two peaks were also observed (RI 1625 and 1799). The first peak corresponded to the expected product, where the amine and the carboxylic group were derivatized; and the second product presented a double derivatization in the amine group, as seen before with other amino acids (Section 3.8.1).

All these findings for amino acids are summarized in Supplementary data Table S2.

3.8.6. Miscellaneous

In this last group, the remaining amino acids showed the expected chromatographic profile based on the literature, with one derivative, except for glycine.

Thus, aromatic amino acids such as phenylalanine, tryptophan, and tyrosine give a single derivative with all the functional groups derivatized (RIs 1738, 2375, 2178, respectively). In the case of tryptophan, the amine located in the indole group was not derivatized, unlike the amine group in the imidazole group of histidine. This result has been obtained in similar analyses performed by other researchers [13,57,62]. The most plausible hypothesis for this phenomenon is based on the pair of electrons in the aromatic ring of tryptophan, which results in the indole nitrogen not acting as a nucleophilic base. In contrast, histidine's

Table 1
Results obtained in this study. Main fragments in accurate mass (intensity order) and experimental Kovats retention index (RI) of 55 compounds derivatized with methyl chloroformate and analyzed by GC-QTOF-MS. *: Proposed derivative for annotation, target or semi-target analysis in untargeted metabolomics. **: compounds that do not present any selective derivative. (m): main peak; (s): secondary peak.

Compound	Derivatives	RI	Main fragments (*m/z*)
Acetic acid	Acetic acid, methyl ester	662	43.0178, 74.0362, 59.0128, 42.0100, 44.0213
Acetoacetic acid	*(m) Acetoacetic acid, methyl ester	854	43.0179, 74.0364, 59.0128, 88.0521, 42.0100
	(s) Acetoacetic acid, acetal, methyl ester	929	43.0179, 88.0519, 60.0206, 82.9450, 61.0281
	(s) Acetoacetic acid, *O*-methoxycarbonyl, methyl ester	1184	99.0427, 115.0376, 59.0126, 68.9972, 102.0536
Alanine	Alanine, *N*-methoxycarbonyl-, methyl ester	1146	102.0550, 58.0651, 70.0288, 59.0127, 42.0338
Arachidonic acid	Arachidonic acid, methyl ester	2256	79.0541, 91.0541, 80.0621, 67.0541, 93.0698
Asparagine	*(m) Asparagine, cyano, *N*-methoxycarbonyl, methyl ester	1422	127.0496, 56.0492, 42.0335, 59.0124, 83.0600
	Aspartic acid, *N*-methoxycarbonyl, dimethyl ester ((s) Asparagine)	1490	160.0597, 128.0336, 118.0494, 59.0128, 86.0233
Aspartic acid**	Aspartic acid, *N*-methoxycarbonyl, dimethyl ester ((s) Asparagine)	1490	160.0597, 128.0336, 118.0494, 59.0128, 86.0233
Benzoic acid	Benzoic acid, methyl ester	1113	105.0322, 77.0383, 136.0519, 51.0226, 106.0345
Butyric acid	Butyric acid, methyl ester	739	74.0362, 71.0491, 43.0542, 87.0441, 59.0128
Cis*-Aconitic acid*	(m) *Cis*-Aconitic acid, trimethyl ester ((s) Citric/Isocitric acid)	1463	156.0418, 153.0182, 125.0235, 184.0366, 157.0482
	(s) *Trans*-Aconitic acid, trimethyl ester	1477	153.0182, 156.0418, 125.0235, 184.0368, 157.0482
Citraconic acid**	Citraconic acid, dimethyl ester ((s) Pyruvic acid)	1101	127.0376, 99.0427, 59.0127, 69.0322, 126.0312
	Mesaconic acid, dimethyl ester ((s) Citraconic acid)	1126	126.0312, 98.0362, 127.0376, 99.0427, 69.0322
Citric acid	(s) Citric acid, trimethyl ester, cycled	1408	101.0233, 143.0336, 59.0127, 57.0335, 68.9970
	(m) *Cis*-Aconitic acid, trimethyl ester ((s) Citric/Isocitric acid)	1464	156.0418, 153.0182, 125.0235, 184.0366, 157.0482
	*(m) Citric acid, trimethyl ester	1496	143.0340, 101.0234, 59.0127, 175.0601, 57.0335
	(s) Citric acid, *O*-methoxycarbonyl, trimethyl ester	1675	115.0376, 153.0169, 125.0260, 157.0482, 189.0744
Cysteine	(s) Cysteine / Serine, *N*-methoxycarbonyl, methyl ester (loss of hydroxyl/thiol group)	1162	127.0262,55.0417, 159.0522, 99.0313, 56.0496
	*(m) Cysteine *N,S*-bis(methoxycarbonyl), methyl ester	1710	192.0234, 117.0004, 132.0114, 176.0137, 146.0341
Fumaric acid**	(m) Maleic acid, dimethyl ester (*cis*) ((m) Fumaric acid / (s) Malic acid)	1013	113.0235, 85.0285, 59.0128, 114.0312, 54.0100
	(s) Fumaric acid, dimethyl ester (*trans*) ((s) Maleic acid / (s) Malic acid)	1019	113.0235, 85.0285, 114.0312, 59.0128, 53.0022
	(m) Malic acid, *O*-methoxycarbonyl, methyl ester ((s) Fumaric acid)	1406	59.0128, 75.0442, 117.0548, 57.0335, 113.0235
Glutamic acid**	Pyroglutamic acid, methyl ester ((s) Glutamic acid/Glutamine)	1418	84.0444, 56.0495, 41.0386, 85.0473, 143.0577
	(s) Glutamic acid, *N*-methoxycarbonyl, dimethyl ester, cycled ((s) Glutamine)	1586	142.0499, 98.0601, 70.0651, 41.0386, 42.0339
	(m) Glutamic acid, *N*-methoxycarbonyl, dimethyl ester ((s) Glutamine)	1608	114.0551, 142.0500, 174.0762, 82.0289, 59.0128
Glutamine	Pyroglutamic acid, methyl ester ((s) Glutamic acid/Glutamine)	1418	84.0444, 56.0495, 41.0386, 85.0473, 143.0577
	*(s) Glutamine, cyano, *N*-methoxycarbonyl, methyl ester	1560	141.0656, 109.0396, 56.0495, 82.0525, 55.0417
	(s) Glutamic acid, *N*-methoxycarbonyl, dimethyl ester, cycled ((s) Glutamine)	1586	142.0499, 98.0601, 70.0651, 41.0386, 42.0339
	(m) Glutamic acid, *N*-methoxycarbonyl, dimethyl ester ((m) Glutamine)	1608	114.0551, 142.0500, 174.0762, 82.0289, 59.0128
Glyceric acid	(s) Glyceric acid, *O*-methoxycarbonyl, methyl ester	1264	43.0179, 59.0127, 87.0076, 42.0100, 45.0335
	*(m) Glyceric acid, *O2,O3*-bis(methoxycarbonyl), methyl ester	1477	59.0128, 91.0390, 133.0495, 103.0389, 59.0478
Glycine	*(m) Glycine, *N*-methoxycarbonyl, methyl ester	1150	88.0394, 44.0496, 56.0131, 59.0129, 115.0264
	(s) Glycine, *N2,N2*-bis(methoxycarbonyl), methyl ester	1385	146.0447, 102.0551, 72.0445, 59.0129, 58.0651
Glycolic acid	*Glycolic acid, *O*-methoxycarbonyl, methyl ester	991	59.0128, 45.0335, 74.0363, 117.0184, 73.0271

(*continued on next page*)

Table 1 (*continued*)

Compound	Derivatives	RI	Main fragments (*m/z*)
	Glycolic acid dimer, *O*-methoxycarbonyl, methyl ester	1395	45.0334, 117.0169, 73.0271, 91.0543, 59.0128
Histidine	(m) Histidine, *N2,N4*-bis(methoxycarbonyl), methyl ester	2076	210.0637, 81.0448, 140.0581,194.0562, 226.0825
Isocitric acid	(s) Isocitric acid, trimethyl ester, cycled	1455	115.0390, 153.0180, 125.0233, 157.0495, 189.0758
	(m) *Cis*-Aconitic acid, trimethyl ester ((s) Citric/Isocitric acid)	1464	156.0418, 153.0182, 125.0235, 184.0366, 157.0482
	*(m) Isocitric acid, trimethyl ester	1532	115.0390, 83.0128, 143.0339, 55.0178, 59.0127
	(s) Isocitric acid, *O*-methoxycarbonyl, trimethyl ester	1723	129.0548, 101.0598, 75.0442, 157.0496, 201.0395
Isoleucine	*(m) Leucine, *N*-methoxycarbonyl, methyl ester	1373	88.0394, 144.1020, 102.0551, 59.0128, 115.0265
	(m) Isoleucine, *N*-methoxycarbonyl, methyl ester	1386	144.1018, 88.0393, 115.0265, 69.0699, 59.0128
	(s) Isoleucine, *N2,N2*-bis(methoxycarbonyl), methyl ester, cycled	1495	115.0263, 159.0158, 83.000, 87.0313, 41.0386
	(s) Isoleucine, *N2,N2*-bis(methoxycarbonyl), methyl ester	1561	202.1075, 134.0449, 173.0319, 102.0185, 69.0699
Isovaleric acid	Isovaleric acid, methyl ester	777	74.0362, 57.0695, 43.01750, 85.0635, 101.0584
Lactic acid	*Lactic acid, *O*-methoxycarbonyl, methyl ester	1026	59.0489, 59.0128, 103.0386, 43.0178, 55.0179
	Lactoyllactic acid, *O*-methoxycarbonyl, methyl ester	1415	59.0491, 131.0325, 103.0376, 59.0127, 55.0177
Lauric acid	Lauric acid, methyl ester	1555	74.0363, 87.0441, 71.0491, 55.0542, 69.0699
Leucine	*(m) Leucine, *N*-methoxycarbonyl, methyl ester	1373	88.0394, 144.1020, 102.0551, 59.0128, 115.0265
	(m) Isoleucine, *N*-methoxycarbonyl, methyl ester	1386	144.1018, 88.0393, 115.0265, 69.0699, 59.0128
	(s) Leucine, *N2,N2*-bis(methoxycarbonyl), methyl ester	1552	202.1071, 134.0446, 57.0698, 102.0548, 69.0698
Levulinic acid	(m) Levulinic acid, methyl ester	974	115.0389, 43.0178, 57.0335, 71.0491, 72.0570
Linoleic acid	Linoleic acid, methyl ester	2095	67.0542, 81.0699, 95.0855, 79.0542, 55.0542
Lysine	Lysine, *N2,N6*-bis(methoxycarbonyl), methyl ester	2012	142.0862, 88.0393, 153.0657, 82.0650, 212.0789
Maleic acid**	(m) Maleic acid, dimethyl ester (*cis*) ((m) Fumaric acid / (s) Malic acid)	1013	113.0235, 85.0285, 59.0128, 114.0312, 54.0100
	(s) Fumaric acid, dimethyl ester (*trans*) ((s) Maleic acid / (s) Malic acid)	1019	113.0235, 85.0285, 114.0312, 59.0128, 53.0022
Malic acid	(m) Maleic acid, dimethyl ester (*cis*) ((m) Fumaric acid / (s) Malic acid)	1013	113.0235, 85.0285, 59.0128, 114.0312, 54.0100
	(s) Fumaric acid, dimethyl ester (*trans*) ((s) Maleic acid / (s) Malic acid)	1019	113.0235, 85.0285, 114.0312, 59.0128, 53.0022
	*(s) Malic acid, dimethyl ester	1143	103.0390, 43.0179, 71.0129, 61.0285, 102.0312
	(m) Malic acid, *O*-methoxycarbonyl, methyl ester ((s) Fumaric acid)	1406	59.0128, 75.0442, 117.0548, 57.0335, 113.0235
Methionine	*(m) Methionine, *N*-methoxycarbonyl, methyl ester	1625	115.0258, 61.0103, 147.0519, 128.0336, 114.0536
	(s) Methionine, *N2,N2*-bis(methoxycarbonyl), methyl ester	1799	173.0307, 205.0568, 61.0102, 114.0541, 59.0124
Methylmalonic acid	Methylmalonic acid, dimethyl ester	948	59.0125, 115.0384, 87.0437, 72.0570, 114.0312
Oleic acid	Oleic acid, methyl ester	2100	55.0542, 69.0699, 74.0362, 67.0542, 87.0427
Ornithine	(s) Ornithine, *N2,N6*-bis(methoxycarbonyl), methyl ester, cycled	1657	128.0705, 139.0501 96.0445, 42.0336, 69.0336
	*(m) Ornithine, *N2,N6*-bis(methoxycarbonyl), methyl ester	1913	128.0707, 115.0259, 198.0635, 88.0394, 230.0911
Oxaloacetic acid**	(m) Pyruvic acid, acetal, methyl ester (Oxaloacetic acid)	930	89.0598, 43.0179, 57.0336, 117.0547, 47.0492
Phenylalanine	Phenylalanine, *N*-(methoxycarbonyl), methyl ester	1738	162.0675, 91.0543, 146.0587, 161.0602, 131.0478
Proline	Proline, *N*-(methoxycarbonyl)-, methyl ester	1419	128.0707, 82.0652, 42.0338, 59.0127, 84.0808
Propionic acid	Propionic acid, methyl ester	698	57.0335, 88.0519, 55.0178, 56.0257, 58.0369
Pyruvic acid**	(m) Pyruvic acid, acetal, methyl ester (Oxaloacetic acid)	928	89.0598, 43.0179, 57.0336, 117.0547, 47.0492
	2-Mehylsuccinic acid, dimethyl ester ((s) Pyruvic acid)	1069	59.0490, 100.0519, 129.0533, 59.0128, 101.0584

(*continued on next page*)

Table 1 (*continued*)

Compound	Derivatives	RI	Main fragments (*m/z*)
	Citraconic acid, dimethyl ester ((s) Pyruvic acid)	1102	127.0391, 99.0442, 59.0128, 69.0335, 126.0312
Serine	(s) Cysteine / Serine, *N*-methoxycarbonyl, methyl ester (loss of hydroxyl/thiol group)	1162	127.0262,55.0417, 159.0522, 99.0313, 56.0496
	*(m) Serine, *N,O*-bis(methoxycarbonyl), methyl ester	1555	100.0380, 59.0128, 42.0338, 146.0434, 127.0264
Succinic acid**	Succinic acid, dimethyl ester ((s) α-ketoglutaric acid)	1029	115.0385, 55.0177, 59.0127, 87.0438, 114.0309
Tartaric acid	(s) Tartaric acid, *O*-methoxycarbonyl, dimethyl ester (loss of hydroxyl group)	1427	115.0391, 68.9971, 59.0127, 143.0340, 47.0491
	(s) Tartaric acid, *O*-methoxycarbonyl, dimethyl ester	1466	101.0234, 59.0128, 68.9971, 85.0283, 45.0335
	*(m) Tartaric acid, *O2,O3*-bis(methoxycarbonyl), dimethyl ester	1683	59.0127, 85.0284, 115.0390, 191.0549, 105.0533
Threonine	*(m) Threonine, *N*-methoxycarbonyl, methyl ester	1390	115.0265, 100.0394, 83.0003, 56.0494, 147.0527
	(s) Threonine, *N,O*-bis(methoxycarbonyl), methyl ester	1552	114.0550, 146.0447, 115.0263, 59.0490, 59.0127
Trans*-Aconitic acid*	(m) *Cis*-Aconitic acid, trimethyl ester ((s) Citric/Isocitric acid)	1463	156.0418, 153.0182, 125.0235, 184.0366, 157.0482
	(s) *Trans*-Aconitic acid, trimethyl ester	1477	153.0182, 156.0418, 125.0235, 184.0368, 157.0482
Tryptophan	Tryptophan, N-(methoxycarbonyl)-, methyl ester	2375	130.0651, 77.0385, 103.0541, 57.0698, 276.1105
Tyrosine	Tyrosine, *N,O*-bis(methoxycarbonyl)-, methyl ester	2178	121.0648, 236.0680, 165.0546, 59.0128, 77.0386
Valeric acid	Valeric acid, methyl ester	814	74.0362, 43.0176, 87.0427, 57.0695, 85.0635
Valine	*(m) Valine, *N*-methoxycarbonyl, methyl ester	1291	130.0857, 98.0596, 115.0259, 55.0540, 59.0128
	(s) Valine, *N2,N2*-bis(methoxycarbonyl), methyl ester, cycled	1405	115.0258, 159.0154, 82.9998, 87.0309, 43.0539
	(s) Valine, *N2,N2*-bis(methoxycarbonyl), methyl ester	1480	188.0908, 112.0751, 156.0645, 55.0538, 59.0124
α-Hydroxybutyric acid	(s) α-Hydroxybutyric acid, methyl ester	819	59.0491, 41.0386, 57.0334, 61.0284, 58.0413
	*(m) α-Hydroxybutyric acid, *O*-methoxycarbonyl, methyl ester	1125	73.0647, 45.0335, 59.0127, 117.0545, 41.0386
α-Hydroxyglutaric acid	*(m) α-Hydroxyglutaric acid, dimethyl ester	1266	85.0285, 57.0335, 86.0284, 59.0125, 55.0165
	(m) α-Hydroxyglutaric acid, *O*-methoxycarbonyl, dimethyl ester	1514	71.0489, 131.0696, 99.0435, 175.0591, 115.0383
α-Hydroxyisobutyric acid	(s) α-Hydroxyisobutyric acid, methyl ester	765	59.0492, 43.0179, 41.0386, 60.0526, 73.0650
	*(m) α-Hydroxyisobutyric acid, *O*-methoxycarbonyl, methyl ester	1039	73.0649, 117.0548, 43.0178, 59.0129, 43.0543
α-Hydroxyisovaleric acid	(s) α-Hydroxyisovaleric acid, methyl ester	874	73.0647, 90.0310, 55.0541, 43.0541, 57.0334
	*(m) α-Hydroxyisovaleric acid, *O*-methoxycarbonyl, methyl ester	1178	55.0542, 87.0804, 131.0703, 59.0127, 114.0676
α-Ketoglutaric acid	Succinic acid, dimethyl ester ((s) α-Ketoglutaric acid)	1029	115.0385, 55.0177, 59.0127, 87.0438, 114.0309
	(s) α-Ketoglutaric acid, dimethyl ester (keto)	1261	115.0385, 55.0177, 87.0438, 59.0127, 88.0380
	*(m) α-Ketoglutaric acid, dimethyl ester (enol)	1295	115.0385, 55.0177, 87.0438, 59.0128, 116.0395
	(s) α-Ketoglutaric acid, bis(*O*-methoxy)	1424	101.0598, 161.0809, 129.0547, 55.0179, 157.0482
	(s) α-Ketoglutaric acid, *O*-methoxy, *O*-methoxycarbonyl, dimethyl ester	1556	129.0533, 101.0584, 55.0179, 128.0468, 59.0127
β-Alanine	β-Alanine, *N*-methoxycarbonyl, methyl ester	1257	101.0472, 88.0394, 44.0495, 102.0536, 70.0288
β-Hydroxybutyric acid	*(m) β-Hydroxybutyric acid, methyl ester	842	43.0177, 74.0360, 45.0334, 71.0126, 103.0386
	(s) β-Hydroxybutyric acid, *O*-methoxycarbonyl, methyl ester	1146	59.0490, 69.0334, 100.0516, 59.0127, 43.0178

imidazole group involves one nitrogen participating in the aromaticity, placing its electrons in a pure -*p* orbital. The other nitrogen possesses a free pair of electrons that are not involved in the conjugation phenomenon. This pair of electrons grants the nitrogen nucleophilicity and basicity required for the reaction.

In the case of proline, a single derivative with pyrrolidine derivatized was detected (RI 1419), confirming that secondary amines also react, although some authors, such as Smart *et al.* [21], claim that they have lower reactivity.

Surprisingly, analysis of glycine showed two derivatives, being the main product the two-fold derivative (RI 1150), as expected from the literature (amine and carboxylic group derivatized), and the secondary

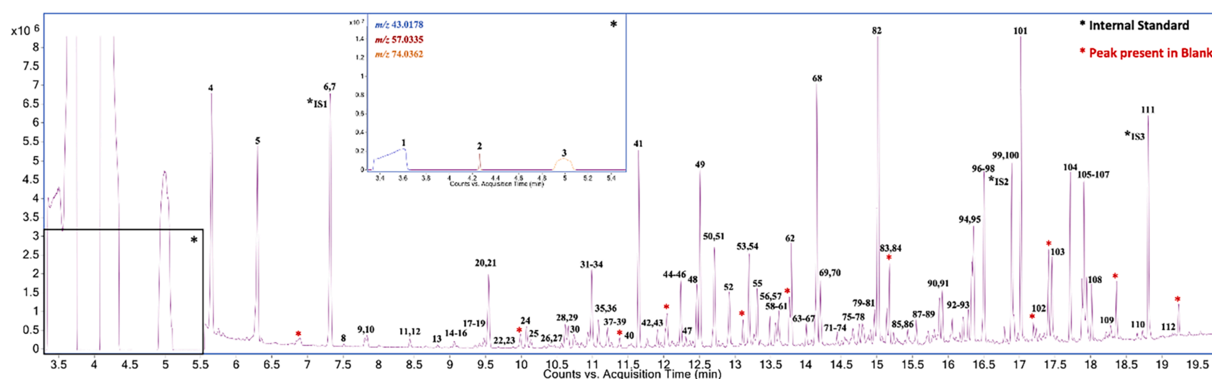

Fig. 12. Total ion chromatographic profile of a pool of feces samples from healthy subjects. Each peak was identified using CEMBIO's PCDL, the corresponding compounds are detailed in Table 2.

peak (RI 1385) the double-derivatized amine and carboxylic group, as it has been reported by other amino acids (Sections 3.8.1 and 3.8.5).

Lastly, both alanine and β-alanine were studied to determine whether the effect of the distance from the carboxylic acid could influence the derivatization of the amine group, as was the case of hydroxy acids. Nevertheless, there were no secondary derivatives, only the expected one was found (RIs 1146 and 1257, respectively), which led to the conclusion that the amine group is more reactive than the hydroxyl group.

4. Author's recommendations for reliable identification and quantification

Substantial efforts have been devoted not only to structurally elucidate each derivative, but also to understand the mechanisms underlying each reaction. This approach has allowed us to establish derivatization rules tailored to the functional groups of each compound and their relative positioning within the whole molecule. To ensure unambiguous identification and facilitate the target/semi-target analysis, it is imperative to select the specific derivative for each metabolite, as proposed in Table 1.

However, for those metabolites lacking a specific derivative, the following strategies are proposed:

- In situations where one or multiple derivatives are formed, and all are shared with another metabolite, the proposed approach involves subtracting the relative area of the peak of the derivative attributed to the other metabolite. For example, aspartic acid produces a single derivative, aspartic acid *N*-methoxycarbonyl dimethyl ester, which is shared with asparagine. Conversely, asparagine yields two derivatives, one (main derivative) that is specific to this metabolite and another (secondary derivative) that is shared with aspartic acid. Therefore, the total peak area of the aspartic acid derivative could be due to the contributions of aspartic acid plus the second derivative of asparagine. For quantification purposes, we propose the use of relative response factors (RRF) to calculate the peak area coming only from aspartic acid as explained below:
- RRF (Peak Area of the secondary derivative / Peak Area of the main derivative) = 0.909. That was calculated with the standard solution containing only asparagine.
- Peak Area of shared derivative in the sample (asparagine contribution) = (RRF) x Peak Area of the main derivative.
- Neat Peak Area in the sample (only aspartic acid) = total Peak Area – (asparagine contribution).

This strategy is also applicable to glutamic acid and glutamine (relative response factor = 0.250), as well as to succinic acid and α-ketoglutaric acid (relative response factor = 0.111). Furthermore, an alternative approach consists of the addition of stable isotopically labeled standards mainly of those that give multiple derivatives: asparagine, glutamine, and α-ketoglutaric acid.

- In instances involving isomers, where there is a nearly equal proportion of *cis* and *trans* configurations, the proposed strategy is to combine the relative areas of both peaks and report them as the mixture of both. Examples include fumaric/maleic acids, *cis-/trans*-aconitic acids, and citraconic/mesaconic acids.
- For pyruvic and oxaloacetic the first strategy does not apply. Here, oxaloacetic acid forms a single derivative shared with pyruvic acid. Despite pyruvic acid producing multiple derivatives, none of them are exclusive to this metabolite. Consequently, in this situation, it is recommended to report both metabolites together.

To sum up, our study covers various circumstances, suggesting practical strategies, while acknowledging the need for further investigation in specific cases. The outlined strategy in Table 1 helps the choice of specific derivatives for quantitation in complex scenarios in any kind of biological sample. Moreover, the consistency in the ratio among multiple derivatives from the same metabolite enhances the reliability of targeted analysis.

5. Application to human feces in the study of microbiota

Finally, once all the derivatives proposed in this study were elucidated, the in-house personal compound data library, PCDL, was updated accordingly, and it was applied to the characterization of the pool of feces samples from healthy individuals. As a result, 112 peaks were identified based on their RT and spectra, as shown in Fig. 12 and Tables 1 and 2, many of which could have been assigned as unknowns without our findings. Summing up, this study allows for the expansion of approximately 50 % of the number of peaks identified in the GC–MS profile to the initial scenario, with reliable identification.

Table 2
List of compounds detected in the chromatogram from Fig. 12, including retention time (RT) in minutes of each derivative found in a pool of feces samples from healthy individuals, and identified with the CEMBIO's PCDL. (m): main peak; (s): secondary peak.

N°	Compound name (PCDL)	RT (min)	N°	Compound name (PCDL)	RT (min)
1	**Acetic acid**, methyl ester	3.62	57	(m) **Malic acid**, *O*-methoxycarbonyl, methyl ester ((s) Fumaric acid)	13.49
2	**Propionic acid**, methyl ester	4.26	58	(s) **Citric acid**, trimethyl ester, cycled	13.52
3	**Butyric acid**, methyl ester	4.98	59	**Pyroglutamic acid**, methyl ester ((s) Glutamic acid/Glutamine)	13.58
4	**Isovaleric acid,** methyl ester	5.68	60	**Proline**, *N*-methoxycarbonyl, methyl ester	13.61
5	**Valeric acid** (pentanoic), methyl ester	6.31	61	(m) **Asparagine**, cyano, *N*-methoxycarbonyl, methyl ester	13.62
6	(s) **2-Hydroxyisovaleric acid**, methyl ester	7.30	62	**Tricarballylic acid**, trimethyl ester	13.79
7	**4-Methylvaleric acid**, methyl ester (IS1)	7.33	63	(m) ***Cis*-Aconitic acid**, trimethyl ester ((s) Citric/Isocitric acid derivative)	13.95
8	**Acetoxyacetic acid**, methyl ester	7.52	64	***N*-Acetyl-Aspartic acid**, dimethyl ester	14.00
9	**Malonic acid**, dimethyl ester	7.81	65	(m) Glyceric acid, O2.O3-bis(methoxycarbonyl), methyl ester	14.07
10	**Caproic acid** (hexanoic), methyl ester	7.84	66	Suberic acid (octanedioic), dimethyl ester	14.08
11	**Methylmalonic acid**, dimethyl ester	8.44	67	(s) **Valine**, N2,N2-bis(methoxycarbonyl), methyl ester	14.09
12	**2-Ketoisocaproic acid**, methyl ester	8.59	68	**Aspartic acid**, *N*-methoxycarbonyl, dimethyl ester ((s) Asparagine derivative)	14.15
13	(m) **Levulinic acid**, methyl ester	8.83	69	(m) **Citric acid**, trimethyl ester	14.20
14	**Isoheptanoic acid**, methyl ester	8.88	70	**5-Aminovaleric acid**, *N*-methoxycarbonyl, methyl ester	14.22
15	**2-Ketobutyric acid**, methyl ester	9.03	71	(m) **2-Hydroxyglutaric acid**, *O*-methoxycarbonyl, dimethyl ester	14.34
16	**Glycolic acid**, *O*-methoxycarbonyl, methyl ester	9.07	72	(m) **Isocitric acid**, trimethyl ester	14.47
17	(m) **Maleic acid**, dimethyl ester (*cis*) ((m) Fumaric acid / (s) Malic acid)	9.35	73	**2-Methylnonanedioic acid**, dimethyl ester	14.53
18	(s) **Fumaric acid**, dimethyl ester (*trans*) ((s) Maleic acid / (s) Malic acid)	9.42	74	**Terephthalic acid**, dimethyl ester	14.55
19	**Enanthic acid** (heptanoic), methyl ester	9.45	75	(s) **Leucine**, N2,N2-bis(methoxycarbonyl), methyl ester	14.62
20	**Lactic acid**, *O*-methoxycarbonyl, methyl ester	9.52	76	(m) **Serine**, *N,O*-bis(methoxycarbonyl), methyl ester	14.64
21	**Succinic acid**, dimethyl ester ((s) 2-Ketoglutaric acid)	9.54	77	**Lauric acid** (dodecanoic), methyl ester	14.65
22	**Ethylmalic acid**, dimethyl ester	9.68	78	(s) **Isoleucine**, N2,N2-bis(methoxycarbonyl), methyl ester	14.68
23	**2-Hydroxyisobutyric acid**, *O*-methoxycarbonyl, methyl ester	9.70	79	**Azelaic acid** (nonanedioic), dimethyl ester	14.79
24	**Methylsuccinic acid**, dimethyl ester	10.07	80	(s) **Glutamic acid**, *N*-methoxycarbonyl, dimethyl ester, cycled ((s) Glutamine)	14.86
25	**Cyclohexanecarboxylic acid**, methyl ester	10.14	81	***N*-Acetyl-Glutamic acid,** dimethyl ester	14.97
26	**Citraconic acid**, dimethyl ester (*cis*) ((s) pyruvic acid)	10.47	82	(m) **Glutamic acid**, *N*-methoxycarbonyl, dimethyl ester ((s) Glutamine)	15.01
27	**5-Ketohexanoic acid**, methyl ester	10.50	83	***N*-Acetyl-iminodiacetic acid**, dimethyl ester	15.10
28	**Benzoic acid**, methyl ester	10.63	84	(m) **Methionine**, *N*-methoxycarbonyl, methyl ester	15.13
29	**Undecane**	10.66	85	**4-Methoxyphenylacetic acid**, methyl ester	15.43
30	**Mesaconic acid**, dimethyl ester (*trans*) ((s) Citraconic acid)	10.76	86	(s) **Citric acid**, *O*-methoxycarbonyl, trimethyl ester	15.46
31	(s) **Malic acid**, dimethyl ester	10.96	87	(m) **Cysteine**, N,S-bis(methoxycarbonyl), methyl ester	15.69
32	**Caprylic acid** (octanoic), methyl ester	10.96	88	**Putrescine**, N1,N4-bis(methoxycarbonyl)	15.75
33	**Alanine**, *N*-methoxycarbonyl, methyl ester	10.99	89	(s) **Isocitric acid**, *O*-methoxycarbonyl, trimethyl ester	15.78
34	(m) **Glycine**, *N*-methoxycarbonyl, methyl ester	10.99	90	**Phenylalanine**, *N*-methoxycarbonyl, methyl ester	15.87
35	**Glutaric acid**, dimethyl ester	11.08	91	**Myristic acid** (tetradecanoic), methyl ester	15.91
36	(s) **Cysteine/Serine**, *N*-methoxycarbonyl, methyl ester (loss of hydroxyl/tiol group)	11.13	92	(s) **Methionine**, N2,N2-bis(methoxycarbonyl), methyl ester	16.26
37	**(s) Levulinic acid**, acetal, methyl ester	11.18	93	**Methylmyristic acid isomer 1** (methylltetradecanoic), methyl ester	16.28
38	**Nicotinic acid**, methyl ester	11.22	94	**Methylmyristic acid isomer 2** (methylltetradecanoic), methyl ester	16.32
39	(m) **2-Hydroxyisovaleric acid**, *O*-methoxycarbonyl, methyl ester	11.34	95	**Amino acid** derivative *	16.35
40	**2-Methylglutaric acid**, dimethyl ester	11.54	96	**Pentadecanoic acid**, methyl ester	16.48
41	**Phenylacetic acid**, methyl ester	11.65	97	**2-(4-(2-Carboxylpropyl)phenyl)propionic acid**, dimethyl ester	16.50
42	**2-Aminobutyric acid**, *N*-methoxycarbonyl, methyl ester	11.96	98	**Indoleacetic acid**, methyl ester	16.56
43	**Butanedioic acid, (1-methylethyl-)**, dimethyl ester	11.96	99	**D31-Palmitic acid**, methyl ester (IS2)	16.89
44	**Pelargonic acid** (nonanoic), methyl ester	12.14	100	(m) **Ornithine**, N2,N6-bis(methoxycarbonyl), methyl ester	16.91
45	**Beta-Alanine**, *N*-methoxycarbonyl, methyl ester	12.16	101	**Palmitic acid** (hexadecenoic), methyl ester	17.02
46	(m) **2-Hydroxyglutaric acid**, dimethyl ester	12.24	102	**Jasmonic acid**, methyl ester	17.30
47	**Adipic acid** (hexanedioic), dimethyl ester	12.31	103	**Lysine**, N2,N6-bis(methoxycarbonyl), methyl ester	17.45
48	(m) **Valine**, *N*-methoxycarbonyl, methyl ester	12.47	104	(m) **Histidine**, N2,N4-bis(methoxycarbonyl), methyl ester	17.76
49	**p-Cresol**, *O*-methoxycarbonyl	12.51	105	**Linoleic acid**, methyl ester	17.88
50	**3-Phenylpropionic acid**, methyl ester	12.71	106	**Oleic acid**, methyl ester	17.90
51	**3-Methyladipic acid**, dimethyl ester	12.73	107	**Elaidic acid**, methyl ester	17.94
52	**Indole**	12.92	108	**Stearic acid** (octadecanoic), methyl ester	18.01
53	(m) **Leucine**, *N*-methoxycarbonyl, methyl ester	13.20	109	**Tyrosine**, *N,O*-bis(methoxycarbonyl), methyl ester	18.28
54	**Pimelic acid** (heptanedioic), dimethyl ester	13.26	110	**Arachidonic acid**, methyl ester	18.65
55	(m) **Isoleucine**, *N*-methoxycarbonyl, methyl ester	13.31	111	**Tricosane** (IS3)	18.81
56	(s) **Valine**, N2,N2-bis(methoxycarbonyl), methyl ester, cycled	13.48	112	**Tryptophan**, *N*-(methoxycarbonyl), methyl ester	19.18

6. Conclusions

This research sheds light on the strengths and weaknesses of methyl chloroformate derivatization in metabolomics. Based on— GC-HRMS and structural elucidation through a one-standard-each approach in the real matrix, these findings bring out helpful information regarding the identification of multiple derivatives coming from polyfunctional metabolites, scarcely or even not reported before. Moreover, it uncovers derivatives in common from metabolites that are geometric isomers of unsaturated acids such as fumaric and maleic acids, along with *cis* and *trans* aconitic acids, among others. Thus, they can undergo a transformation from one geometric isomer to another, enhancing the *cis*- over

the *trans*-form. In addition, unsaturation can go through hydration and derivatization in the form of *O*-methoxycarbonyl.

In this investigation, the reaction order of functional groups is unveiled along with the distance effect. Thus, α-hydroxy acids and α-hydroxy amines are always derivatized. Nevertheless, alternative positions of the hydroxyl group hinder the process of derivatization, then the derivatized form becomes secondary compared to the free form (main derivative). Moreover, whenever this last case occurs, dehydrated forms could also appear as secondary derivatives.

In addition, cycled intramolecular forms of derivatives can be produced and detected as secondary products, especially in polyfunctional compounds. The cyclization is also the source of some transformations not described previously. Moreover, the findings provided can demystify with a piece of evidence, some published statements not previously confirmed experimentally, such as the absent peaks in the GC analysis of arginine after treatment with MCF. Correspondingly, alkyl chloroformate derivatization is not recommended for either the study of carbohydrates or polyols. Eventually, this research has provided a detailed description of the metabolic phenotype of stool samples from healthy individuals.

In conclusion, high-resolution mass spectrometry coupled to gas chromatography has turned out to be a fundamental tool in the structural elucidation of derivatives of polyfunctional metabolites obtained after derivatization with MCF. The findings enhance the reliability of results in targeted and untargeted metabolomics studies using this derivatization procedure, contributing to a better understanding of biochemical mechanisms in diseases and aiding in the identification of biomarkers and therapeutic targets.

CRediT authorship contribution statement

M. Bajo-Fernández: Data curation, Formal analysis, Investigation, Validation, Writing – original draft. **G. Montero:** Data curation, Formal analysis, Investigation, Software, Writing – original draft. **V. Alonso-Herranz:** Formal analysis, Investigation, Methodology. **C. Barbas:** Conceptualization, Formal analysis, Resources, Supervision, Writing – review & editing, Funding acquisition. **F. Rey-Stolle:** Formal analysis, Methodology, Resources, Software, Supervision, Validation, Visualization, Writing – original draft. **A. García:** Conceptualization, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This research was funded by the Ministry of Science and Innovation of Spain (MCIN) MCIN/AEI/10.13039/501100011033 and by ERDF A way of making Europe, grant number PID2021-122490NB-I00. M. Bajo-Fernández acknowledges Fundación Universitaria San Pablo CEU for her PhD fellowship. The authors are very grateful to M. Paz Martínez-Alcázar for her scientific support.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.chroma.2024.464656.

References

[1] A. Mastrangelo, A. Ferrarini, F. Rey-Stolle, A. García, C. Barbas, From sample treatment to biomarker discovery: a tutorial for untargeted metabolomics based on GC-(EI)-Q-MS, Anal. Chim. Acta 900 (2015) 21–35, https://doi.org/10.1016/j.aca.2015.10.001.

[2] F. Rey-Stolle, D. Dudzik, C. Gonzalez-Riano, M. Fernández-García, V. Alonso-Herranz, D. Rojo, C. Barbas, A. García, Low and high resolution gas chromatography-mass spectrometry for untargeted metabolomics: a tutorial, Anal. Chim. Acta 1210 (2022) 339043, https://doi.org/10.1016/j.aca.2021.339043.

[3] J.R. Marchesi, J. Ravel, The vocabulary of microbiome research: a proposal, Microbiome 3 (2015) 31, https://doi.org/10.1186/S40168-015-0094-5.

[4] S.L. Prescott, History of medicine: origin of the term microbiome and why it matters, Hum. Microb. J. 4 (2017) 24–25, https://doi.org/10.1016/J.HUMIC.2017.05.004.

[5] L.R. Hegstrand, R.J. Hine, Variations of brain histamine levels in germ-free and nephrectomized rats, Neurochem. Res. 11 (1986) 185–191, https://doi.org/10.1007/BF00967967.

[6] J. Whipps, K. Lewis, R. Cooke, Mycoparasitism and plant disease control, in: M. Burge (Ed.), Mycoparasitism and plant disease control, Fungi Biol Control Syst (1988) 161–187.

[7] F.M. Gribble, F. Reimann, Function and mechanisms of enteroendocrine cells and gut hormones in metabolism, Nat. Rev. Endocrinol. 15 (2019) 226–237, https://doi.org/10.1038/s41574-019-0168-8.

[8] L. Mayo-Martínez, F.J. Rupérez, G. Martos-Moreno, M. Graell, C. Barbas, J. Argente, A. García, Unveiling metabolic phenotype alterations in anorexia nervosa through metabolomics, Nutrients 13 (2021) 4249, https://doi.org/10.3390/NU13124249/S1.

[9] J.J. Goedert, J.N. Sampson, S.C. Moore, Q. Xiao, X. Xiong, R.B. Hayes, J. Ahn, J. Shi, R. Sinha, Fecal metabolomics: assay performance and association with colorectal cancer, Carcinogenesis 35 (2014) 2089–2096, https://doi.org/10.1093/CARCIN/BGU131.

[10] J. Tan, C. McKenzie, M. Potamitis, A.N. Thorburn, C.R. Mackay, L. Macia, The role of short-chain fatty acids in health and disease, Adv. Immunol. 121 (2014) 91–119, https://doi.org/10.1016/B978-0-12-800100-4.00003-9.

[11] H. Zeng, S. Hamlin, B. Safratowich, W.-H. Cheng, L. Johnson, Superior inhibitory efficacy of butyrate over propionate and acetate against human colon cancer cell proliferation via cell cycle arrest and apoptosis, Curr. Dev. Nutr. 4 (2020) 364, https://doi.org/10.1093/CDN/NZAA044_063.

[12] D.C. Montrose, X.K. Zhou, L. Kopelovich, R.K. Yantiss, E.D. Karoly, K. Subbaramaiah, A.J. Dannenberg, Metabolic profiling, a non-invasive approach for the detection of experimental colorectal neoplasia, Cancer Prev. Res. (Phila) 5 (2012) 1358–1367, https://doi.org/10.1158/1940-6207.CAPR-12-0160.

[13] P. Hušek, Rapid derivatization and gas chromatographic determination of amino acids, J. Chromatogr. 552 (1991) 289–299.

[14] J. Fiori, S. Turroni, M. Candela, R. Gotti, Assessment of gut microbiota fecal metabolites by chromatographic targeted approaches, J. Pharm. Biomed. Anal. 177 (2020) 112867, https://doi.org/10.1016/J.JPBA.2019.112867.

[15] P. Husek, J.A. Rijks, P.A. Leclercq, C.A. Cramers, Fast esterification of fatty acids with alkyl chloroformates. Optimization and application in gas chromatography, J. High Resol. Chromatogr. 13 (1990) 633–638, https://doi.org/10.1002/JHRC.1240130910.

[16] S.G. Villas-Bôas, K.F. Smart, S. Sivakumaran, G.A. Lane, Alkylation or silylation for analysis of amino and non-amino organic acids by GC-MS? Metabolites 1 (2011) 3–20, https://doi.org/10.3390/metabo1010003.

[17] L. Zhao, Y. Ni, M. Su, H. Li, F. Dong, W. Chen, R. Wei, L. Zhang, S.P. Guiraud, F. P. Martin, C. Rajani, G. Xie, W. Jia, High throughput and quantitative measurement of microbial metabolome by gas chromatography/mass spectrometry using automated alkyl chloroformate derivatization, Anal. Chem. 89 (2017) 5565–5577, https://doi.org/10.1021/acs.analchem.7b00660.

[18] H.F.N. Kvitvang, T. Andreassen, T. Adam, S.G. Villas-Bôas, P. Bruheim, Highly sensitive GC/MS/MS method for quantitation of amino and nonamino organic acids, Anal. Chem. 83 (2011) 2705–2711, https://doi.org/10.1021/ac103245b.

[19] Y. Qiu, M. Su, Y. Liu, M. Chen, J. Gu, J. Zhang, W. Jia, Application of ethyl chloroformate derivatization for gas chromatography-mass spectrometry based metabonomic profiling, Anal. Chim. Acta 583 (2007) 277–283, https://doi.org/10.1016/j.aca.2006.10.025.

[20] Z. Fu, Q. Jia, H. Zhang, L. Kang, X. Sun, M. Zhang, Y. Wang, P. Hu, Simultaneous quantification of eleven short-chain fatty acids by derivatization and solid phase microextraction - Gas chromatography tandem mass spectrometry, J. Chromatogr. A 1661 (2022) 462680, https://doi.org/10.1016/J.CHROMA.2021.462680.

[21] K.F. Smart, R.B.M. Aggio, J.R. Van Houtte, S.G. Villas-Bôas, Analytical platform for metabolome analysis of microbial cells using methyl chloroformate derivatization followed by gas chromatography-mass spectrometry, Nat. Protoc. 5 (2010) 1709–1729, https://doi.org/10.1038/nprot.2010.108.

[22] P. Husek, P. Simek, Alkyl chloroformates in sample derivatization strategies for GC analysis. Review on a decade use of the reagents as esterifying agents, Curr. Pharm. Anal. 2 (2006) 23–43, https://doi.org/10.2174/157341206775474007.

[23] P. Hušek, Chloroformates in gas chromatography as general purpose derivatizing agents, J. Chromatogr. B Biomed. Sci. Appl. 717 (1998) 57–91.

[24] F.W. McLafferty, F. Tureček, Interpretation of Mass Spectra, University Science Books, Mill Valley, 1993, https://doi.org/10.1002/RCM.1290080616. Fourth ed.

[25] C. Dass, Fundamentals of Contemporary Mass Spectrometry, John Wiley & Sons, Inc., 2007, https://doi.org/10.1002/0470118490.

[26] J.T. Watson, O.D. Sparkman, Electron Ionization, in: Introduction to Mass Spectrometry: Instrumentation, Applications and Strategies for Data Interpretation, John Wiley & Sons, Ltd, Chichester, UK, 2007, pp. 315–448, https://doi.org/10.1002/9780470516898.CH6. Fourth ed.

[27] A. Behan, Mass spectrometry. A textbook, John Wiley & Sons, Ltd, 2005, https://doi.org/10.1002/ANIE.200385205. By Jürgen H. Gross.

[28] C.F. Poole, J.H. Gross, Mass Spectrometry. A Textbook, Springer, 2018, https://doi.org/10.1007/S10337-017-3400-5. Third ed.

[29] R.M. Smith, Understanding Mass Spectra: A Basic Approach, John Wiley & Sons, Inc., 2004, https://doi.org/10.1002/0471479357.

[30] P. Hušek, Z. Švagera, D. Hanzlíková, L. Řimnáčová, H. Zahradníčková, I. Opekarová, P. Šimek, Profiling of urinary amino-carboxylic metabolites by in-situ heptafluorobutyl chloroformate mediated sample preparation and gas chromatography-mass spectrometry 2, J. Chromatogr. A 1443 (2016) 211–232, https://doi.org/10.1016/j.chroma.2016.03.019.

[31] S.G. Villas-Bôas, D.G. Delicado, M. Akesson, J. Nielsen, Simultaneous analysis of amino and nonamino organic acids as methyl chloroformate derivatives using gas chromatography-mass spectrometry, Anal. Biochem. 322 (2003) 134–138, https://doi.org/10.1016/j.ab.2003.07.018.

[32] Y.P. Silva, A. Bernardi, R.L. Frozza, The role of short-chain fatty acids from gut microbiota in gut-brain communication, Front. Endocrinol. (Lausanne) 11 (2020), https://doi.org/10.3389/fendo.2020.00025.

[33] K.K. Dhillon, S. Gupta, Biochemistry, Ketogenesis, StatPearls, 2022. https://www.ncbi.nlm.nih.gov/books/NBK493179/(accessed. September 5, 2022.

[34] P. Husek, Simultaneous profile analysis of plasma amino and organic acids by capillary gas chromatography, J. Chromatogr. B Biomed. Appl. 669 (1995) 352–357.

[35] P. Hušek, Derivatization and gas chromatographic determination hydroxycarboxylic acids treated with chloroformates, J. Chromatogr. A 547 (1991) 307–314.

[36] L. Tretter, A. Patocs, C. Chinopoulos, Succinate, an intermediate in metabolism, signal transduction, ROS, hypoxia, and tumorigenesis, Biochim. Biophys. Acta 1857 (2016) 1086–1101, https://doi.org/10.1016/J.BBABIO.2016.03.012.

[37] NORD Member Organizations, Acidemia, Methylmalonic - NORD (National Organization for Rare Disorders). https://rarediseases.org/rare-diseases/acidemia-methylmalonic/(accessed May 31, 2022).

[38] R. Wojcieszak, F. Santarelli, S. Paul, F. Dumeignil, F. Cavani, R.V. Gonçalves, Recent developments in maleic acid synthesis from bio-based chemicals, Sustain. Chem. Process. 3 (2015), https://doi.org/10.1186/s40508-015-0034-5.

[39] J.S. Meek, The determination of a mechanism of isomerization of maleic acid to fumaric acid, J. Chem. Educ. 52 (1975) 541–543, https://doi.org/10.1021/ED052P541.

[40] W.B. Dunn, N.J.C. Bailey, H.E. Johnson, Measuring the metabolome: current analytical technologies, Analyst 130 (2005) 606–625, https://doi.org/10.1039/B418288J.

[41] I.J. Joye, Acids and Bases in Food, in: Encyclopedia of Food Chemistry, Elsevier, 2018, pp. 1–9, https://doi.org/10.1016/B978-0-08-100596-5.21582-5.

[42] S.R. Bhumireddy, G. Rocchetti, P. Pallerla, L. Lucini, P. Sripadi, A combined targeted/untargeted screening based on GC/MS to detect low-molecular-weight compounds in different milk samples of different species and as affected by processing, Int. Dairy. J. 118 (2021) 105045, https://doi.org/10.1016/j.idairyj.2021.105045.

[43] P. Wang, C. Mai, Y.L. Wei, J.J. Zhao, Y.M. Hu, Z.L. Zeng, J. Yang, W.H. Lu, R.H. Xu, P. Huang, Decreased expression of the mitochondrial metabolic enzyme aconitase (ACO2) is associated with poor prognosis in gastric cancer, Med. Oncol. 30 (2013) 552, https://doi.org/10.1007/S12032-013-0552-5.

[44] M. Akram, Citric acid cycle and role of its intermediates in metabolism, Cell Biochem. Biophys. 68 (2014) 475–478, https://doi.org/10.1007/s12013-013-9750-1.

[45] Y. Qiao, Z. Gao, Y. Liu, Y. Liu, Y. Cheng, M. Yu, L. Zhao, Y. Duan, Breath ketone testing: a new biomarker for diagnosis and therapeutic monitoring of diabetic ketosis, Biomed. Res. Int. (2014) 869186, https://doi.org/10.1155/2014/869186, 2014.

[46] D. Klysz, X. Tai, P.A. Robert, M. Craveiro, G. Cretenet, L. Oburoglu, C. Mongellaz, S. Floess, V. Fritz, M.I. Matias, C. Yong, N. Surh, J.C. Marie, J. Huehn, V. Zimmermann, S. Kinet, V. Dardalhon, N. Taylor, Glutamine-dependent α-ketoglutarate production regulates the balance between T helper 1 cell and regulatory T cell generation, Sci. Signal. 8 (2015) ra97, https://doi.org/10.1126/SCISIGNAL.AAB2610/SUPPL_FILE/8_RA97_SM.PDF.

[47] N. Wu, M. Yang, U. Gaur, H. Xu, Y. Yao, D. Li, Alpha-ketoglutarate: physiological functions and applications, Biomol. Ther. 24 (2016) 1–8, https://doi.org/10.4062/biomolther.2015.078.

[48] P.B. Smith, W.J. Steinbach, Candida Species, Principles and Practice of Pediatric Infectious Diseases: Fourth Edition, 2012, pp. 1196–1202, https://doi.org/10.1016/B978-1-4377-2702-9.00245-2, e4.

[49] P. Husek, H.M. Liebichb, Organic acid profiling by direct treatment of deproteinized plasma with ethyl chloroformate, J. Chromatogr. B Biomed. Appl. 656 (1994) 37–43.

[50] S.G. Villas-Bôas, A. Koulman, G.A. Lane, Analytical methods from the perspective of method standardization, Top. Curr. Genet. 18 (2007) 11–52, https://doi.org/10.1007/4735_2007_0217/Published.

[51] E.P.J.G. Neis, C.H.C. Dejong, S.S. Rensen, The role of microbial amino acid metabolism in host metabolism, Nutrients 7 (2015) 2930–2946, https://doi.org/10.3390/NU7042930.

[52] N.G. Todua, J.E. Camara, J.A. Murray, A.I. Mikaia, Stepwise extraction, chemical modification, GC-MS separation, and determination of amino acids in human plasma #, Sep. Sci. Plus. 1 (2018) 177–189, https://doi.org/10.1002/sscp.201700043.

[53] B.S. Reddy, V.N. Chary, P. Pavankumar, S. Prabhakar, Characterization of N-methylated amino acids by GC-MS after ethyl chloroformate derivatization, J. Mass Spectrom. 51 (2016) 638–650, https://doi.org/10.1002/jms.3788.

[54] K. Dettmer, N. Nürnberger, H. Kaspar, M.A. Gruber, M.F. Almstetter, P.J. Oefner, Metabolite extraction from adherently growing mammalian cells for metabolomics studies: optimization of harvesting and extraction protocols, Anal. Bioanal. Chem. 399 (2011) 1127–1139, https://doi.org/10.1007/s00216-010-4425-x.

[55] T.G. Sobolevsky, A.I. Revelsky, B. Miller, V. Oriedo, E.S. Chernetsova, I. A. Revelsky, Comparison of silylation and esterification/acylation procedures in GC-MS analysis of amino acids, J. Sep. Sci. 26 (2003) 1474–1478, https://doi.org/10.1002/JSSC.200301492.

[56] H. Kaspar, K. Dettmer, W. Gronwald, P.J. Oefner, Automated GC–MS analysis of free amino acids in biological fluids, J. Chromatogr. B Analyt. Technol. Biomed. Life Sci. 870 (2008) 222–232, https://doi.org/10.1016/J.JCHROMB.2008.06.018.

[57] R.S. Croxton, M.G. Baron, D. Butler, T. Kent, V.G. Sears, Development of a GC-MS method for the simultaneous analysis of latent fingerprint components, J. Forensic Sci. 51 (2006) 1329–1333, https://doi.org/10.1111/j.1556-4029.2006.00203.x.

[58] J. Wang, Z.-H. Huang, D.A. Gage, J.T. Watson, Analysis of amino acids by gas chromatography-flame ionization detection and gas chromatography-mass spectrometry: simultaneous derivatization of functional groups by an aqueous-phase chloroformate-mediated reaction, J. Chromatogr. A 663 (1994) 71–78.

[59] C. González-Riano, D. Dudzik, A. Garcia, A. Gil-De-La-Fuente, A. Gradillas, J. Godzien, A. López-Gonzálvez, F. Rey-Stolle, D. Rojo, F.J. Ruperez, J. Saiz, C. Barbas, Recent developments along the analytical process for metabolomics workflows, Anal. Chem. 92 (2020) 203–226, https://doi.org/10.1021/ACS.ANALCHEM.9B04553/ASSET/IMAGES/ACS.ANALCHEM.9B04553.SOCIAL.JPEG_V03.

[60] D.J. Beale, F.R. Pinu, K.A. Kouremenos, M.M. Poojary, V.K. Narayana, B. A. Boughton, K. Kanojia, S. Dayalan, O.A.H. Jones, D.A. Dias, Review of recent developments in GC–MS approaches to metabolomics-based research, Metabolomics 14 (2018) 152, https://doi.org/10.1007/s11306-018-1449-2.

[61] P. Simek, P. Husek, H. Zahradnícková, Gas chromatographic-mass spectrometric analysis of biomarkers related to folate and cobalamin status in human serum after dimercaptopropanesulfonate reduction and heptafluorobutyl chloroformate derivatization, Anal. Chem. 80 (2008) 5776–5782, https://doi.org/10.1021/AC8003506.

[62] F.M. Perrine, B.G. Rolfe, M.F. Hynes, C.H. Hocart, Gas chromatography-mass spectrometry analysis of indoleacetic acid and tryptophan following aqueous chloroformate derivatisation of Rhizobium exudates, Plant Physiol. Biochem. 42 (2004) 723–729, https://doi.org/10.1016/j.plaphy.2004.07.008.