Università degli Studi di Ferrara

DOCTORAL COURSE IN
"CHEMICAL SCIENCES"

CYCLE XXXVIII

COORDINATOR Prof. Massi Alessandro

Expanding the chemical space through non-targeted analysis and multidimensional gas chromatography coupled with mass spectrometry

Scientific/Disciplinary Sector (SDS) CHEM/01A

Candidate
Dott. De Poli Marco

_____
(signature)

Supervisor
Prof. Franchina Flavio Antonio

_____
(signature)

Co-Supervisor
Prof. Surname - Name

_____
(signature)

Possible second Co-Supervisor
Prof. Surname - Name

_____
(signature)

Year 2023/2025

Table of contents

Chapter 1
General introduction and scope 1
References 4

Chapter 2
Definitions and concepts for targeted and non-targeted chemical analysis 7
2.1. Introduction 7
2.2. Conceptual foundations 8
2.2.1. Chemical space 8
2.2.2. Targeted analysis, non-targeted analysis and suspect screening 10
2.2.3. Profiling and fingerprint 12
2.3. Resolution 14
2.3.1. Chromatographic resolution 14
2.3.2. Mass spectrometric resolution 15
2.4. Methodological aspects of non-targeted analysis: data processing and interpretation 16
2.4.1. Spectral library search 16
2.4.2. Confidence in annotations 17
2.4.3. Feature generation methods 19
2.4.4. Post-analytical treatments and data exploration 20
2.5. Analytical information continuity (AIC) 24
2.5.1. Sample preparation as an information filter 25
2.5.2. Sample introduction as a selective interface 25
2.5.3. Chromatographic separation 26
2.5.3. Mass spectrometry and informational depth 26
2.5.4. Cumulative effects across the analytical workflow 26
References 29

Chapter 3
Fundamentals of comprehensive two-dimensional gas chromatography 35
3.1. The concept of multidimensionality 35
3.2. Why 2D-GC? 38
3.2. From GC-GC to GC×GC 41
3.3. Advantages of GC×GC 46
3.4. Basic instrumentation 48
3.4.1. Injectors and injection systems 49
3.5. Column combinations 53
3.6. Modulators 54
3.7. Mass spectrometry detection 57
References 62

Chapter 4
Non-targeted terpenoid profiling in Cannabis inflorescences by sorbent-based sampling with two-stage thermal desorption coupled to GC×GC-TOFMS 71

4.1. Introduction 71
4.2. Materials and methods 72
4.2.1. Chemicals, standards and samples 72
4.2.2. Sampling and extraction 73
4.2.3. Instrumental and experimental conditions 73
4.3. Results and discussion 74
4.3.1. Optimization of thermal desorption and sorbent selection 74
4.3.2. Volatile organic compounds profiling of different cannabis chemovars 76
4.4. Conclusions 83
References 84

Chapter 5
Investigation of solvent-dependent metabolomic profiles in human serum via comprehensive two-dimensional gas chromatography and mass spectrometry 87
5.1. Introduction 87
5.2. Materials and methods 89
5.2.1. Chemicals, standards and solvents 89
5.2.2. Extraction and derivatization 89
5.2.3. Instrumental experimental conditions 90
5.3. Results and discussion 90
5.3.1. Putative selection of metabolomic features 90
5.3.2. Evaluation of extraction performance 96
5.4. Conclusions 101
References 103

Chapter 6
Hard (EI) and soft (PCI, NCI) ionization techniques as a complementary approach to expand PFAS chemical space by gas chromatography coupled to high-resolution mass spectrometry 107
6.1. Introduction 107
6.2. Materials and methods 109
6.2.1. Chemicals, standards and solvents 109
6.2.2. Instrumental experimental conditions 109
6.3. Results and discussion 110
6.3.1. PFAS fragmentation patterns 110
6.3.2. Ion source performance comparison 119
6.4. Conclusions 124
References 125

List of publications 129

This page was intentionally left blank

Chapter 1

General introduction and scope

In recent years, laboratories have faced a growing demand to identify new and unknown chemicals in complex samples [1]. Although modern analytical instruments can detect thousands of analytes in a single measurement, targeted analysis is inherently constrained by its predefined analytical scope. By focusing on the identification and quantification of a predefined set of compounds, targeted methods are just inappropriate to capture the entire chemical diversity present in a sample. Established focus approaches are therefore not feasible to address the new environmental and public health challenges, and non-targeted analysis (NTA) has rapidly emerged to fill the gaps that targeted methods cannot easily solve [1, 2]. Three factors, according to Millman *et al.* [3], determine the popular trend of non-targeted analysis: (i) the growing need for analytical determinations in environmental and biological domains, (ii) the rapid advancement of analytical instrumentation, and (iii) the development of informatics, databases, and data-processing tools.

The growing demand for analytical investigation represents a key driving force, as it is closely related to human and environmental safety [3]. In particular, the frequently uncontrolled use of chemical compounds in industrial production, agriculture, consumer goods, and pharmaceuticals has led to a situation in which societies are continuously exposed to numerous arrays of chemical substances. These compounds reach the environment and the human body through multiple pathways, including inhalation, ingestion, and absorption through the skin. Over time, this continuous exposure has become a significant concern for both scientists and health authorities [4]. For example, recent investigations have demonstrated that functional textiles, such as firefighter gear, can emit volatile per- and polyfluoroalkyl substances (PFAS) even at mild temperatures (≈38 °C), or similarly, plastic food contact materials (FCMs) have been highlighted for the migration of numerous non-intentionally added substances (NIAS) such as phthalates or hydrocarbon oligomers [5, 6]. Moreover, some of these substances may persist in the environment, bioaccumulate, or exhibit unexpected toxicological properties, making their detection and identification a matter of growing urgency [4]. In this sense, endocrine-disrupting chemicals (EDCs), such as bisphenols and parabens, represent a significant challenge, as their occurrence is well documented in surface water and blood plasma [7–9]. Then, in the biological domain, non-targeted metabolomic profiling on human serum has been used for various purposes, for example, to investigate the metabolic effects of persistent organic pollutants (POPs) exposure. By combining gas chromatography-mass spectrometry (GC-MS) and multivariate analysis, Carrizo *et al.* found significant alterations in sphingolipid and glycerophospholipid levels, suggesting lipid metabolism disruption as a potential biomarker of chronic exposure [10]. Also, Di Giovanni *et al.*, using comprehensive two-dimensional gas chromatography coupled to high-resolution time-of-flight mass spectrometry (GC×GC-(HR)TOFMS), identified through a non-targeted workflow 33 significant metabolites that distinguished

Crohn's disease patients from controls with high sensitivity and specificity [11]. Beyond aspects related to human and environmental health, the development of new methods for investigating aromatic and flavouring compounds in matrices of commercial interest is another field where non-targeted analysis is becoming increasingly popular. Essential oils, Cannabis inflorescence, and coffee products are representative examples, where comprehensive non-targeted workflows enable quality control, authenticity verification, and product characterization [12–15]. For example, Franchina *et al.* demonstrated how the combination of purge-and-trap sampling, comprehensive two-dimensional gas chromatography, and mass spectrometry enables a non-targeted exploration of aroma diversity in commercial fruit beers, allowing discrimination among products based on their volatile fingerprints [16].

The second factor, as previously mentioned, is related to the current high level of development of analytical instruments, resulting from the advancement of new chromatographs and mass spectrometers [3]. The increasingly sophisticated and automated analytical tools introduced to the market are developed and released at a pace that often exceeds the ability of laboratory operators to acquire the right skills to fully master the instruments. Moreover, the developments of GC separation and novel MS systems offer levels of resolution and sensitivity that extend far beyond the detection of only a few dozen of analytes. Indeed is now possible to simultaneously monitor thousands of molecular features within a single analytical run, pushing the boundaries of analytical scope toward a more comprehensive exploration of chemical complexity [17]. In this context, it is important to note that gas chromatography-based approaches inherently introduce a degree of selectivity, as only volatile compounds can be effectively analyzed. While this selectivity may be considered a limitation, it plays a key role in expanding the observable chemical space. In particular, gas chromatography enables access to volatile, semi-volatile, and non-polar compounds that are often poorly represented by alternative analytical techniques.

Lastly, the third factor reflects the rapid development of informatics science, the emergence of new databases, and the creation of algorithms and software that efficiently manipulate large volumes of data [3]. In the field of multidimensional gas chromatography, this progress has profoundly transformed how analytical data are handled and interpreted. The complexity of the multidimensional datasets, often comprising thousands of variables per run, has driven the development of advanced data analysis pipelines capable of turning raw signals into meaningful chemical information [18]. A wide range of commercial software solutions provides automated workflows for peak detection, spectral deconvolution, and library matching. Also, the increasing application of pattern recognition and multivariate statistical tools has opened new frontiers in non-targeted analysis. Principal component analysis (PCA), hierarchical cluster analysis (HCA), and Fisher ratio (FS) are techniques commonly used and incorporated into commercial software to uncover relevant chemical patterns across large datasets [19].

The aim of this PhD thesis is to highlight how non-targeted analytical strategies can expand the observable chemical space (see Chapter 2, section 2.2.1) using advanced analytical techniques such as comprehensive two-dimensional gas chromatography coupled to mass spectrometry (GC×GC-MS), all tuned with the proper sample preparation approach [20]. Non-targeted workflows were developed and demonstrated with three application domains: plants, human metabolites, and environmental persistent pollutants. In the first case, thanks to the high chromatographic resolution provided by GC×GC-MS and the use of a cryogenic trap, it was possible to develop a method for comparing different thermal desorption tubes in the non-targeted profiling of terpenes and terpenoids in Cannabis inflorescence (Chapter 4). Moreover, a comparison of different solvent extractions method for metabolite profiling in human serum was conducted to highlight the importance of optimizing sample preparation in defining the accessible chemical space (Chapter 5). Finally, a comprehensive investigation

of fragmentation patterns and analytical responses using hard and soft ionization techniques was conducted on a PFAS mixture by gas chromatography (GC-) and GC×GC-(HR)MS, demonstrating that varying the ionization conditions enhances the structural information and quantification of specific analytes rather than using a single ionization approach (Chapter 6).

References

[1] C. M. Fisher, K. T. Peter, S. R. Newton, A. J. Schaub, and J. R. Sobus, "Approaches for assessing performance of high-resolution mass spectrometry–based non-targeted analysis methods," *Anal. Bioanal. Chem.*, 2022, doi: 10.1007/s00216-022-04203-3.

[2] J. Hollender, E. L. Schymanski, H. P. Singer, and P. L. Ferguson, "Nontarget Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go?," *Environ. Sci. Technol.*, vol. 51, no. 20, Oct. 2017, doi: 10.1021/acs.est.7b02184.

[3] B. L. Milman and L. K. Zhurkovich, "Present-Day Practice of Non-Target Chemical Analysis," *Journal of Analytical Chemistry*, vol. 77, no. 5, May 2022, doi: 10.1134/S1061934822050070.

[4] M. Paszkiewicz, K. Godlewska, H. Lis, M. Caban, A. Białk-Bielińska, and P. Stepnowski, "Advances in suspect screening and non-target analysis of polar emerging contaminants in the environmental monitoring," *TrAC - Trends in Analytical Chemistry*, vol. 154, Sep. 2022, doi: 10.1016/j.trac.2022.116671.

[5] R. Aranda-Rodriguez *et al.*, "PFAS emissions from functional textiles using micro-chamber and thermal desorption coupled to two-dimensional gas chromatography-time of flight mass spectrometry (TD-GC×GC-TOF MS)," *J. Chromatogr. A*, vol. 1733, Sep. 2024, doi: 10.1016/j.chroma.2024.465219.

[6] I. Skjevrak *et al.*, "Non-targeted multi-component analytical surveillance of plastic food contact materials: Identification of substances not included in EU positive lists and their risk assessment," *Food Addit. Contam.*, vol. 22, no. 10, pp. 1012–1022, Oct. 2005, doi: 10.1080/02652030500090877.

[7] S. Shekhar *et al.*, "Detection of phenolic endocrine disrupting chemicals (EDCs) from maternal blood plasma and amniotic fluid in Indian population," *Gen. Comp. Endocrinol.*, vol. 241, Jan. 2017, doi: 10.1016/j.ygcen.2016.05.025.

[8] L. Martín-Pozo, M. D. C. Gómez-Regalado, I. Moscoso-Ruiz, and A. Zafra-Gómez, "Analytical methods for the determination of endocrine disrupting chemicals in cosmetics and personal care products: A review," *Talanta*, vol. 234, Nov. 2021, doi: 10.1016/j.talanta.2021.122642.

[9] J. L. Zhao *et al.*, "Determination of phenolic endocrine disrupting chemicals and acidic pharmaceuticals in surface water of the Pearl Rivers in South China by gas chromatography-negative chemical ionization-mass spectrometry," *Science of the Total Environment*, vol. 407, no. 2, Jan. 2009, doi: 10.1016/j.scitotenv.2008.09.048.

[10] D. Carrizo *et al.*, "Untargeted metabolomic analysis of human serum samples associated with exposure levels of Persistent organic pollutants indicate important perturbations in Sphingolipids and Glycerophospholipids levels," *Chemosphere*, vol. 168, 2017, doi: 10.1016/j.chemosphere.2016.11.001.

[11] N. Di Giovanni, M. A. Meuwis, E. Louis, and J. F. Focant, "Untargeted Serum Metabolic Profiling by Comprehensive Two-Dimensional Gas Chromatography-High-Resolution Time-of-Flight Mass Spectrometry," *J. Proteome Res.*, vol. 19, no. 3, Mar. 2020, doi: 10.1021/acs.jproteome.9b00535.

[12] D. J. Beale, P. D. Morrison, A. V. Karpe, and M. S. Dunn, "Chemometric analysis of lavender essential oils using targeted and untargeted GC-MS acquired data for the

rapid identification and characterization of oil quality," *Molecules*, vol. 22, no. 8, Aug. 2017, doi: 10.3390/molecules22081339.

[13] C. Ingallina *et al.*, "Cannabis sativa L. inflorescences from monoecious cultivars grown in central Italy: An untargeted chemical characterization from early flowering to ripening," *Molecules*, vol. 25, no. 8, Apr. 2020, doi: 10.3390/molecules25081908.

[14] C. de C. Couto, D. W. H. Chávez, E. M. M. Oliveira, O. Freitas-Silva, and S. Casal, "SPME-GC-MS untargeted metabolomics approach to identify potential volatile compounds as markers for fraud detection in roasted and ground coffee," *Food Chem.*, vol. 446, Jul. 2024, doi: 10.1016/j.foodchem.2024.138862.

[15] P. Q. Tranchida, F. A. Franchina, and L. Mondello, "Analysis of essential oils through comprehensive two-dimensional gas chromatography: General utility," *Flavour Fragr. J.*, vol. 32, no. 4, pp. 218–227, Jul. 2017, doi: 10.1002/ffj.3383.

[16] F. A. Franchina, D. Zanella, E. Lazzari, P. H. Stefanuto, and J. F. Focant, "Investigating aroma diversity combining purge-and-trap, comprehensive two-dimensional gas chromatography, and mass spectrometry," *J. Sep. Sci.*, vol. 43, no. 9–10, May 2020, doi: 10.1002/jssc.201900902.

[17] S. E. Prebihalo *et al.*, "Multidimensional Gas Chromatography: Advances in Instrumentation, Chemometrics, and Applications," *Anal. Chem.*, vol. 90, no. 1, Jan. 2018, doi: 10.1021/acs.analchem.7b04226.

[18] P. H. Stefanuto, A. Smolinska, and J. F. Focant, "Advanced chemometric and data handling tools for GC×GC-TOF-MS: Application of chemometrics and related advanced data handling in chemical separations," *TrAC - Trends in Analytical Chemistry*, vol. 139, Jun. 2021, doi: 10.1016/j.trac.2021.116251.

[19] T. J. Trinklein, C. N. Cain, G. S. Ochoa, S. Schöneich, L. Mikaliunaite, and R. E. Synovec, "Recent Advances in GC×GC and Chemometrics to Address Emerging Challenges in Nontargeted Analysis," *Anal. Chem.*, vol. 95, no. 1, Jan. 2023, doi: 10.1021/acs.analchem.2c04235.

[20] F. A. Franchina, D. Zanella, L. M. Dubois, and J. F. Focant, "The role of sample preparation in multidimensional gas chromatographic separations for non-targeted analysis with the focus on recent biomedical, food, and plant applications," *J. Sep. Sci.*, p. jssc.202000855, Oct. 2020, doi: 10.1002/jssc.202000855.

This page was intentionally left blank

Chapter 2

Definitions and concepts for targeted and non-targeted chemical analysis*

2.1. Introduction

Over the past decades, chemical analysis has undergone extensive transformations. As highlighted in Chapter 1, advances in instrumentation and computational capacity have made possible the development of highly sensitive, robust, and high-resolution analytical platforms. These technological evolutions not only increase the number and diversity of compounds that can be separated and detected but also reshape the strategies for resolving complex chemical mixtures.

The Chemical Abstracts Service (CAS) Registry currently lists millions of compounds, and more than 350,000 substances have been listed for global production [1]. Online resources, such as ChemSpider and PubChem, as well as mass spectral libraries like the National Institute of Standards and Technology (NIST), provide access to millions of chemical structures and their corresponding MS fragmentation patterns. This vast and growing chemical space places an increasing demand on analytical workflows, particularly in metabolomics, environmental science, food chemistry, and the petroleum industry, where sample matrices often contain hundreds to thousands of components, many of which are unknown or unexpected [2].

Analytical chemists can approach chemical analysis with two complementary strategies: targeted analysis, which focuses on detecting known predefined compounds with high specificity, and non-targeted analysis, which aims to capture a broader chemical profile without prior selection of analytes [3]. Non-targeted analysis is nowadays the driving approach in modern omics sciences for the discovery and understanding of complex systems.

While this distinction is operationally well established and widely adopted, it does not, on its own, account for the full range of factors that influence which chemical information is ultimately retained, excluded, or transformed throughout the analytical process.

*This section has been adapted from: **De Poli M.**, Polidoro A., Franchina F.A.; "*Chapter 1: Definitions and concepts for targeted and non-targeted chemical analysis";* From: *"Gas Chromatography-Mass Spectrometry for Omics Applications",* Wiley, ISBN: 978-1-394-20927-9 (2026).

A more integrated perspective should consider how sample preparation, chromatographic and spectrometric resolution, and data processing strategies interact to determine which components of a sample become chemically accessible and which are excluded. These methodological choices are not merely technical; they directly influence the scope, quality, and structure of the information obtained from an analytical system. In this context, it is important to recognize that what is detected, interpreted, and reported is shaped by tools, procedures, and experimental approaches employed to investigate the sample [4].

2.2. Conceptual foundations

2.2.1. Chemical space

The concept of "chemical space" can be described as the theoretical set of all chemical compounds that may be present within a given sample [5] and in the framework of non-targeted analysis, the detected features are confined within this space. Generally, the organization of a chemical space can be described using various molecular descriptors. This approach essentially involves plotting the relevant physicochemical properties of a set of molecules and assessing their distribution in a multidimensional space defined by the number of variables identified. Common descriptors include boiling point, density, and heat capacity. As an illustrative example, Figure 2.1 shows a representation of a chemical space obtained by plotting molecular weight (MW < 1250 Da; x-axis) against the polarity coefficient (logKow; y-axis).

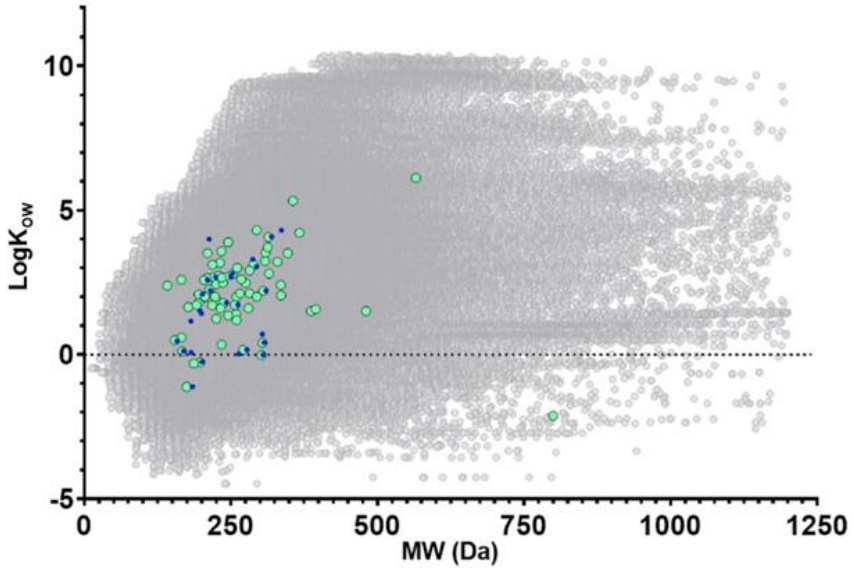


Figure 2.1. *Representation of a chemical space showing how compounds of interest are distributed according to their molecular weight (MW; x-axis) and polarity (log $K_o w$; y-axis). The plot, according to [6], includes 699,013 organic compounds (molecular mass < 1200 Da) from the Distributed Structure-Searchable Toxicity (DSSTox) database, shown as grey dots. Green and blue dots represent a subset of pharmaceuticals and personal care products (PPCPs) [6].*

Chemists can theoretically synthesize an unlimited number of chemical compounds, and many more can be made by nature, making the chemical space theoretically infinite [7]. However, considering specific molecule borders, the range of accessible entities is considerably restricted [2]. For instance, the number of potentially synthesizable compounds according to the elemental relationship is approximately 400 million, as represented by the ZINC database [8] mentioned in Table 2.1. Another subset includes compounds present in the previously mentioned Chemical Abstract Service (CAS) Registry [9], with over 250 million compounds, and around half are included in already cited databases such as PubChem [10] or ChemSpider [11].

Table 2.1. *Chemical space and subspaces. Adaptation from [2].*

Space	Comments, information source	Compounds
All possible compounds	MW < 500 Da [7]	10^{20}-10^{200}
Synthesizable compounds	ZINC database [8]	~4×10^{8}
All known compound	Registered in Chemical Abstracts [9]	>250 000 000
Chemical compounds of biological significance	PubChem database [10]	>120 000 000
Common chemical compounds	ChemSpider database [11]	>100 000 000
Metabolites	Metabolites, bio compounds in the METLIN database [12]	>960 000
	Human Metabolome database (HMBD) [13]	>220 000
Environmental compounds	EPA's CompTox Chemistry Dashboard database [14]	>800 000

Moreover, considering that an actual sample composition is notably different from the entire known space, the number of substances commonly found in various matrices is generally lower. In the biological domain, it falls below a million, according to the Metlin [12] and Human Metabolome Database (HMDB) [13] databases. Similarly, environmental matrices contain a comparable number of prevalent compounds, as shown by the EPA database [14]. By continuing this procedure, which increasingly narrows the field of observation, the number of compounds in the chemical space becomes increasingly reduced. However, these values differ greatly from the detectable chemical space of the sample, which is influenced not only by the field of study (environmental rather than biological) but also by numerous decisions during the method development.

In this sense, Gabrielle Black *et al.* [15] proposed an informatic tool conceived by eight key filtering parameters for defining detectable chemical space, as a basis for mapping: (1) sample matrix type, (2) extraction solvent, (3) extract pH, (4) extraction/cleanup media, (5) elution buffers, (6) instrument platform, (7) ionization type, and (8) ionization mode. In the proposed resource, each parameter yields a list of compounds, occupying specific regions within chemical space, and ultimately, these eight lists of amenable compounds are then compared with the overlapping compounds that define the detectable space (Figure 2.2).

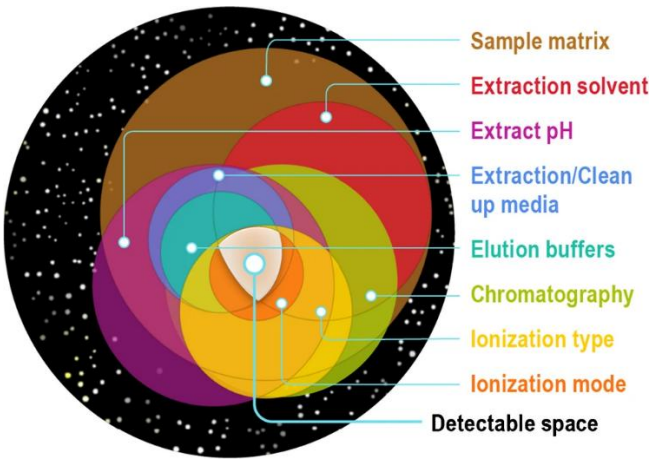


Figure 2.2. *Illustration of the intersection of individual filtering steps to define the detectable space [15].*

The concept of chemical space throughout this PhD thesis serves as a unifying element across the various experimental applications, providing a conceptual tool to support the non-targeted analytical strategies.

2.2.2. Targeted analysis, non-targeted analysis, and suspect screening

The result of a chemical analysis is defined by how the information is selected, isolated, and interpreted across the various experimental conditions or workflows available. Targeted analysis (TA) aims to detect and quantify a predefined set of compounds, often referred to as "known" [16], using reference standards and optimized analytical conditions to achieve maximum sensitivity and specificity [3]. This approach is especially prevalent in fields such as environmental monitoring and food safety, where regulatory compliance and the unambiguous quantification of specific and highly characterized substances are required, due to hazardous properties and safety issues [17].

The use of reference standards in targeted analysis enables the unequivocal identification of compounds, typically through matching retention times and characteristic mass spectra or ion ratios, thereby ensuring analytical reliability and reproducibility. The entire analytical workflow is tailored for achieving maximum selectivity. Moreover, the sample preparation is deliberately designed to isolate target analytes at regulatory limits while removing matrix components considered as interferences [18].

In contrast, non-targeted analysis (NTA) offers a different paradigm: instead of limiting the scope to a predefined set of analytes, NTA includes a system of strategies/schemes which aims to detect, characterize, and, when possible, identify as many chemical constituents as possible in a given sample, the so-called "unknown unknowns", regardless of prior expectations or database availability [19]. This holistic approach is particularly valuable in contexts where awareness or knowledge may be limited, such as environmental monitoring for emerging pollutants or the discovery of novel biomarkers in biological matrices. To achieve a comprehensive chemical understanding of a given sample, multiple methodologies should ideally be combined, since each has its own selectivity, and no single technique is universally applicable.

Unlike TA, where selectivity is deliberately maximized to reduce complexity, NTA workflows emphasize "chemical breadth". Sample preparation is usually designed to capture the widest possible spectrum of compounds, employing broad extraction methods. This necessity for inclusiveness, however, introduces a fundamental trade-off. As the range of recovered analytes expands, the potential matrix effects and analytical challenges related to sample complexity increase [20, 21].

Moreover, the massive volume and complexity of data generated often demand time-consuming pre- and post-processing, as well as advanced chemometric and statistical approaches for data mining and interpretation. In this sense, the lack of standardized protocols, incomplete reference databases, and the computational demand and uncertainty associated with handling thousands of features per sample constitute the current challenges of non-targeted analysis [22].

An intermediate strategy between targeted and non-targeted analysis is suspect screening (SS), which focuses on detecting a predefined list of compounds or chemical classes whose presence in the sample is uncertain. This terminology is mostly used for environmental applications, and unlike targeted analysis, suspect screening does not rely on the availability of analytical reference standards. Instead, it utilizes prior information, such as molecular formulas, exact masses, predicted fragments, or retention indices, often derived from databases or in silico predictions. These compounds are often referred to as "known unknowns", since their structures are known, but their occurrence in the sample is not. By narrowing the scope of study while maintaining a non-targeted acquisition approach, suspect screening reduces the computational and interpretive challenge typical of non-targeted workflows, while still allowing for exploratory flexibility aligned with a specific research objective [21, 23]. A straightforward application of this strategy is demonstrated in the study by Casey *et al.*, who employed a suspect screening strategy using GC-(HR)MS to analyze incineration stack emissions that potentially contained PFAS. Even in the absence of analytical standards, several fluorinated compounds were tentatively identified by combining library matching, observation of diagnostic fragments, and confirmation of molecular ions across multiple ionization modes (EI, PCI, NCI) [24].

In Figure 2.3, a concentrical relationship among TA, SS, and NTA is shown, taken from the efforts of the BP4NTA community [25]. The first panel, "Concept", represents the different levels of chemical knowledge, moving from compounds already characterized (*known* and *known unknowns*) to those yet unrevealed (*unknown unknowns*). The second panel, "Category", relates these concepts to analytical strategies: TA focuses on specific, predefined compounds; SS extends the investigation on compounds suspected to be present; and NTA represents the broadest approach, aiming to detect any compound. The last panel, "In Practice", highlights how these categories align with laboratory and routine classifications, distinguishing between compounds from reference standards, substances contained in predefined set lists, and those not yet characterized.

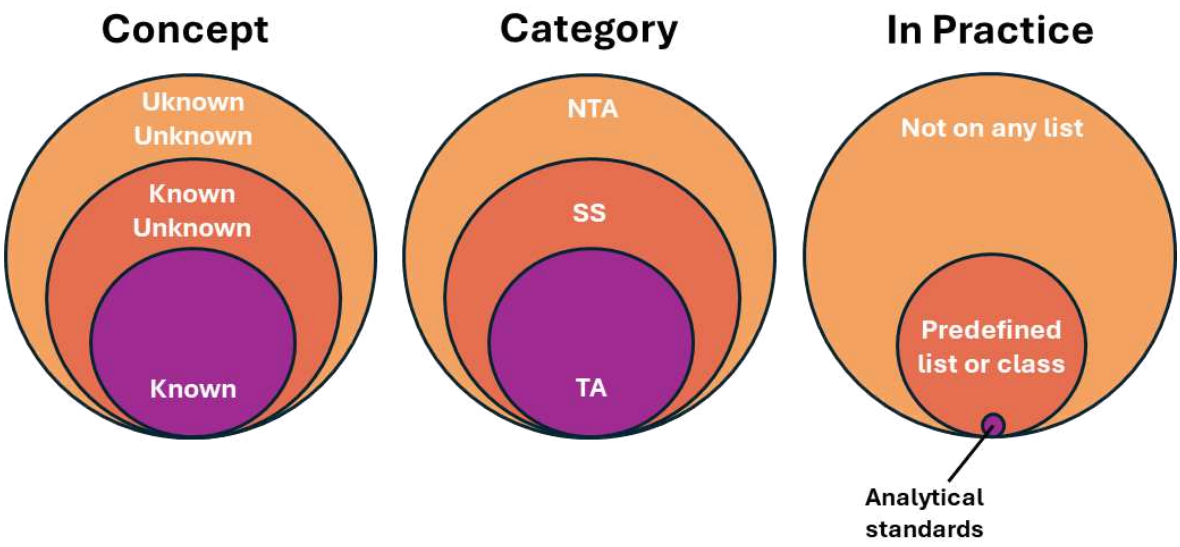


Figure 2.3. *Visual representation of targeted, suspect screening, and non-targeted analysis across multiple visions, from conceptual foundations to an operational point of view. Taken and adapted from [25].*

2.2.3. Profiling and fingerprint

Within the broad domain of non-targeted analysis, the literature also distinguishes between profiling (P) and fingerprinting (F) strategies. Although the terms are sometimes used interchangeably, they represent distinct approaches. Profiling is designed to extract comprehensive information about the constituents of a sample, including retention time, spectral data, and intensity profiles. When profiling is conducted without prior selection of analytes, it is typically described as non-targeted profiling [26]. For instance, Wong *et al.* [27] conducted a non-targeted profiling of *Aquilaria malaccensis* (agarwood) essential oils using GC×GC, detecting approximately 550 chromatographic peaks. Among these, 155 compounds were tentatively identified, mainly monoterpenes and sesquiterpenes, which characterize the chemical complexity of agarwood.

In contrast, fingerprinting prioritizes pattern recognition rather than compound identification. The goal is not to identify or quantify individual analytes, but to capture global chemical signatures that differentiate samples based on similarities and differences in their overall signal profiles. This approach is particularly useful for classification problems, where analytical discrimination is achieved through the multivariate comparison of features, and the number of detected peaks or features directly contributes to the robustness of the fingerprint [17, 28]. An example of this strategy is the study developed by Caratti *et al.*, who applied non-targeted GC×GC-MS/FID analysis, combined with multivariate fingerprinting, to differentiate the volatile profiles of balsamic vinegar. This approach distinguished samples based on geographical origin and fermentation practices through both unsupervised and supervised models [29].

Figure 2.4 offers a conceptual visualization of the analytical strategies discussed, mapped within a three-dimensional space defined by three typical descriptors: the number of analyzed samples (X-axis), the number of investigated analytes (Y-axis), and the overall volume of chemical information obtained from the sample (Z-axis). Each strategy is represented as a “cloud” because it is often difficult to draw clear borders and extensions, emphasizing the absence of strict boundaries between such categories and groups.

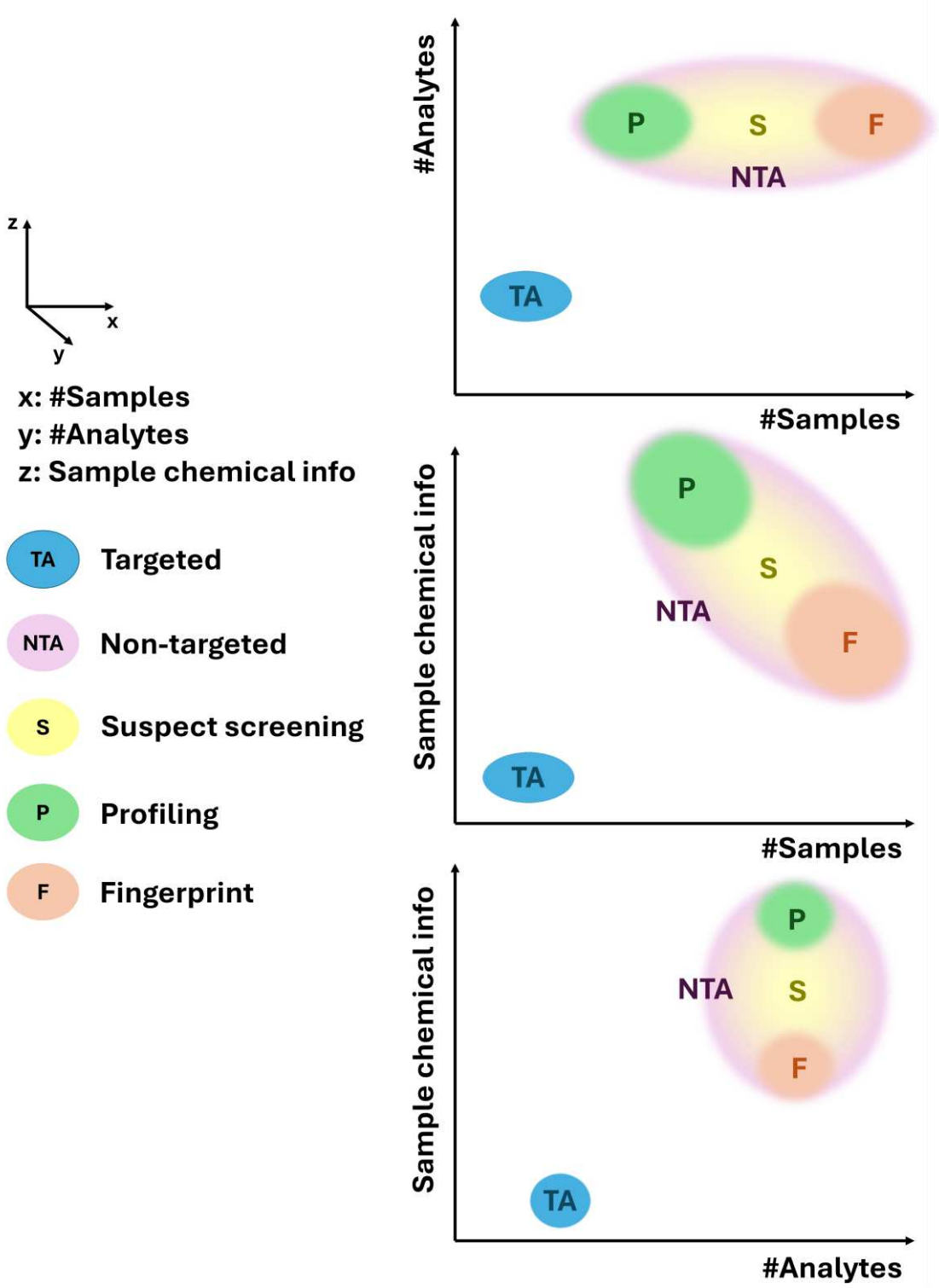


Figure 2.4. *Analytical strategies positioned in a 3D space defined by three descriptors: number of samples (X), number of analytes (Y), and chemical information (Z), illustrating their overlap and gradation across methodological dimensions.*

In this representation, TA occupies a relatively narrow region, defined by a limited number of analytes and a focused informational output. Even when large numbers of samples are processed, each one is typically assessed independently by comparing it to a calibration curve. Direct comparisons between samples are not inherent to the strategy, though they can be added later depending on the study design (*e.g.*, seasonal variation in air contaminant levels).

In contrast, non-targeted analysis (NTA cloud) encompasses a much broader region, covering a wider array of analytes and samples, while generating a substantially larger volume of chemical data. Within this non-targeted domain, the three sub-strategies are illustrated. Suspect screening (S in Figure 1), located near the center of the NTA space, focuses on compounds with known structures but unknown presence (so-called "known unknowns"), and typically relies on information such as accurate mass, predicted fragmentation patterns, or retention index to support identification. Profiling (P in Figure 2.4) is positioned higher along the Z-axis, as it emphasizes the detailed chemical characterization of each sample. Fingerprinting (F in Figure 2.4), in contrast, extends along the X-axis, reflecting its focus on comparing global signal patterns across multiple samples without requiring compound identification. This multidimensional perspective highlights that analytical strategies are not isolated categories but rather interconnected approaches that vary along continuous gradients of selectivity, coverage, and informational output. There are no fixed borders, only regions shaped by methodological choices and analytical goals.

If high-resolution data are recorded (*e.g.*, GC×GC-(HR)MS), it is up to the analyst to decide how the data are subsequently exploited. Provided that sample preparation and analytical conditions are sufficiently broad to own all the chemical information, the acquired data can be used as fingerprinting (just patterns, no weight on chemical identification which might be accessible afterwards), targeted (by focusing only to the predetermined analytes of interest), or post-targeted (*e.g.*, looking retrospectively for the presence of newly regulated contaminants that were not regulated at the time of the acquisition).

2.3. Resolution

Resolution in analytical chemistry is more than a numerical parameter; it defines the level of detail at which the chemical information can be accessed and understood. In this context, the chromatographic and mass spectrometric resolution defines how finely the chemical space can be resolved into distinguishable components.

2.3.1. Chromatographic resolution

Chromatographic resolution quantifies how well a separation system can distinguish adjacent analytes based on retention time. It is formally expressed as:

$$R = \frac{t_{R2} - t_{R1}}{\frac{W_1 + W_2}{2}} \qquad \text{Eq. 2.1}$$

where t_{R2} and t_{R2} are retention times, and W_1 and W_2 are peak widths. This parameter is influenced not only by retention but also by selectivity and efficiency, which play a significant role in overall chromatographic resolution [30].

While this formula offers a standard definition, chromatographic resolution should be understood as a strategic decision rather than a fixed metric. In targeted analysis, chromatographic resolution is best seen as a fit-for-purpose parameter, as it only needs to be high enough to separate the analytes of interest from nearby interferences. Beyond that, increasing resolution may add time and complexity without improving results [31].

In contrast, non-targeted strategies require broader and more efficient separation capabilities. When the analytical aim is to explore or monitor a wide array of unknowns, unresolved coelutions can result in spectral entanglement, misannotation, or complete loss of signal. In such cases, resolution plays a defining role in determining how much of the sample's chemical complexity can be separated, detected, and interpreted. Chromatographic efficiency, a parameter that reflects how effectively a column can produce narrow,

symmetrical peaks with minimal broadening, is also critical as peak sharpness and symmetry directly influence the separation, the quality of mass spectra, and the success of deconvolution algorithms [30, 32]

Multidimensional chromatographic systems, such as GC×GC, are especially valuable in this context. By combining multiple separation mechanisms and distributing analytes across two retention axes, they expand peak capacity while maintaining, or even increasing, efficiency. Additionally, the use and combination of such multidimensional approaches/techniques contribute to the reduction of sample dimensionality [33].

Such multidimensional, high-resolution, and high-efficiency chromatograms not only enhance the number of components detected but also exhibit structured elution patterns that reveal meaningful chemical class information, enabling semi-supervised exploration even in the absence of analytical standards, and facilitating downstream interpretation and multivariate analysis [32, 34].

Therefore, chromatographic resolution and efficiency together are not simply about how many peaks can be separated, but about how well the chemical information is organized for interpretation.

2.3.2. Mass spectrometric resolution

Mass spectrometric resolution is defined as:

$$R = \frac{m}{\Delta m} = \frac{m/z}{\Delta\, m/z} \qquad \text{Eq. 2.2}$$

where m is the *m/z* of a selected ion, and Δm is the smallest mass difference that can be resolved.

Low-resolution instruments, such as quadrupoles operating at unit resolution, are often well-suited for targeted workflows. This approach is useful for monitoring predefined *m/z* with high precision and reproducibility. However, their capacity to resolve complex mixtures or identify unexpected species is limited [35].

High-resolution mass analyzers, including Time-of-Flight (TOF) and Orbitrap systems, can resolve closely spaced ions and deliver mass accuracies often within a few parts per million. This enables the discrimination of isobaric compounds, supports the assignment of elemental formulas, and facilitates the annotation of unknown features in complex matrices. In contexts where spectral libraries are incomplete or unavailable, high mass resolution serves as a surrogate for chemical knowledge, enabling inference where direct identification is not possible [35].

Notably, mass resolution also improves signal-to-noise ratios and reduces the ambiguity of peak assignments, which are crucial for both qualitative and quantitative analyses. Furthermore, when combined with accurate mass and isotope pattern recognition, high-resolution data enhances confidence in structural hypotheses [36, 37].

As in chromatographic resolution, non-targeted strategies greatly benefit from high mass spectrometric resolution, both in qualitative and separation capabilities. Indeed, when high-resolution mass spectrometry is combined with soft ionization techniques (*e.g.*, PCI, NCI), it supports more reliable chemical identification, including for unknowns. On the other hand, the ability to accurately distinguish closely spaced ions becomes critical for spectral deconvolution [36, 38, 39].

Both chromatographic and mass spectrometric resolution define the degree to which a given chemical domain can be detailed into its single components. Analytical techniques that operate with low resolution settings may offer higher throughput at the cost of losing information about individual components. Conversely, maximizing resolution increases detailed analytical information, although it presents challenges in terms of data volume, complexity, and processing demands [36, 38, 40].

Understanding resolution not only as a stand-alone, traditional metric (chromatographic and mass spectrometric), but also as a unified concept of overall analytical resolution, is necessary for successful non-targeted strategies.

2.4. Methodological aspects of non-targeted analysis: data processing and interpretation

2.4.1. Spectral library search

Without reference standards, compound identification can be tentatively achieved in a non-targeted workflow by a spectral library search after data acquisition. This is particularly true for chromatography-mass spectrometry-based techniques (*e.g.*, gas chromatography, GC), where library search with electron impact is undoubtedly a pivotal step. This process involves assigning a match value to the acquired spectrum based on the spectrum stored in a spectral library. Through this approach, the analyst can rapidly compare experimental mass spectra with reference data to propose a "tentative" or "putative" identification, a fundamental aspect of non-targeted workflows. The basic aspect of this process is the availability of the reference spectral libraries, which not only enable tentative compounds identification but also give access to an extensive set of information accumulated over the decades, such as CAS number or retention index, all of which together support high-confidence annotations [41].

In gas chromatography, electron ionization (EI) represents the most adopted ion formation technique due to its robustness and the exceptional reproducibility of fragmentation patterns across instruments [1, 42]. Consequently, EI-based analyses benefit from the availability of the most extensive mass spectral libraries, such as the Wiley and NIST databases [43, 44]. These libraries, in combination with dedicated matching algorithms, allow the calculation of parameters including similarity, reverse match, and probability, which provide quantitative criteria for assessing the reliability of compound identification. Also, MS/MS spectral references exist for fewer compounds (NIST 23, MS^2, [43]) and dedicated databases to specific fields (biological domains) are common as well (mzCloud, [45]). Table 2.2 presents a brief list of mass spectral libraries, including the number of compounds per database.

Table 2.2. *Main libraries of mass spectra of low-molecular weight compounds. Adapted from [41].*

Name	Compounds	Comments
Wiley Registry 13th, EI-MS^1 [44]	~ 750 000	Volatile compounds
NIST 23, EI-MS^1 [43]	~ 450 000	Volatile compounds
NIST 23, MS^2 [43]	~ 35 000	MS/MS for volatile compounds
mzCloud [45]	~ 25 000	MS^n, pharmaceuticals, metabolites

Typically, match scores above 900 are considered excellent, while values between 700 and 900 indicate acceptable similarity, depending on the complexity of the sample matrix [23]. While EI libraries are well established, the same cannot be said for advanced mass spectrometric techniques such as high-resolution MS or soft ionization approaches, for which available spectral data continue to be limited also due to the limited fragmentation patterns [19]. Finally, spectral reference not only relies on an available database, but also, it can be obtained either experimentally, through the measurement of analytical standards, or theoretically, by computational approaches such as *in silico* fragmentation [21].

Despite these methodological advancements, several challenges remain in non-targeted GC-MS analysis. One critical issue concerns the ionization conditions, where variation in detector sensitivity and acquisition parameters can result in spectral variations that hinder consistent library matching [19, 46]. Then, a key point is that nowadays, no other methods can replace definitive structural confirmation with authentic chemical standards [46]. Finally, a last consideration should be pointed to the rapid growth of non-targeted data requirements, such as large storage or sharing infrastructures. Individual measurement files can easily exceed one gigabyte in size, making repositories essential for long-term archiving and retrospective analysis [47]. Some open-access digital platforms, such as the NORMAN database system [48], are already available to store and share useful information, demonstrating the potential of collaborative efforts.

2.4.2. Confidence in annotations

When a spectral similarity search does not yield a conclusive library match, the peak detected by the software remains unidentified in the peak table list; however, even in the absence of spectral references, if available in the laboratory, the use of high-resolution mass spectrometry (HRMS) techniques provides an opportunity to perform accurate compound identification. In particular, HRMS instruments suitable for routine analysis are now capable of sensitive, non-targeted detection of thousands of substances within the short time frames required by chromatographic separations. These instruments, essentially TOF or Orbitrap analyzer, can achieve mass accuracy of less than 1 ppm and an operational mass range of up to 1000 Da [3]. These analytical capabilities enable the detection of compounds with exact mass up to 5 digits and the generation of detailed molecular information.

In this context, a benchmarking publication for non-targeted analysis is the 5-level classification scheme introduced by Schymanski in 2014 [37], which aims to effectively communicate confidence in identification using HRMS (including both GC- and LC-). This scheme is widely accepted by the scientific community and serves as a concise, easy-to-

follow manual to guide users in correctly identifying compounds. The illustration of these confidence levels is proposed in Figure 2.5.

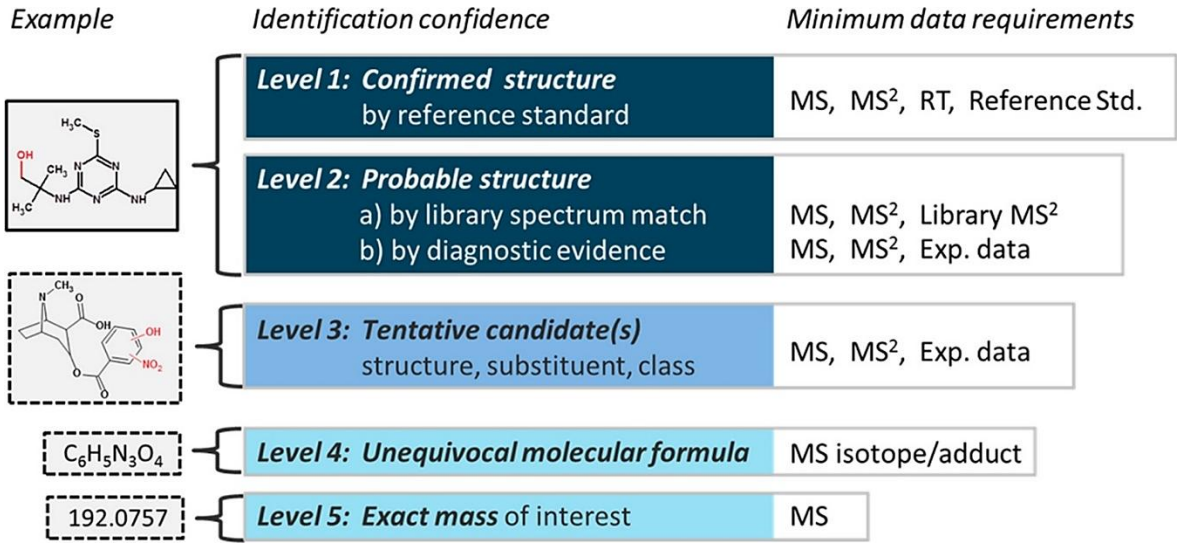


Figure 2.5. *Confidence levels in high-resolution mass spectrometry identification. Taken from [37]*

Starting from the bottom, Level 5 represents the measure of exact mass. This is the common starting point for further identification efforts, where prior analysis has led to high interest in the exact mass due to, for example, high intensity or high frequency of occurrence in the sample. Moving up one position, Level 4 describes the assignment of an unequivocal molecular formula based on the exact mass (Level 5). This annotation is not trivial, as interpreting molecular formulas is challenging, especially for high molecular masses or in the absence of a characteristic isotope pattern. Elements commonly found in organic molecules exhibit characteristic isotopic distributions that can be exploited to determine molecular formulas. Among these, bromine and chlorine show particularly distinctive isotope patterns due to their high relative abundances of heavier isotopes (M + 2 peaks), while other elements, such as carbon, contribute to smaller isotopic variations. Level 3 represents a "grey zone" according to the author of the publication, where "*evidence exists for possible structure(s), but insufficient information for one exact structure only (e.g., positional isomers)*". This level of confidence typically arises when spectral data are consistent with several structures, such as stereoisomers, that cannot be distinguished solely based on mass spectrometric information. For instance, identical elemental compositions and nearly indistinguishable fragmentation patterns may prevent the precise localization of functional groups within a molecule. Moreover, Level 2 refers to a degree of confidence where the available analytical evidence supports the assignment of a probable structure. In particular, Level 2 is divided into two sub-levels, 2a and 2b. In Level 2a, identification is based on a direct match between the experimental and reference spectra. The spectral similarity must be unambiguous and reproducible, and ideally, additional supporting information, such as collision energy, resolution, or retention index, should be considered to further strengthen the identification. In Level 2b, instead, no reference spectrum is available, and in such cases, characteristic ions provide structural clues allowing the analyst to propose one plausible molecular structure. At the top of the scheme, Level 1 represents the highest goal in non-targeted analysis, indicating the greatest confidence in compound annotation where the proposed molecular structure has been definitively verified through direct comparison with an authentic analytical standard.

Despite its broad acceptance and practical value, the Schymanski framework suffers from some limitations and ambiguities that hinder its consistent application. For example, the

number of fragment ions necessary to support probable structure identification, as well as the thresholds for spectral similarity, such as the minimum match factor required to assign a compound to Level 2 based on library comparison [47]. However, overall, this type of scheme represents one of the concrete efforts to harmonize non-targeted workflow strategies.

2.4.3. Feature generation methods

By definition, non-targeted analysis involves manipulating large amounts of data to detect as many analytes as possible in the sample. In chromatographic separations, especially when using GC×GC-MS, the data essentially consist into a number of peaks corroborated by chromatographic and spectrometric properties, commonly listed in a peak table. In scientific language, these attributes are also referred to as "features" to emphasize the importance of each individual detection beyond simple spectral library searches. As discussed in Chapter 1, a single chromatographic run can generate a huge number of analytes, ranging from a few hundred to thousands; this data must then be multiplied by the number of samples and their respective replicates within the analysis batch. As a result, it became progressively clear over time that manipulating these big data using the same strategies employed to process a few dozen analytes was no longer feasible, as the time required to extract useful information from the sample became far longer than the analysis itself. In this context, thanks to advances in computer science, chemists had to reinvent the way to manipulate and make use of this type of data. Stilo *et al.*, Reichenbachet *et al.* and Caratti *et al.* [28, 49, 50] have proposed and summarized in three different publications, the main data processing approaches for GC×GC-MS, with a particular focus on non-targeted cross-sample analysis. These strategies, essentially for feature generation, include: datapoints, tiles, regions, peaks, and peak-region features.

According to the three authors, the simplest type of representation is called a datapoint feature, in which each intensity value, and thus each pixel in the 2D chromatogram, is handled as a distinct feature. This method ensures full coverage of all analytes, but it produces incredibly large datasets with many redundant data points per analyte, increasing the risk of errors and misalignment.

Moving further, peak features are derived from the detection and integration of chromatographic peaks, each representing an individual analyte. Ideally, there is a one-to-one correspondence between features and analytes, which allows direct chemical interpretation. However, in non-targeted analyses of complex mixtures, automatic peak detection and matching across many chromatograms can be difficult due to coelution, noise, missing peaks, and retention-time variability. As a result, comprehensive matching of peak features across multiple datasets is often incomplete or, most commonly, error-prone.

To reduce the complexity, tile features and region features aggregate multiple datapoints into larger, defined portions of the chromatographic plane. Tiles are typically rectangular, while regions can have irregular or chemically meaningful shapes. These approaches reduce the number of variables and are less sensitive to chromatographic drift; however, they are less selective, as multiple analytes may elute within the same area, and single peaks may be split across different tiles or regions. Besides that, region-based features are particularly useful in group-type analyses, such as petrochemical analyses, where the cumulative signal of compounds within a defined chemical region (*e.g.*, paraffines, olefines, aromatic compounds) is of greater interest than the signals of individual compounds.

Finally, peak-region features have been developed as a hybrid solution combining the strengths of region- and peak-based approaches. In this method, chromatograms are first aligned and combined into a composite chromatogram that includes all analyte peaks from a dataset. Each detected peak defines a specific region in a "feature template", which can then be mapped back to every chromatogram in the set. In Figure 2.6, a feature template is

reproduced from Magagna *et al.* [51], where a peak-region approach was adopted to discriminate the volatile profile of Cocoa from different regions and treatments by automated headspace solid-phase microextraction (HS-SPME) with comprehensive two-dimensional gas chromatography (GC×GC) coupled to thermal modulation (TM) and flow modulation (FM), followed by mass spectrometric detection.

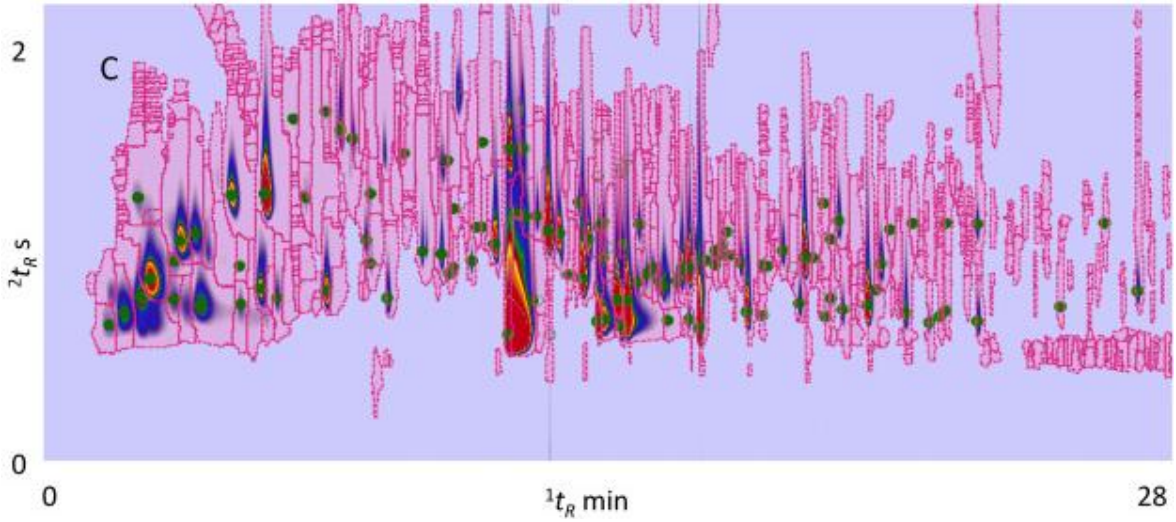


Figure 2.6. *Peak-region features template from roasted cocoa sample. The feature template consists of the sum of the contributions of each peak in the sample. Every peak region is highlighted in pink and outlined in purple, arranged in irregular shapes due to the separation in the first and second dimensions. Taken and adapted from [51].*

2.4.4. Post-analytical treatments and data exploration

Prior to statistical analysis, data preprocessing represents an essential step in the treatment of GC- and GC×GC data. This stage aims to remove nonchemical variability and to extract chemical information from raw instrumental signals. Proper preprocessing allows chromatograms from different runs to be directly comparable by minimizing instrumental noise, baseline drifts, and retention time variability that may otherwise obscure important chemical differences [52, 53]. Moreover, data pre-treatment steps can be tailored to retain as many molecular features as possible or introduce filtering steps based on the analytical goal. For example, extracted features can be further filtered based on abundance cut-offs (*e.g.*, data processing signal-to-noise ratio), or mass ranges and retention time windows of interest [54].

Among the various operations, baseline correction removes minor background signals caused by stationary-phase bleed, detector drift, or column degradation [53, 55, 56]. If uncorrected, these effects may alter quantitative data or hinder peak detection. Baseline correction can be performed through several approaches; the simplest is the subtraction of a blank chromatogram from the acquired dataset.

The subsequent stage involves peak picking and deconvolution, which are employed to identify and quantify individual analyte signals [35, 42, 57]. Peak-picking algorithms detect chromatographic peaks in extracted ion chromatograms (EICs) based on criteria such as maximum peak width and signal-to-noise (S/N) ratio. Moreover, despite the enhanced separation provided by GC×GC-MS, peak overlap can still occur, especially in complex samples. Spectral deconvolution, therefore, involves algebraic manipulations (*e.g.,* PARAFAC) performed by dedicated algorithms that resolve co-eluting compounds by reconstructing their individual signal contributions across both chromatographic and spectral

dimensions [35, 57]. This step enhances peak purity and provides more reliable spectral information for compound identification.

Finally, retention time alignment is designed to compensate for temporal variations in chromatographic conditions, such as fluctuations in pressure, carrier gas flow, or temperature. These variations can cause minor but significant shifts in analyte retention times between runs. If not corrected, these alterations can result in mismatched features across datasets, complicating comparative analyses. Alignment algorithms correct these discrepancies by adjusting peak positions so that identical compounds across multiple chromatograms share consistent retention times [52, 53, 55].

Once the data has been settled and well organized, the next step often correspond in the application of a series of statistical tests for data exploration. These serve to highlight differences between samples or even develop predictive models to classify samples into different categories. The application of statistical and mathematical models to chemical data is a growing field of study known as chemometrics. This science tends to classify approaches into two categories, supervised and unsupervised, which are often referred to simply as non-targeted approaches in literature (Figure 2.7). Supervised methods incorporate prior sample information, such as class labels, directly into the model-building process. In contrast, unsupervised methods do not use class information as an input; instead, they aim to explore the intrinsic structure of the data, revealing natural groupings or patterns among samples [52].

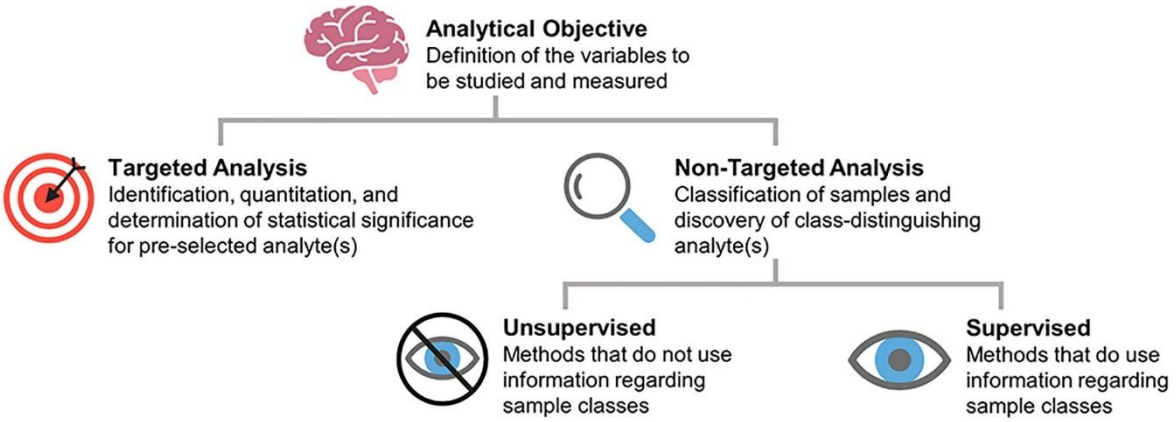


Figure 2.7. *Schematic representation of targeted and non-targeted analysis, highlighting the role of supervised and unsupervised methods according to the analytical objective [52].*

Among unsupervised techniques, the most popular for exploratory data is, without any doubt, principal component analysis (PCA), which is often included in scientific publications. PCA enables reducing data dimensionality by transforming a large set of correlated variables into a smaller set of orthogonal ones, called principal components (PCs), by producing a linear combination of the original variables that preserves the original variance. This algebraic transformation enables visualization of very complex datasets, facilitating the identification of patterns and potential outliers. Two main graphical outputs are typically used to interpret PCA results: the scores plot, which illustrates how individual samples are distributed in the principal component space, and the loadings plot, which reveals how strongly each variable contributes to and influences the corresponding principal components [22]. By combining scores and loadings, PCA provides a qualitative understanding of the main sources of variance in the dataset and the relationships between samples and features. Franchina *et al.* [58], for example, used a PCA score plot (Figure 2.8) to show the separation in the space of the principal components of different Sativa and Indica

Cannabis inflorescences by developing a stir bar sorptive extraction (SBSE) thermal desorption by GC×GC-MS.

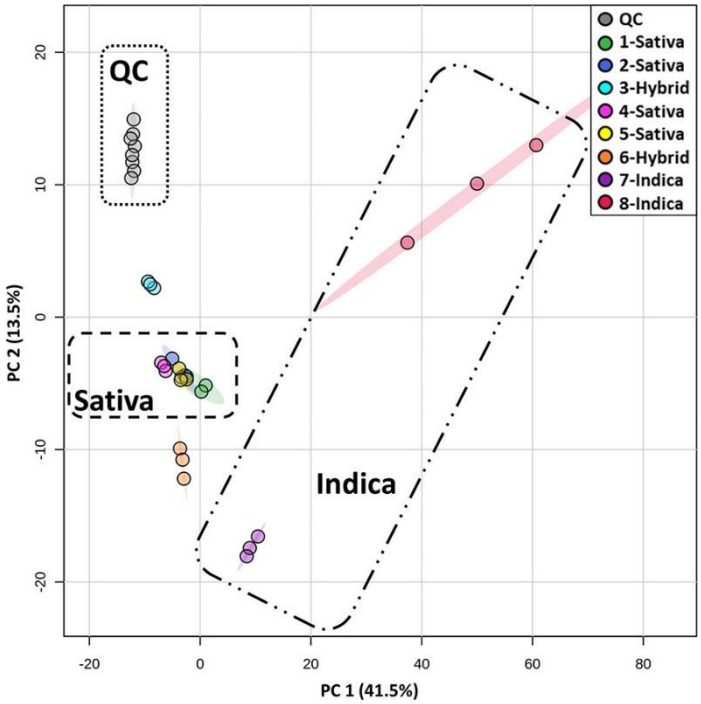


Figure 2.8. *PCA score plot from the SBSE-GC×GC-MS non-targeted analysis on Cannabis inflorescences. Taken from [58].*

Another unsupervised approach commonly applied in multivariate analysis is hierarchical cluster analysis (HCA), which measures the similarity (or dissimilarity) between objects within a dataset. HCA evaluates distances between samples in a multidimensional space and successively groups them hierarchically. The result is represented as a graphic called dendrogram, which visualizes the degree of correlation among samples. Two fundamental parameters define the clustering process: the distance metric and the linkage criterion. The first one determines how similarity is quantified in terms of distances (*e.g.*, Euclidean or Manhattan distances), while the second specifies how distances between clusters are computed during consecutive merging steps (*e.g.*, average linkage, median linkage) [35]. This exploratory tool becomes particularly powerful when combined with a heat map, as the resulting visualization not only displays the hierarchical relationships among objects but also shows the magnitude of their similarities or differences through variations in color and intensity. An example of HCA is given by the work of Lukic' *et al.* (Figure 2.9, [59]), to emphasize the separation of Croatian monovarietal wines using 60 volatile aroma compounds detected from GC-MS and GC×GC-MS analysis.

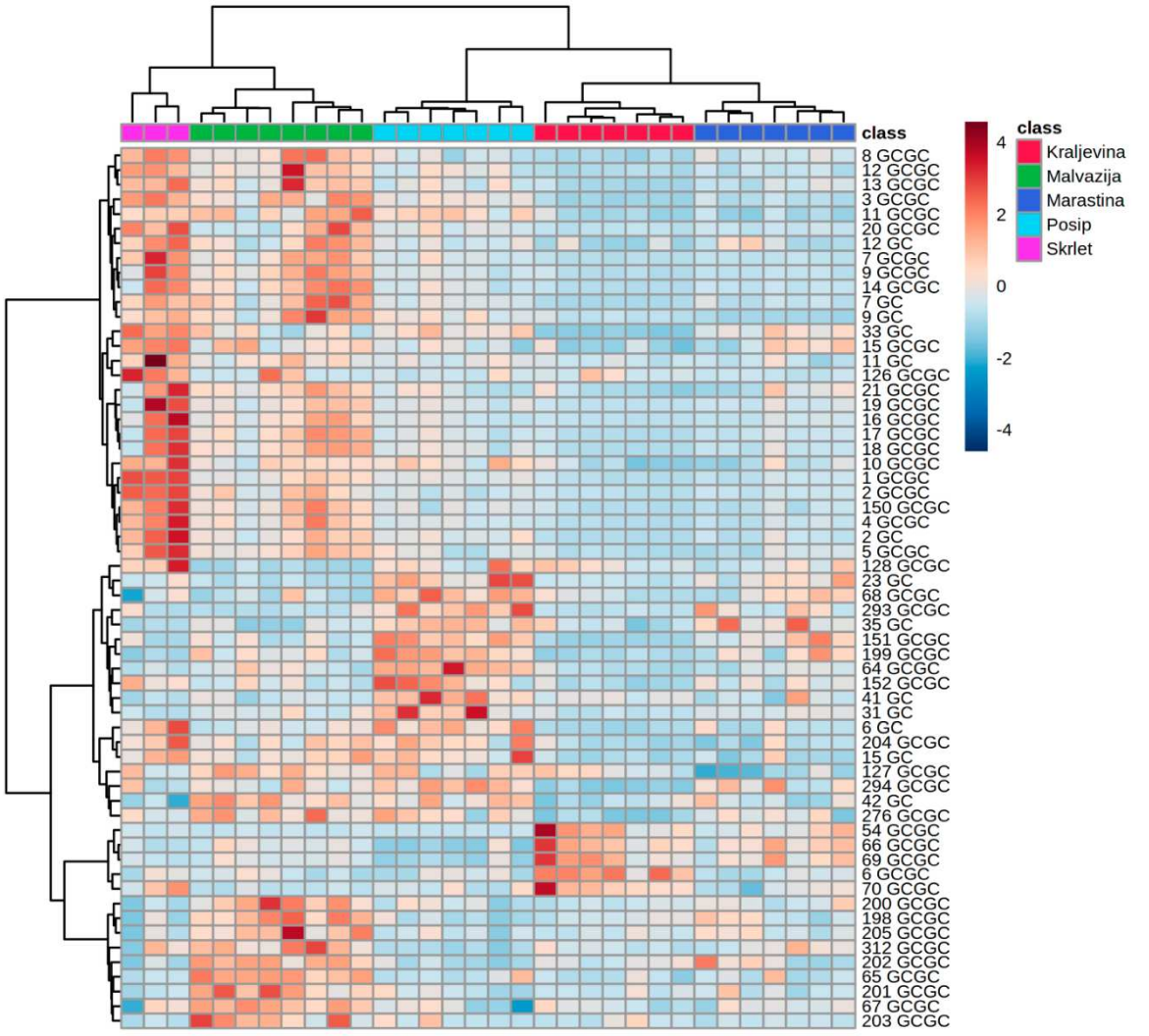


Figure 2.9. *Hierarchical clustering analysis of volatile aroma compound profiles of Croatian monovarietal wines obtained by GC-MS and GC×GC-TOF-MS analysis. The rows in the heatmap represent compounds, and the columns indicate samples (also highlighted in different colors in the first row, starting from the top). Above and on the side of the figure, the dendrogram structure displays the similarity relationships between objects, while within the heatmap, the colors indicate the abundance of the compounds, with red indicating higher abundance and blue indicating lower predominance [59].*

On the supervised method side, one of the simplest techniques is Fisher Ratio Analysis (FRA), which is based on the statistical principles of one-way analysis of variance (ANOVA). The method calculates the F-ratio, defined as the ratio of between-class to within-class variance, for each chromatographic variable (which can be areas, pixels, or tiles). Then, variables with higher F-ratios indicate greater contribution to class separation, and these data are typically a ranked "hit list" of features, ordered in descending F-ratio values [56]

Partial Least Squares-Discriminant Analysis (PLS-DA) is also a supervised technique that has been frequently reported in the literature. This technique uses multivariate regression to correlate chromatographic variables with categorical variables (the "supervised" sample classes). Similar to PCA, PLS-DA generates a linear combination of the original variables, known as latent variables; however, in this case, the direction is oriented to maximize the separation between predictors (*e.g.*, chromatographic properties, in the chromatographic context) and class membership, thus identifying the features responsible for the class

differences [60]. An example of PLS-DA outputs is shown in Figure 2.9, taken from the work of Feng *et al.* [61], for distinguishing rice origin based on GC-MS analysis.

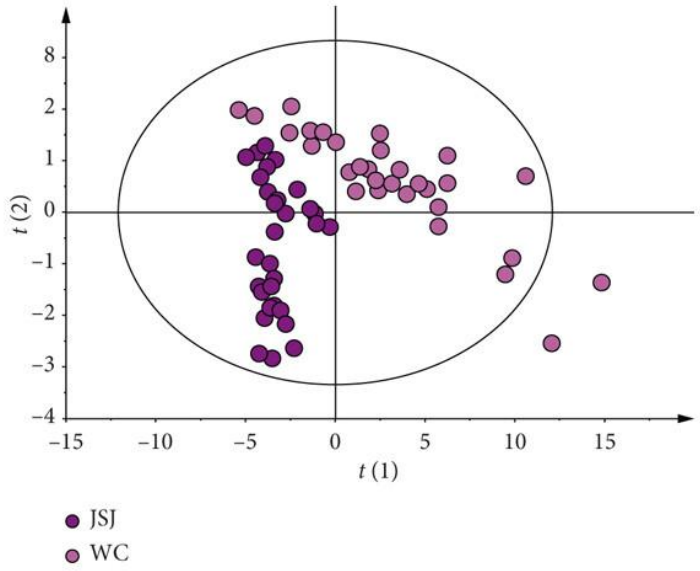


Figure 2.9 *PLS-DA score plot of rice samples from two geographical origins (Wuchang, WC; and Jiansanjiang, JSJ). Each point represents one rice sample, and the axes (t1 and t2) correspond to the first two latent variables of the model. The clear separation between the two groups indicates significant differences in their metabolic profiles, suggesting that the cultivation origin has a strong influence on rice metabolite composition. Taken from [61].*

2.5. Analytical information continuity (AIC)

The analytical workflow defines experimentally what can and cannot be accessed in the sample. In targeted and non-targeted strategies, methodological choices on the sample preparation, chromatographic separation, and mass spectrometric detection together determine which portions of the chemical space become analytically visible/accessible. These three dimensions govern the informational output of any analysis, not only shaping the detectability of compounds but also confining the knowledge that can be drawn from samples. Existing terminology often fails to capture how these three dimensions interact across the analytical workflow.

To address this conceptual gap, this PhD thesis refer to analytical information continuity (AIC) as a concept to evaluate the analytical workflows based on their capacity to preserve and transmit the chemical complexity of a sample throughout all methodological stages, from sample preparation to data utilization/interpretation, focusing on the degree to which each analytical decision retains the informational depth/level of the original sample.

AIC can be defined as the extent to which the chemical information present in a sample is retained across the analytical workflow and made available for interpretation. It reflects a continuum, not a binary division. At one end, workflows are highly selective and narrow, optimized for specific analytes under controlled conditions, typically seen in classical TA. At the other end, methods attempt to capture the broadest possible chemical profile, as in NTA, using minimal sample manipulation and high-resolution techniques.

This concept shifts the evaluative focus from analytical intent (*what are we looking for?*) to analytical capacity (*how much of the sample's chemical space can be accessed?*). In doing so, AIC becomes not only a practical consideration but also a measure of informational depth; it quantifies what becomes analytically accessible. This concept is illustrated in Figure

2.10, where the analytical workflow acts as a selective and transformative interface that governs how chemical reality is converted into chemical knowledge.

During this process, information loss is unavoidable due to the intrinsic selectivities of analytical procedures and strategies, regardless of their nature.

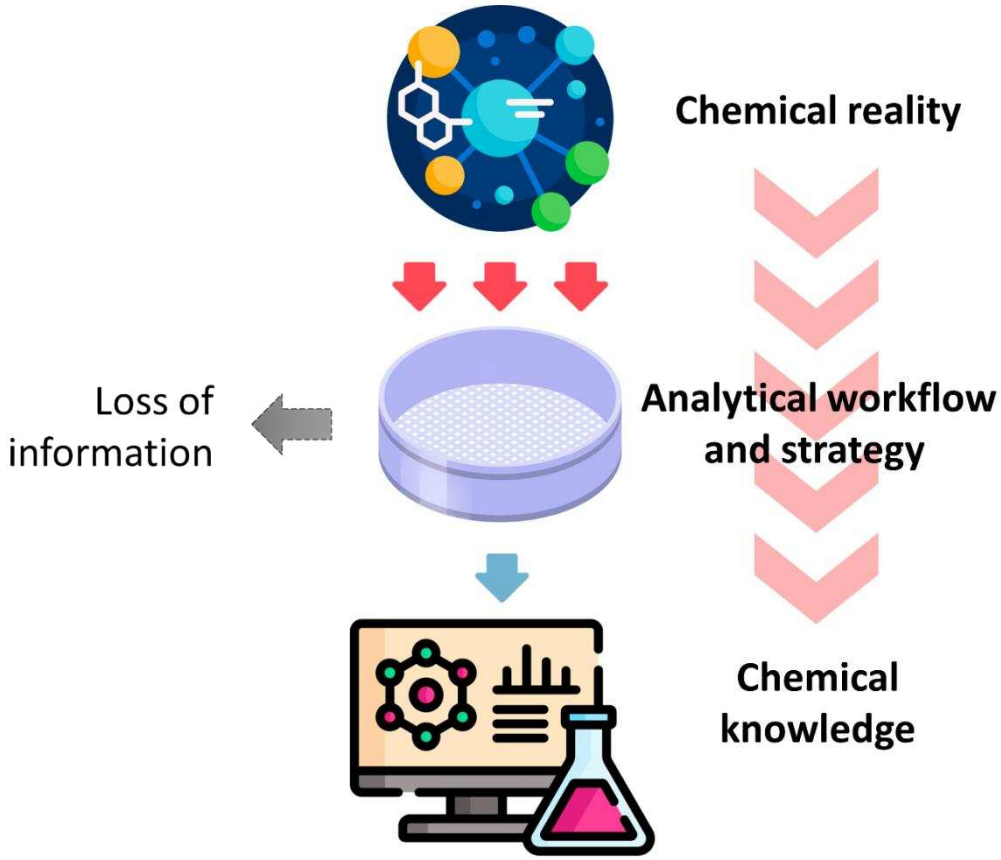


Figure 2.10. *Schematic representation of how the analytical workflow mediates the transformation of chemical reality into chemical knowledge.*

2.5.1. Sample preparation as an information filter

Sample preparation is the first and one of the most critical steps in ensuring the continuity of analytical information. Every extraction, purification, or derivatization step acts as a filter, selectively enriching certain analytes while excluding others. Highly specific preparation procedures (*e.g.*, solid-phase extraction tailored for particular analyte classes) enhance sensitivity and reproducibility but reduce the scope of detectable compounds [62]. This results in low AIC, since much of the sample's chemical content is discarded.

Conversely, non-selective sample preparation strategies such as dilute-and-shoot, thermal desorption, or direct injections are designed to preserve the broadest possible chemical profile. These approaches maximize the initial AIC by allowing more of the sample's chemical reality to enter the analytical pipeline.

However, it is important to note that less selective or minimal sample preparation does not necessarily translate into greater AIC. While broader extraction strategies may allow more compounds to enter the analytical system, they also increase the risk of matrix effects, signal suppression, and co-elutions, all of which can compromise detection and interpretation [63]. The effectiveness of sample preparation must therefore be evaluated in conjunction with the resolution and sensitivity of the downstream analytical techniques. In this sense, sample

preparation is not an isolated step, but rather a process that must be tuned with the entire analytical workflow and research objective.

2.5.2. Sample introduction as a selective interface

Even before chromatographic separation, the act of introducing a sample into the analytical system is inherently selective. This has profound implications for non-targeted analysis. Although NTA is often outlined as an inclusive strategy aimed at capturing the broadest possible array of chemical constituents, the operational reality is that inclusion is conditional. For GC-based NTA, analytes must be volatile, thermally stable, and compatible with the carrier gas and liner materials. These requirements introduce implicit selectivity, constraining the accessible chemical domain even when no prior target list is defined [64]. Thus, while NTA is conceptually open-ended, its analytical information continuity is always instrumentally mediated.

In addition, the injection step plays a decisive role in chromatographic performance, since high injection efficiency ensures narrow, well-defined bands entering the column, which is fundamental for achieving high-resolution GC separations. From the perspective of AIC, the injection system is therefore not a passive interface but both a filter and a performance factor. This physical and chemical threshold not only determines what becomes accessible but also how effectively it can be resolved downstream in the analysis process (*i.e.*, in the column, in the mass spectrometer, and in data elaboration and interpretation).

2.5.3. Chromatographic separation

Chromatographic efficiency and resolution directly affect the number of distinguishable chemical signals that can be extracted from a sample, thereby influencing the degree of AIC. High-efficiency systems generate narrow, symmetrical peaks and minimize band broadening, enabling a larger number of compounds to be separated, detected, and eventually identified, while higher resolution enables a larger number of compounds to be separated, detected, and eventually identified.

Capillary GC inherently offers very high separation efficiency due to the narrow internal diameter, uniform stationary phase coating, and optimal flow dynamics. When coupled with the structured modulation of GC×GC, this efficiency is further amplified, producing exceptionally sharp and well-resolved peaks. Such performance not only maximizes peak capacity but also sets a strong foundation for accurate mass spectrometric detection, feature deconvolution, and reliable compound identification, all of which are indispensable for achieving high AIC.

In GC×GC, the separative logic enhances detectability by distributing compounds more evenly across a 2D chromatographic space, resolving co-elutions that would remain hidden in a single column. Structured chromatograms, band compression, and various separation mechanisms all help preserve the sample's complexity throughout the separation process. In contrast, workflows relying on conventional one-dimensional chromatography might sacrifice a portion of the chemical space due to peak overlap and limited resolving power. Although often sufficient for targeted applications, their AIC is constrained.

2.5.3. Mass spectrometry and informational depth

Mass spectrometry contributes to the information continuity as well. For instance, excluding a range of *m/z* values from the EI-MS acquisition could limit or erroneously affect the recorded information. MS also defines the level of structural information available for each detected signal. High-resolution instruments, such as TOF and Orbitrap analyzers already mentioned, provide precise mass measurements and high resolving power, which are essential for distinguishing isobaric compounds, assigning elemental formulas, and

deconvoluting co-eluting peaks, resulting in high AIC [35]. These capabilities are especially critical in non-targeted strategies, where analytes may not be present in spectral libraries and accurate mass becomes a necessary feature for tentative identification. In such contexts, high mass resolution extends the interpretive depth of the analysis, allowing greater confidence in structural assignments and facilitating more reliable annotation of unknown features [37, 65].

In contrast, unit-resolution instruments, such as quadrupoles, while effective for monitoring specific transitions in targeted analyses, they impose a limitation on what can be learned from the data. They lack detail to resolve fine mass differences and are generally unsuited for exploratory or high-complexity workflows [35]. Thus, mass resolution plays a decisive role not just in detection but in shaping the scope of chemical knowledge that can be extracted. When considering MS/MS systems, the choice mainly depends on the nature of the analyzers and the acquisition mode. For example, triple quadrupoles in tandem MS mode exclude most of the information by filtering only the selected parent/daughter ions and are used exclusively for target analytes, which, however, results in limited AIC.

In GC-based workflows, mass spectrometry also benefits from the extensive EI spectral libraries, which greatly enhance the identification power and are highly reliable and standardized across different instrument suppliers. Beyond standard EI, the use of complementary ionization modes such as soft EI, negative chemical ionization (NCI), and positive chemical ionization (PCI) can selectively emphasize different structural features or compound classes [66]. This multimodal approach enhances AIC by expanding the range of detectable analytes, increasing confidence in identification, and facilitating the characterization of compounds that may be underrepresented or poorly ionized under a single ionization mode.

2.5.4. Cumulative effects across the analytical workflow

The continuity of analytical information is not defined by any isolated single step, but by the cumulative effect of all stages. A method with high-resolution MS but selective sample preparation remains limited in what it can reveal. Likewise, universal sample extraction paired with low-resolution separation and detection yields an incomplete picture. It is the preservation of information across the entire analytical chain, from sampling to preparation, separation, detection, and interpretation, that defines a method's AIC. Because each of these steps introduces its own selectivity, achieving a truly comprehensive chemical picture often requires combining multiple analytical methodologies. No single technique can capture all compounds present in a complex sample due to the diverse physicochemical properties and concentrations. By integrating complementary techniques, such as GC and LC, or multiple ionization modes, analysts can mitigate limitations and expand the accessible chemical space.

Finally, to fully maximize the recorded chemical data, efforts must also be made during data analysis by selecting suitable tools and tailoring them to the analytical strategy and objective [67]. Otherwise, most of the information would remain hindered, offering little benefit and not balancing the efforts of experimental conditions [68].

This holistic view allows AIC to be conceptualized as a gradient shaped by trade-offs (Figure 2.11). Essentially, every methodological choice, what to extract, what to separate, what to detect, and how to process, affects not only the data obtained but also the knowledge that can be produced.

Workflows with high AIC are better suited for exploratory studies, which aim to unravel unknowns and identify unexpected compounds (non-targeted analysis). In contrast, low-AIC

workflows are optimized for quantification and regulatory compliance, but at the cost of discarding potentially relevant information (targeted analysis).

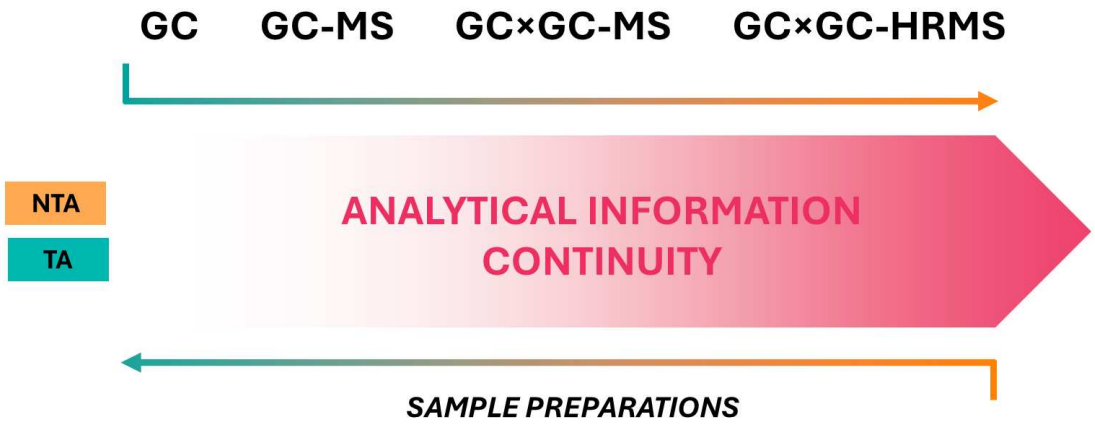


Figure 2.11. *Gradient of AIC. The central arrow changes color as the AIC increases, starting from the left with targeted approaches (involving invasive sample preparation and one-dimensional techniques, e.g., GC) and moving to the right toward more non-targeted approaches (requiring minimal sample preparation and employing advanced analytical techniques, e.g., GC×GC-HRMS).*

References

[1] D. Ciccarelli *et al.*, “Bridging knowledge gaps in human chemical exposure via drinking water with non-target screening,” *Crit. Rev. Environ. Sci. Technol.*, vol. 55, no. 3, 2025, doi: 10.1080/10643389.2024.2396690.

[2] B. L. Milman and I. K. Zhurkovich, “The chemical space for non-target analysis,” *TrAC - Trends in Analytical Chemistry*, vol. 97, Dec. 2017, doi: 10.1016/j.trac.2017.09.013.

[3] N. Z. Ballin and K. H. Laursen, “To target or not to target? Definitions and nomenclature for targeted versus non-targeted analytical food authentication,” *Trends Food Sci. Technol.*, vol. 86, Apr. 2019, doi: 10.1016/j.tifs.2018.09.025.

[4] I. Hcking, *Representing and Intervening: Introductory Topics in the Philosophy of Natural Science.* Cambridge University Press, 1983. doi: 10.1017/CBO9780511814563.

[5] K. E. Manz *et al.*, “Non-targeted analysis (NTA) and suspect screening analysis (SSA): a review of examining the chemical exposome,” *J. Expo. Sci. Environ. Epidemiol.*, vol. 33, no. 4, Jul. 2023, doi: 10.1038/s41370-023-00574-6.

[6] B. Schulze *et al.*, “An assessment of quality assurance/quality control efforts in high resolution mass spectrometry non-target workflows for analysis of environmental samples,” *TrAC - Trends in Analytical Chemistry*, vol. 133, Dec. 2020, doi: 10.1016/j.trac.2020.116063.

[7] R. Van Deursen and J. L. Reymond, “Chemical space travel,” *ChemMedChem*, vol. 2, no. 5, May 2007, doi: 10.1002/cmdc.200700021.

[8] “ZINC 15, A free database of commercially-available compounds.” Accessed: Nov. 12, 2025. [Online]. Available: http://zinc.docking.org

[9] “CAS, Chemical Abstracts Service.” Accessed: Nov. 12, 2025. [Online]. Available: https://pubchem.ncbi.nlm.nih.gov/

[10] “PubChem dabase.” Accessed: Nov. 12, 2025. [Online]. Available: https://pubchem.ncbi.nlm.nih.gov/

[11] “ChemSpider dabase.” Accessed: Nov. 12, 2025. [Online]. Available: http://www.chemspider.com/

[12] “METLIN database.” Accessed: Nov. 12, 2025. [Online]. Available: https://metlin.scripps.edu/landing_page.php?pgcontent=mainPage

[13] “Human Metabolome database.” Accessed: Nov. 12, 2025. [Online]. Available: http://www.hmdb.ca/

[14] “EPA’s CompTox Chemistry Dashboard database.” Accessed: Nov. 12, 2025. [Online]. Available: https://comptox.epa.gov/dashboard

[15] G. Black *et al.*, “Exploring chemical space in non-targeted analysis: a proposed ChemSpace tool,” *Anal. Bioanal. Chem.*, vol. 415, no. 1, Jan. 2023, doi: 10.1007/s00216-022-04434-4.

[16] J. L. Little, C. D. Cleven, and S. D. Brown, “Identification of ‘known unknowns’ utilizing accurate mass data and chemical abstracts service databases,” *J. Am. Soc.*

Mass Spectrom., vol. 22, no. 2, pp. 348–359, Feb. 2011, doi: 10.1007/s13361-010-0034-3.

[17] J. Devers, D. I. Pattison, A. B. Hansen, and J. H. Christensen, "Comprehensive two-dimensional gas chromatography as a tool for targeted and non-targeted analysis of contaminants of emerging concern in wastewater," *Talanta*, vol. 282, Jan. 2025, doi: 10.1016/j.talanta.2024.127032.

[18] F. Hernández, T. Portolés, E. Pitarch, and F. J. López, "Gas chromatography coupled to high-resolution time-of-flight mass spectrometry to analyze trace-level organic compounds in the environment, food safety and toxicology," *TrAC - Trends in Analytical Chemistry*, vol. 30, no. 2, Feb. 2011, doi: 10.1016/j.trac.2010.11.007.

[19] B. Gruber, F. David, and P. Sandra, "Capillary gas chromatography-mass spectrometry: Current trends and perspectives," *TrAC - Trends in Analytical Chemistry*, vol. 124, Mar. 2020, doi: 10.1016/j.trac.2019.04.007.

[20] P. Hajeb, L. Zhu, R. Bossi, and K. Vorkamp, "Sample preparation techniques for suspect and non-target screening of emerging contaminants," *Chemosphere*, vol. 287, Jan. 2022, doi: 10.1016/j.chemosphere.2021.132306.

[21] M. Pourchet *et al.*, "Suspect and non-targeted screening of chemicals of emerging concern for human biomonitoring, environmental health studies and support to risk assessment: From promises to challenges and harmonisation issues," *Environ. Int.*, vol. 139, Jun. 2020, doi: 10.1016/j.envint.2020.105545.

[22] B. J. Pollo *et al.*, "Chemometrics, Comprehensive Two-Dimensional gas chromatography and 'omics' sciences: Basic tools and recent applications," *TrAC - Trends in Analytical Chemistry*, vol. 134, Jan. 2021, doi: 10.1016/j.trac.2020.116111.

[23] J. Hollender *et al.*, "NORMAN guidance on suspect and non-target screening in environmental monitoring," *Environ. Sci. Eur.*, vol. 35, no. 1, Dec. 2023, doi: 10.1186/s12302-023-00779-4.

[24] J. S. Casey, S. R. Jackson, J. Ryan, and S. R. Newton, "The use of gas chromatography – high resolution mass spectrometry for suspect screening and non-targeted analysis of per- and polyfluoroalkyl substances," *J. Chromatogr. A*, vol. 1693, Mar. 2023, doi: 10.1016/j.chroma.2023.463884.

[25] "https://nontargetedanalysis.org/reference-content/methods/study-design/."

[26] F. Stilo, C. Bicchi, A. M. Jimenez-Carvelo, L. Cuadros-Rodriguez, S. E. Reichenbach, and C. Cordero, "Chromatographic fingerprinting by comprehensive two-dimensional chromatography: Fundamentals and tools," *TrAC - Trends in Analytical Chemistry*, vol. 134, Jan. 2021, doi: 10.1016/j.trac.2020.116133.

[27] Y. F. Wong, S. T. Chin, P. Perlmutter, and P. J. Marriott, "Evaluation of comprehensive two-dimensional gas chromatography with accurate mass time-of-flight mass spectrometry for the metabolic profiling of plant-fungus interaction in Aquilaria malaccensis," *J. Chromatogr. A*, vol. 1387, Mar. 2015, doi: 10.1016/j.chroma.2015.01.096.

[28] S. E. Reichenbach, X. Tian, C. Cordero, and Q. Tao, "Features for non-targeted cross-sample analysis with comprehensive two-dimensional chromatography," *J. Chromatogr. A*, vol. 1226, Feb. 2012, doi: 10.1016/j.chroma.2011.07.046.

[29] A. Caratti *et al.*, "Translating 2D-Chromatographic Fingerprinting to Quantitative Volatilomics: Unrevealing Compositional Changes in Maize Silage Volatilome for

Robust Marker Discovery," *J. Agric. Food Chem.*, vol. 72, no. 42, Oct. 2024, doi: 10.1021/acs.jafc.4c05877.

[30] P. De Coning and J. Swinley, *A practical guide to gas analysis by gas chromatography*. Elsevier, 2019. ISBN: 978-0-12-818889-7

[31] K. Dettmer-Wilde and W. Engewald, *Practical Gas Chromatography*. Elsevier, 2014. doi: 10.1007/978-3-642-54640-2.

[32] L. Mondello, A. C. Lewis, and K. D. Bartle, *Multidimensional chromatography*. Wiley, 2002.

[33] J. C. Giddings, "Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation," *J. Chromatogr. A*, vol. 703, 1995, doi: 10.1016/0021-9673(95)00249-M.

[34] D. Zanella, J. F. Focant, and F. A. Franchina, "30th Anniversary of comprehensive two-dimensional gas chromatography: Latest advances," *Analytical Science Advances*, vol. 2, no. 3–4, Apr. 2021, doi: 10.1002/ansa.202000142.

[35] P. Q. Tranchida and L. Mondello, *Hyphenations of capillary chromatography with mass spectrometry*. Elsevier, 2020. ISBN: 978-0-12-809638-3

[36] A. Kaufmann, "The current role of high-resolution mass spectrometry in food analysis," *Anal. Bioanal. Chem.*, vol. 403, no. 5, pp. 1233–1249, May 2012, doi: 10.1007/s00216-011-5629-4.

[37] E. L. Schymanski *et al.*, "Identifying small molecules via high resolution mass spectrometry: Communicating confidence," *Environ. Sci. Technol.*, vol. 48, no. 4, Feb. 2014, doi: 10.1021/es5002105.

[38] E. L. Schymanski *et al.*, "Non-target screening with high-resolution mass spectrometry: Critical review using a collaborative trial on water analysis," *Anal. Bioanal. Chem.*, vol. 407, no. 21, Aug. 2015, doi: 10.1007/s00216-015-8681-7.

[39] T. Kind and O. Fiehn, "Advances in structure elucidation of small molecules using mass spectrometry," *Bioanal. Rev.*, vol. 2, no. 1, pp. 23–60, 2010, doi: 10.1007/s12566-010-0015-9.

[40] O. Pani and T. Górecki, "Comprehensive two-dimensional gas chromatography (GC×GC) in environmental analysis and monitoring," *Anal. Bioanal. Chem.*, vol. 386, no. 4, 2006, doi: 10.1007/s00216-006-0568-1.

[41] B. L. Milman and I. K. Zhurkovich, "Big Data in Modern Chemical Analysis," *Journal of Analytical Chemistry*, vol. 75, no. 4, Apr. 2020, doi: 10.1134/S1061934820020124.

[42] L. Lacalle-Bergeron, D. Izquierdo-Sandoval, J. V. Sancho, F. J. López, F. Hernández, and T. Portolés, "Chromatography hyphenated to high resolution mass spectrometry in untargeted metabolomics for investigation of food (bio)markers," *TrAC - Trends in Analytical Chemistry*, vol. 135, Feb. 2021, doi: 10.1016/j.trac.2020.116161.

[43] "The NIST 20 Mass spectral library." Accessed: Nov. 12, 2025. [Online]. Available: https://www.msp.ch/wp/wp-content/uploads/2023/03/NIST_2020_MS_Libraries.pdf

[44] "Wiley Registry of Mass Spectral Data, 12th ed." Accessed: Nov. 12, 2025. [Online]. Available: http://www.sisweb.com/software/wiley-registry.htm#1

[45] "mzCloud." Accessed: Nov. 12, 2025. [Online]. Available: http://www.mzcloud.org/

[46] R. A. Hites and K. J. Jobst, "Is Nontargeted Screening Reproducible?," *Environ. Sci. Technol.*, vol. 52, no. 21, Nov. 2018, doi: 10.1021/acs.est.8b05671.

[47] J. Hollender *et al.*, "High resolution mass spectrometry-based non-target screening can support regulatory environmental monitoring and chemicals management," *Environ. Sci. Eur.*, vol. 31, no. 1, Dec. 2019, doi: 10.1186/s12302-019-0225-x.

[48] "https://www.norman-network.com/nds/."

[49] F. Stilo, C. Bicchi, A. Robbat, S. E. Reichenbach, and C. Cordero, "Untargeted approaches in food-omics: The potential of comprehensive two-dimensional gas chromatography/mass spectrometry," *TrAC - Trends in Analytical Chemistry*, vol. 135, Feb. 2021, doi: 10.1016/j.trac.2020.116162.

[50] A. Caratti *et al.*, "Boosting comprehensive two-dimensional chromatography with artificial intelligence: Application to food-omics," *TrAC - Trends in Analytical Chemistry*, vol. 174, May 2024, doi: 10.1016/j.trac.2024.117669.

[51] F. Magagna *et al.*, "Advanced fingerprinting of high-quality cocoa: Challenges in transferring methods from thermal to differential-flow modulated comprehensive two dimensional gas chromatography," *J. Chromatogr. A*, vol. 1536, pp. 122–136, Feb. 2018, doi: 10.1016/j.chroma.2017.07.014.

[52] T. J. Trinklein, C. N. Cain, G. S. Ochoa, S. Schöneich, L. Mikaliunaite, and R. E. Synovec, "Recent Advances in GC×GC and Chemometrics to Address Emerging Challenges in Nontargeted Analysis," *Anal. Chem.*, vol. 95, no. 1, Jan. 2023, doi: 10.1021/acs.analchem.2c04235.

[53] K. M. Pierce, B. Kehimkar, L. C. Marney, J. C. Hoggard, and R. E. Synovec, "Review of chemometric analysis techniques for comprehensive two dimensional separations data," *J. Chromatogr. A*, vol. 1255, Sep. 2012, doi: 10.1016/j.chroma.2012.05.050.

[54] J. C. Altamirano, S. Yin, L. Belova, G. Poma, and A. Covaci, "Exploring the hidden chemical landscape: Non-target and suspect screening analysis for investigating solid waste-associated environments," *Environ. Res.*, vol. 245, Mar. 2024, doi: 10.1016/j.envres.2023.118006.

[55] F. N. Castañeda, R. B. Pellegrino Vidal, and J. Aspromonte, "Untargeted chromatographic methods coupled with chemometric strategies for the analysis of food and related samples," *TrAC - Trends in Analytical Chemistry*, vol. 173, Apr. 2024, doi: 10.1016/j.trac.2024.117650.

[56] Ramos Lourdes, *Comprehensive Analytical Chemistry*, vol. 55. Elsevier, 2009. ISBN: 978-0-444 53237-4.

[57] N. H. Snow, *Basic multidimensional gas chromatography*. Academic Press, 2020. ISBN: 978-0-12-813746-8

[58] F. A. Franchina, L. M. Dubois, and J. F. Focant, "In-Depth Cannabis Multiclass Metabolite Profiling Using Sorptive Extraction and Multidimensional Gas Chromatography with Low- A nd High-Resolution Mass Spectrometry," *Anal. Chem.*, vol. 92, no. 15, p. 10520, Aug. 2020, doi: 10.1021/acs.analchem.0c01301.

[59] I. Lukić, S. Carlin, and U. Vrhovsek, "Comprehensive 2d gas chromatography with tof-ms detection confirms the matchless discriminatory power of monoterpenes and provides in-depth volatile profile information for highly efficient white wine varietal differentiation," *Foods*, vol. 9, no. 12, Dec. 2020, doi: 10.3390/foods9121787.

[60] S. E. Prebihalo *et al.*, "Multidimensional Gas Chromatography: Advances in Instrumentation, Chemometrics, and Applications," *Anal. Chem.*, vol. 90, no. 1, Jan. 2018, doi: 10.1021/acs.analchem.7b04226.

[61] Y. Feng, T. X. Fu, L. Zhang, C. Wang, and D. Zhang, "Research on differential metabolites in distinction of rice (Oryza sativa L.) origin based on GC-MS," *J. Chem.*, vol. 2019, 2019, doi: 10.1155/2019/1614504.

[62] F. A. Franchina, D. Zanella, L. M. Dubois, and J. F. Focant, "The role of sample preparation in multidimensional gas chromatographic separations for non-targeted analysis with the focus on recent biomedical, food, and plant applications," *J. Sep. Sci.*, vol. 44, no. 1, Jan. 2021, doi: 10.1002/jssc.202000855.

[63] B. K. Matuszewski, M. L. Constanzer, and C. M. Chavez-Eng, "Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC-MS/MS," *Anal. Chem.*, vol. 75, no. 13, Jul. 2003, doi: 10.1021/ac020361s.

[64] H. Qin *et al.*, "Recent advances in sample preparation and chromatographic analysis of pharmaceuticals and personal care products in environment," *TrAC - Trends in Analytical Chemistry*, vol. 164, Jul. 2023, doi: 10.1016/j.trac.2023.117112.

[65] J. Hollender, E. L. Schymanski, H. P. Singer, and P. L. Ferguson, "Nontarget Screening with High Resolution Mass Spectrometry in the Environment: Ready to Go?," *Environ. Sci. Technol.*, vol. 51, no. 20, pp. 11505–11512, Oct. 2017, doi: 10.1021/acs.est.7b02184.

[66] D. M. Mazur, E. A. Detenchuk, A. A. Sosnova, V. B. Artaev, and A. T. Lebedev, "GC-HRMS with complimentary ionization techniques for target and non-target screening for chemical exposure: Expanding the insights of the air pollution markers in Moscow snow," *Science of the Total Environment*, vol. 761, Mar. 2021, doi: 10.1016/j.scitotenv.2020.144506.

[67] P. H. Stefanuto, A. Smolinska, and J. F. Focant, "Advanced chemometric and data handling tools for GC×GC-TOF-MS: Application of chemometrics and related advanced data handling in chemical separations," *TrAC - Trends in Analytical Chemistry*, vol. 139, Jun. 2021, doi: 10.1016/j.trac.2021.116251.

[68] M. Gaida, C. N. Cain, R. E. Synovec, J. F. Focant, and P. H. Stefanuto, "Tile-Based Random Forest Analysis for Analyte Discovery in Balanced and Unbalanced GC × GC-TOFMS Data Sets," *Anal. Chem.*, vol. 95, no. 36, Sep. 2023, doi: 10.1021/acs.analchem.3c01872.

This page was intentionally left blank

Chapter 3

Fundamentals of comprehensive two-dimensional gas chromatography

3.1. The concept of multidimensionality

The chemical space we are enclosed in is characterized by an enormous degree of heterogeneity. Mixtures, such as natural olive oil, are composed of a relatively small number of constituents [1], whereas others, such as fuels, are instead highly complex, including hundreds to thousands of compounds covering multiple chemical classes and wide ranges of polarity and volatility [2]. Chromatography is, without doubt, the technique of choice for the separation of a real-world sample; however, due to this heterogeneity, separating compounds of interest from a complex sample could be difficult with a single chromatographic run, leading to the adoption over the years of a multidimensional approach [3].

The first multidimensional separation was performed by Consden *et al.* in 1944 to separate amino acids on cellulose [4]. As the authors mentioned, various solvents were tested, and the arrangement of amino acids in the chromatogram depends on the solvent type. In each analysis, two solvents were used, and the combination of them to be employed was chosen based on retention factor (RF) values (the ratio between the movement of the band and the movement of the solvent front). In this way, "*by development first in one direction with one solvent followed by development in a direction at right angles with another solvent, amino acids (e.g. a drop of protein hydrolysate) placed near the corner of a sheet of paper become distributed in a pattern across the sheet to give a two-dimensional chromatogram characteristic of the pair of solvents used*". Employing this approach with an appropriate combination of solvents significantly enhances the occupation of the two-dimensional separation space. In this example, the use of collidine (a basic organic solvent composed of trimethylpyridine isomers) in the first dimension and a phenol-ammonia solution in the second dimension enables effective separation of amino acid mixtures by maximizing orthogonality between the two dimensions. This strategy is just one of the many combinations of various separation mechanisms that can be used to create a multidimensional system [5], and the result of this first multidimensional separation is presented in Figure 3.1.

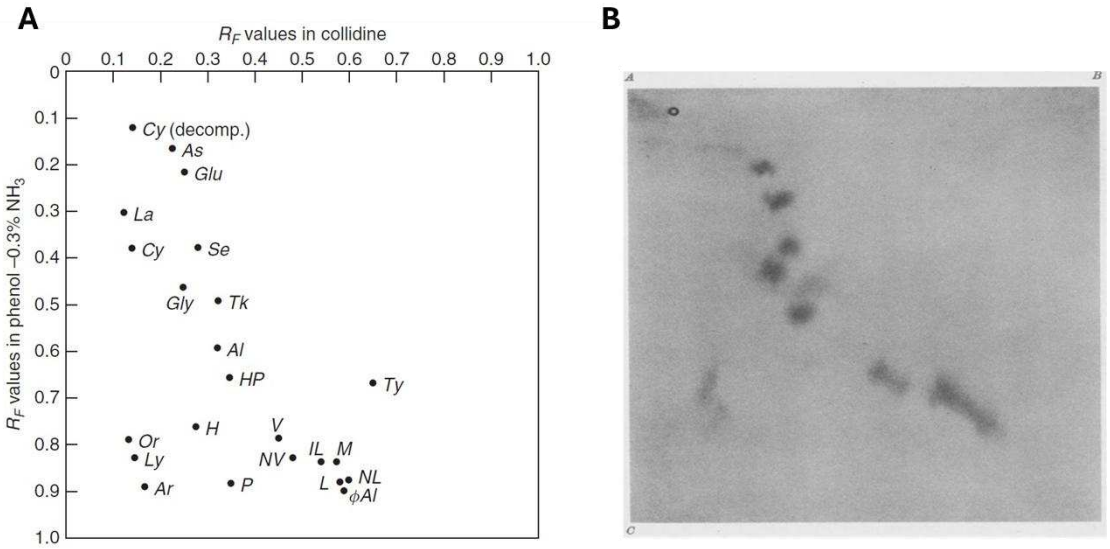


Figure 3.1. *A) Expected positions of 22 amino acids on a two-dimensional plane. B) Two-dimensional separation of 22 amino acids on paper, carried out using collidine in the first dimension (upper side of the paper) and a phenol-ammonia mixture for the second separation (vertical side of the paper) [4].*

Consden *et al.* identified this separation mechanism in 1944, long before the comprehensive GC×GC system proposed by Liu and Philips in 1991 [6]. In fact, throughout the decades that followed, several researchers expanded the concept of multidimensionality [7, 8].

A truly multidimensional separation system is not simply the result of connecting two or more columns in series, but rather the consequence of two fundamental requirements. These conditions were formulated by Giddings in 1987 [9], and can be summarized as follows:

- Sample components of a sample must undergo two or more largely independent separative displacements.

By this first point, Giddings emphasizes that the separation mechanism in each dimension must rely on different physical and chemical properties, such as volatility and polarity [10]. In this way, if the migration of an analyte in one dimension is uncorrelated with its movement in the second, two components that may co-elute in the first dimension probably get separated in the second. This fact essentially corresponds to the dramatic increase in peak capacity (see Section 3.2) usually observed in multidimensional systems.

- The resolution achieved in one dimension must not be lost (or decreased) in the subsequent dimension.

In this case, to maintain the separation already achieved in the first dimension and thus satisfy the second condition, only a small, well-defined fraction of effluent from the first column is transferred to the second. Essentially, this prevents remixing of the peaks, as each fraction is transferred as an independent subsample, maintaining the temporal separation previously achieved.

Although a certain degree of correlation is always present, when the separation dimensions are based on different interaction mechanisms, the separation can be considered "orthogonal" [11]. There are several ways to represent orthogonality in multidimensional separation [12], but the fundamental concept is that the correlation between the dimensions will produce redundant information, where a separation effort with one mechanism corresponds, for every analyte, to a predictable effort of separation with the second one.

In Figure 3.2, three different degrees of correlation in a two-dimensional separation system are presented:

A. Orthogonal separation: the peaks are distributed throughout the plane.

The two dimensions are entirely uncorrelated: the retention behavior of the analytes in one dimension is independent of the other, resulting in a spread distribution of points across the 2D plane. This represents the ideal scenario, in terms of peak capacity, for multidimensional separations, as orthogonality maximizes the accessible separation.

B. Separation with partial correlation: the peaks are centered along the diagonal.

Here, analytes show some degree of related behavior across the two separation mechanisms, leading to clustering within a restricted angular region. Although some orthogonality is retained, the effective separation space is reduced, and consequently, the gain in peak capacity is less pronounced than in the fully uncorrelated case.

C. Separation with total correlation: the peaks are distributed along the diagonal, resulting in the equivalent 1D separation.

The analytes align along a narrow diagonal trend, indicating that their migration in the second dimension is largely predictable from the first. This high degree of correlation ruins the effective separation space, pushing it back towards one dimension, providing no advantage over a conventional one-dimensional system. Nevertheless, under these conditions, processes such as thermal modulation may still provide analytical benefits, including peak focusing and enhanced signal-to-noise ratio.

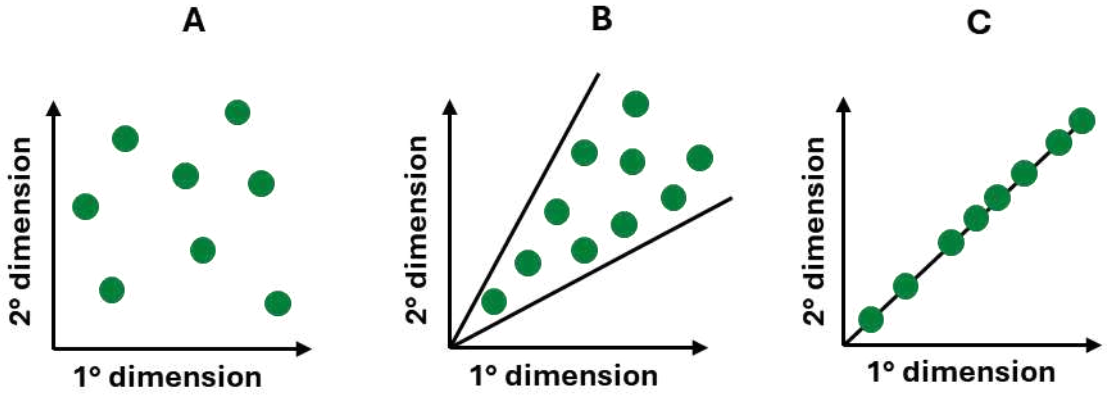


Figure 3.2. *Representation of three different degrees of correlation between two separation dimensions. A) Orthogonal separation. B) Separation with correlation. C) Separation with total correlation. Adapted from [13].*

Selecting the correct separation dimensions is crucial for creating efficient multidimensional systems [14], as a rational choice is vital to achieving well-organized distributions of compounds, which directly increases the information obtained. This requirement is directly linked to the concept of "sample dimensionality" introduced by Giddings in 1995 [15], which defines the number of independent variables characterizing the physicochemical diversity of a mixture. Such variables may include π-π aromatic interactions, hydrogen-bonding capacity, functional groups, volatility or carbon number. These molecular descriptors constitute the pool of properties to be introduced as complementary separation axes for creating an orthogonal separation system.

3.2. Why 2D-GC?

The two fundamental aspects that govern all chromatography processes are [3]:

- peak capacity (n_c)
- stationary-phase selectivity

The former parameter can be mathematically expressed in different ways. One possible description is given by Eq. 3.1 [16]:

$$n_C = 1 + \frac{\sqrt{N}}{4R_s} ln \frac{t_{R2}}{t_{R1}} \qquad \text{Eq. 3.1}$$

In Eq. 3.1, N represents the theoretical plate number, R_s is the desired resolution, and t_{R2} and t_{R1} are the maximum and minimum retention times, respectively, at which peaks may elute.

Peak capacity can be understood as the maximum number of chromatographic peaks that can theoretically be accommodated within the time window of a separation. For example, if each peak has a base width (W_b) of 30 seconds and the total chromatographic run lasts 30 minutes, then approximately 60 peaks could fit within that space, giving a peak capacity of exactly 60 units [10]. A conceptual visualization of how peak capacity changes the pattern of a chromatogram is given in Figure 3.3[16].

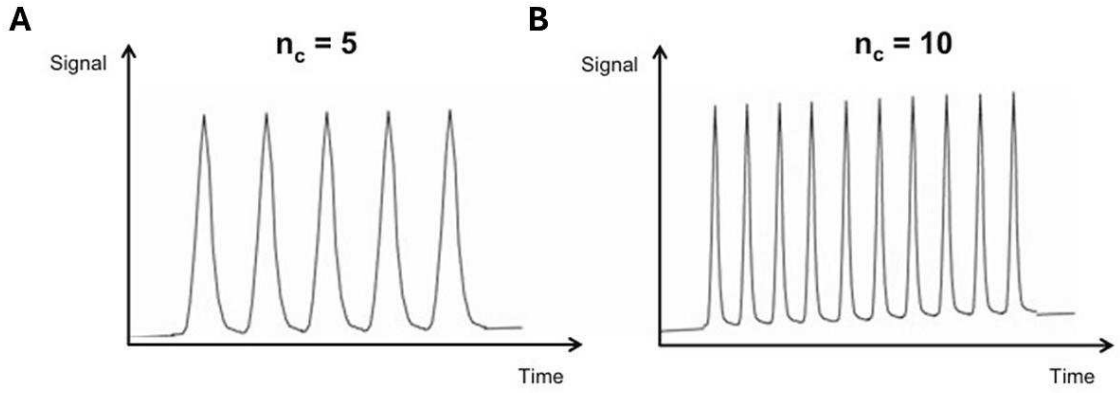


Figure 3.3. *Graphical representation of the variations of the peak capacity in a chromatogram. A) Peak capacity of 5 (n_c=5). Exactly five peaks are distributed in the chromatogram. B) Peak capacity of 10 (n_c=10). Increasing the peak capacity to a value of 10, ten peaks are accommodated in the same chromatographic region [16].*

Peak capacity (n_c) depends both on the intrinsic characteristics of the column, such as its length, internal diameter, and on the other hand, on the experimental conditions applied, including flow rate, temperature, and outlet pressure. Together, these factors determine how narrow the peaks can be and how many of them can be arranged side by side without overlap, thus defining the practical limits of separation in a one-dimensional chromatographic system [3, 17].

The advantages of a multidimensional separation system become evident when examining each parameter in Eq. 3.1. The first thing to consider is that peak capacity (n_c) is significantly increased from the effective column efficiency (N), as in the case of connecting two columns, the column length is extended to increase the plate number. Second, the required resolution

(R_s) also favors the use of orthogonal separation mechanisms, as compounds co-eluting in one dimension are likely to be separated in the other, thereby reducing the minimum resolution required between adjacent peaks and effectively increasing the number of recognizable components. Finally, the last logarithmic term (ln t_{R2}/t_{R1}) represents the separation window, which is significantly larger than in a dimensional separation space, as peaks can span a much wider range of retention behaviors than any single dimension alone can offer [17].

The other parameter mentioned above, the stationary-phase selectivity, is a feature directly related to the chemical composition of the stationary phase and, hence, to the specific type of analyte-stationary phase interactions (*e.g.*, dispersion, dipole-dipole, and hydrogen bonding). Selectivity, however, is influenced not only by analyte-stationary phase interactions but also by analyte solubility in the mobile phase whenever this interaction occurs [18]. In this sense, although the carrier gas is typically considered an inert transport medium, its physical properties, such as the viscosity, can notably modulate analyte behavior, especially in modulational processes, where pressure and pulses can influence the repartitioning equilibrium in each dimension [19].

As discussed, the use of an additional separation dimension provides much greater insight into a sample's complexity by significantly increasing peak capacity. However, even in 1987 [9], Giddings demonstrated that even a highly efficient one-dimensional chromatographic system is fundamentally limited. According to his work, as the number of analytes increases, many peaks inevitably fall close together, meaning that only a small portion of the system's theoretical peak capacity can be actually exploited to resolve individual components. In fact, the author showed that no more than about one-third of this capacity contributes to a real peak separation, and that only around 18% of the peaks observed in a typical one-dimensional chromatogram correspond to a single pure compound, while the remainder represent overlaps of co-eluting species. This limitation does not arise from poor instrumental performance, but from the intrinsic constraint that a single separation axis does not provide enough “space” to accommodate the diversity present in complex mixtures.

Considering also the familiar formula for the calculation of resolution between two adjacent compounds, is it possible to highlight the advantages of translating from one to a multidimensional separation space [20].

$$\mathrm{R}_s = \frac{\sqrt{N}}{4}\frac{\alpha - 1}{\alpha}\frac{k_2}{k_2 - 1} \qquad \text{Eq. 3.2}$$

In Eq. 3.2, N represents the plate number, α the selectivity, and k_2 the retention factor of the second analyte.

A higher resolution is well desired in all the possible dimensions of separation, and as in the case of the peak capacity, even here, an increase of the plate number (N), which corresponds to adding a subsequent dimension (as extending the column length), leads to an increase of the resolution. By hypothetically maintaining the other parameters of Eq. 3.2 constant, the resolution increases by the characteristic pattern of the square root of the plate number. This effect can be seen in Figure 3.4, where the separations of two analytes (k1 = 4.8; k2 = 5.0; α = 1.05) on a GC column, under fixed conditions, are considered [3].

In the case of the retention factor (k_2), the contribution to the resolution only appears when this parameter is very low, as can be seen for the first exponential part of $\frac{k-1}{k}$. This means

essentially that "*k"* has a great influence on the resolution of early eluting peaks. However, as indicated by the trend, after k > 5, further increases in the retention factor provide little improvement in resolution. In practice, increasing "*k*" simply shifts peaks to longer retention times.

Finally, the most significant enhancement is given by the selectivity, which grows rapidly even for small changes in "*α*". A slight difference in how two analytes interact with the stationary phase results in a much larger separation of their retention times, which directly translates into higher resolution.

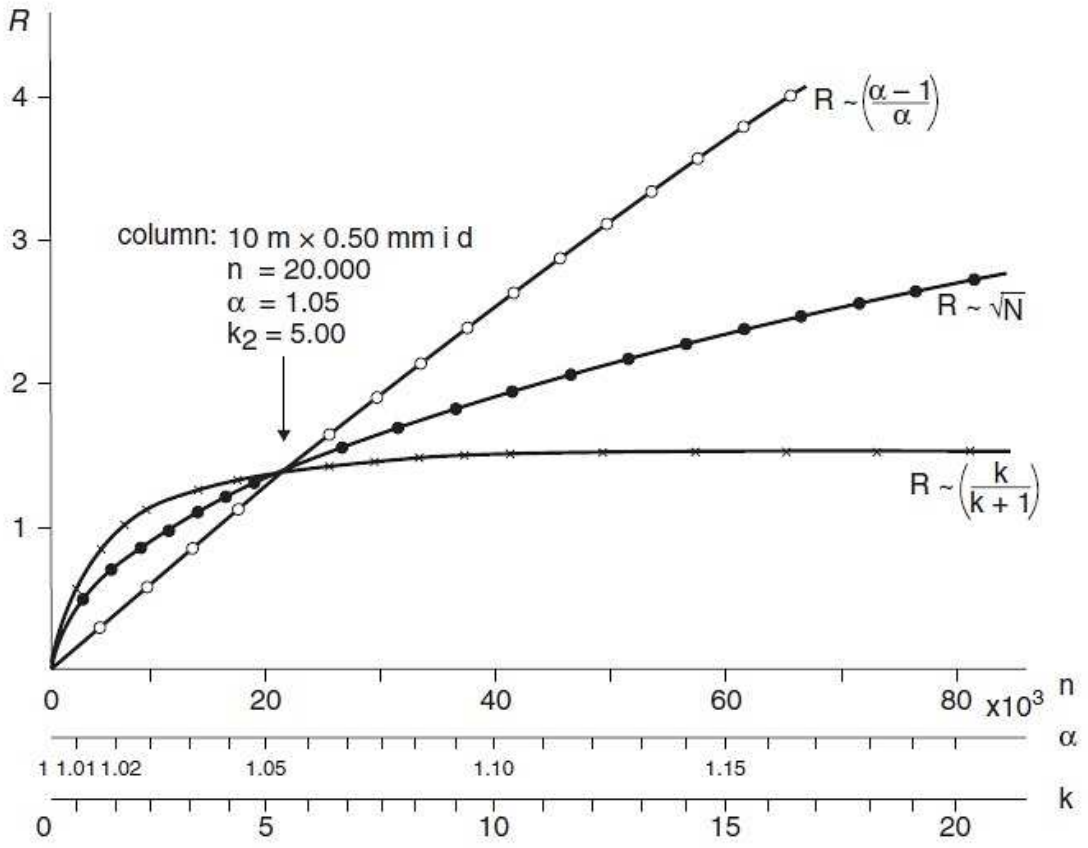


Figure 3.4. *Influence of N, k, and α on resolution. Taken from [3].*

All the parameters involved in increasing separation efficiency could be improved by simply increasing the column length, as a traditional strategy in one-dimensional gas chromatography; however, this approach is substantially constrained by the fact that analysis time increases. Consequently, meaningful gains in resolution can only be achieved at the cost of prohibitively long run times, making further extension of one-dimensional GC impractical for the analysis of complex mixtures. For this reason, the most effective way to enhance the separations is to employ a multidimensional chromatographic system [3].

In the end, in light of the discussion in this section, it becomes clear that for multiple reasons the pursuit of multidimensional separations, including the development of comprehensive two-dimensional gas chromatography (GC×GC), represents nowadays the key technique to separate complex mixtures, from driving the identification and quantification of components in various matrices (*i.e.* biologically and environmentally) to leading the generation of new informatic and chemometrics workflows [21–26].

3.2. From GC-GC to GC×GC

Multidimensional gas chromatography is typically distinguished into two different modalities [27]:

- Heart-cut GC, which is usually indicated with a dash between two "GC" (GC-GC).
- Comprehensive multidimensional GC, which is typically indicated by a times sign between the two "GC" (GC×GC).

In GC-GC, two different columns are used, but only a small portion of the effluent eluting from the first column ("heart-cut") is introduced for further separation into the second column; however, the number of heart-cuts can be increased if the time allowed for separating the cuts in the second dimension is also decreased. When the number of heart-cuts gets high enough (and the time for their separation too), a comprehensive separation can be achieved, in which the entire sample is subjected to separation in both dimensions (Figure 3.5) [28]. Essentially, GC×GC can be considered an advancement of traditional heart-cut GC [29].

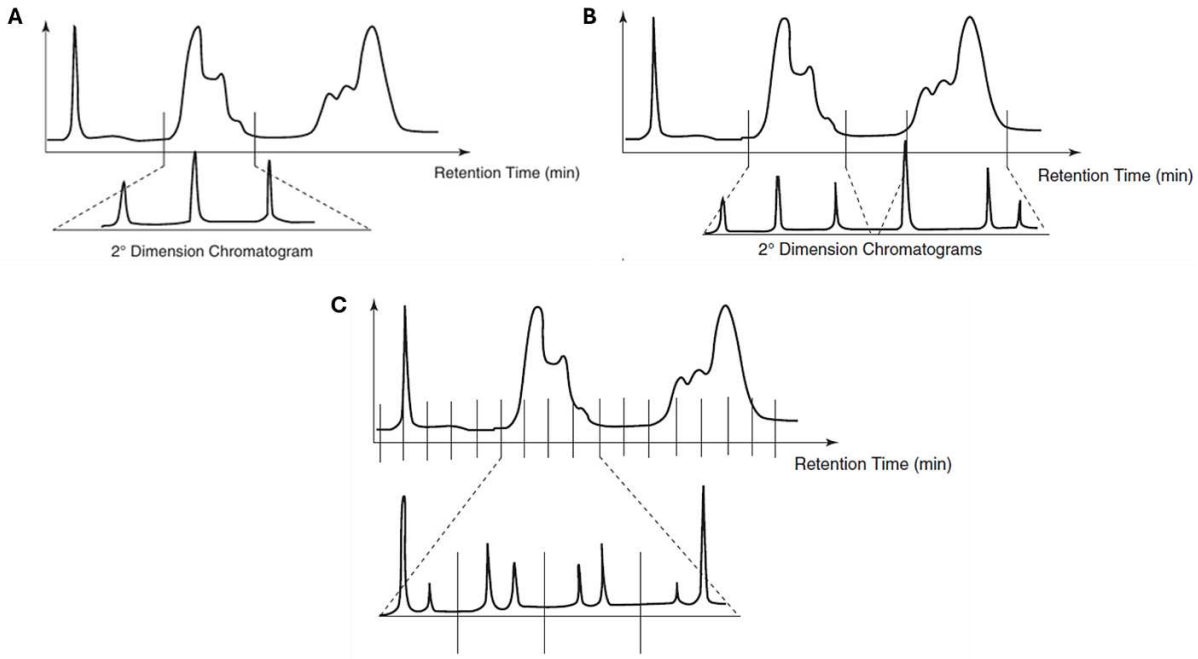


Figure 3.5. *The concept of multidimensional GC. A) Single heart-cut GC analysis; a significant portion of the effluent from the primary column with coelutions is diverted to the second dimension. B) Dual heart-cut GC analysis; two regions with coelutions are diverted to the second dimension column. C) Comprehensive two-dimensional GC analysis; the size of the sequential heart-cuts and the second dimension chromatograms are very short, enabling a comprehensive separation of the sample components [28].*

Considering the peak capacity for the process involved in Figure 3.5A, using a single heart-cut of about one peak width, the theoretical peak capacity of the system ($n_{GC\text{-}GC}$) equals the sum of the peak capacities of both dimensions [30], as indicated in Eq. 3.3.

$$n_{GC-GC} = n_1 + n_2 \qquad \text{Eq. 3.3}$$

In the second case, if multiple heart-cuts are generated, the peak capacity of the second dimension (n_2) should be multiplied by the number (x) of cuts, and consequently, the Eq. 3.4 becomes:

$$n_{GC-GC} = n_1 + (x \times n_2) \qquad \text{Eq. 3.4}$$

This expression represents an ideal mathematical description of this multidimensional approach, as it assumes perfect transfer between the separation domains. In practice, deviations are expected, as this approach imposes a significant operational burden on the operator. Essentially, heart-cut separations require manual or semi-manual transfer of selected fractions between two (or more) instruments, based on different separation mechanisms, or within the same instrument, at the cost of completing the first analytical run, changing the separation mechanism, reconditioning the device, and reinjecting each sample fraction one by one. In fact, heart-cut multidimensional separations are traditionally more widely implemented in liquid chromatography, where fraction collection and selective transfer of eluate are technically easier to achieve and well supported by commercial instrumentation [31].

The very first heart-cutting multidimensional separation was carried out in 1958 by Simmons *et al.* [32]. Through the use of an intricated mechanism that resembles an archaeological manufactory to the view of an operator used to modern separation systems (see picture in [32]), the author achieved "*separations that were not normally possible with conventional single columns or multiple columns connected in series arrangements*". This old first heart-cut separation, however, is based on principles that are still valid today. In particular, Simmons *et al.* described that by combining a first dimension non-polar column (reflecting boiling point separation) with a second dimension polar column, and by selectively transferring only narrow "cuts" of interest, they were able to avoid the remixing effects that normally occurred when two columns of different selectivity were directly connected in series. The results showed a dramatic improvement in resolving power and allowed high-quality separations of key analytes in hydrocarbon fractions C_6, C_7, and C_8 from catalytic platformer streams Figure 3.6 [32].

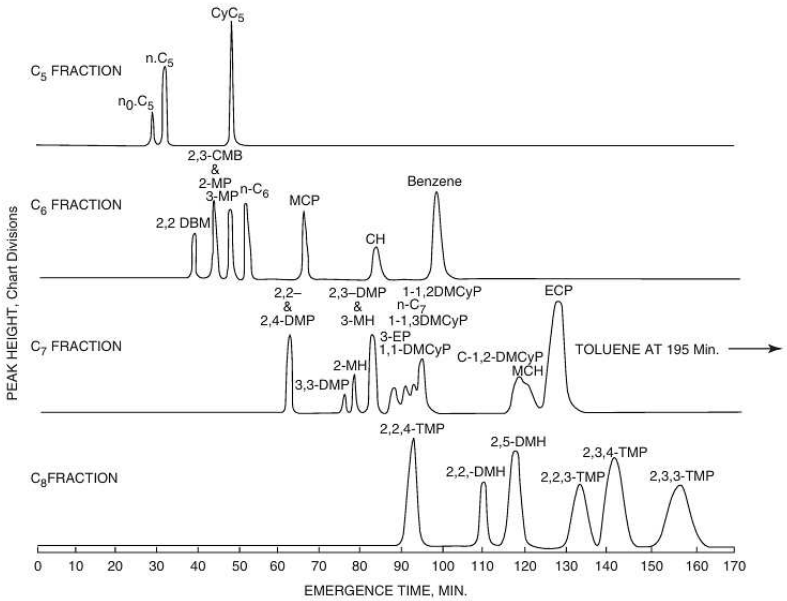


Figure 3.6. *Second dimension GC separations of four (C_5 , C_6, C_7, C_8) hydrocarbon fractions [32].*

Heart-cut multidimensional separations offer several advantages; for example, they enable highly targeted analysis of specific regions of the chromatogram, enhance selectivity for

difficult analyte pairs, and reduce data complexity compared to fully comprehensive techniques [33]. However, these benefits also come with some limitations. Since only the selected portion of the first chromatogram is forwarded to subsequent separations, the peak capacity gain is lower (as evaluated in the previous section), and coelutions outside the heart-cut region remain unresolved. Then, the method's success depends heavily on the analyst's ability to decide which fractions transfer, making the approach less suitable for global profiling (see Chapter 2, Section 2.2.4). For these reasons, while heart-cut remains a valuable tool for some target separations, comprehensive multidimensional techniques such as GC×GC (and LC×LC) provide far greater resolving power for complex mixtures [31].

In comprehensive two-dimensional gas chromatography, the entire sample loaded through the injector is subjected to the two chromatographic separations by means of a transfer device (the modulator), which sequentially transfers "heart-cuts" from the first to the second column. This is possible because the sample transfer in the second dimension is conducted in very small fractions, and the second column is shorter and frequently smaller in diameter than the main one [34].

The peak capacity in GC×GC is given by the product of the peak capacities of each column, as can be seen in Eq. 3.5.

$$n_{GC\times GC} = n_1 \times n_2 \qquad \text{Eq. 3.5}$$

This increase in separation power is graphically illustrated in Figure 3.7.

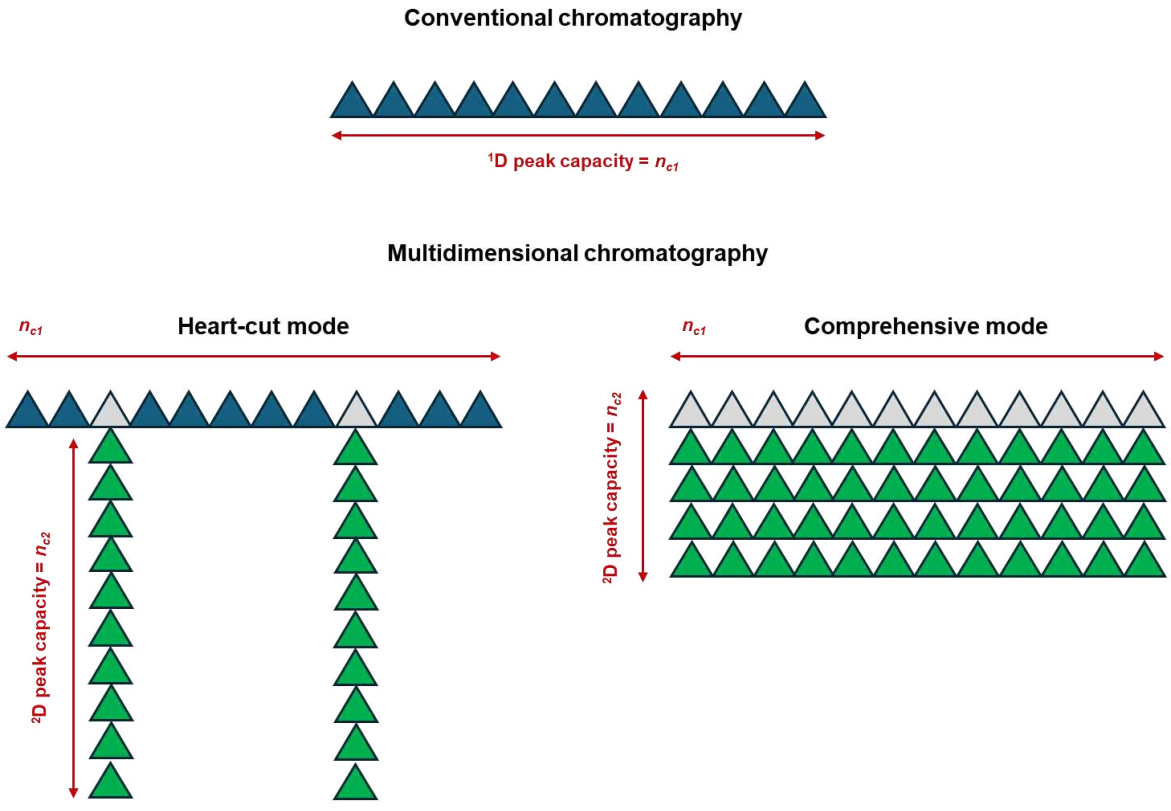


Figure 3.7. *Comparison between the peak capacity of a one-dimensional and two-dimensional separations (heart-cut and comprehensive). Adapted from [34].*

From an operational point of view, obtaining experimental peak capacity values close to those predicted by the theory is not straightforward. As highlighted by Blumberg *et al.*, some discrepancies prevent the theoretical assumption from being fulfilled. This discrepancy arises, for example, because the theoretical model assumes a uniform peak distribution

across two-dimensional space, whereas experimental chromatograms frequently exhibit an oriented structure, which dramatically limits the usable separation area [35]. Then, also some instrumental limitations prevent the full realization of the theoretical peak capacity, such as the injection pulses (the actual duration of reinjection into the second column), which, if not narrow, tend to broaden the second dimension peaks and significantly reduce the effective separation space [35, 36]. However, even in the light of this, according to Blumberg *et al.*, "*the peak capacity of optimal GC×GC can be more than an order of magnitude larger than that of optimal 1D-GC*" [35].

Regarding gas chromatography, the first fully comprehensive two-dimensional gas separations were obtained by Liu and Phillips in 1991 [6, 37]. In contrast to the earlier heart-cut multidimensional approaches, the two authors were capable of continuously collecting and reinjecting narrow pulses of effluent from a first column into a second shorter column by an "*on-column thermal desorption modulator interface*". This important innovation enabled the generation of a complete "*high-speed*" chromatogram for each slice of the primary separation, ensuring that every compound from a fourteen-component hydrocarbon mixture passed through both columns. Using this orthogonal system, they achieved clear separation of the analytes, as shown in Figure 3.8, where components are visualized by their characteristic circular peak shapes, as modern chromatographers are accustomed to seeing. This work laid the technological foundation for all modern GC×GC systems and remains one of the most important contributions in the history of multidimensional chromatography.

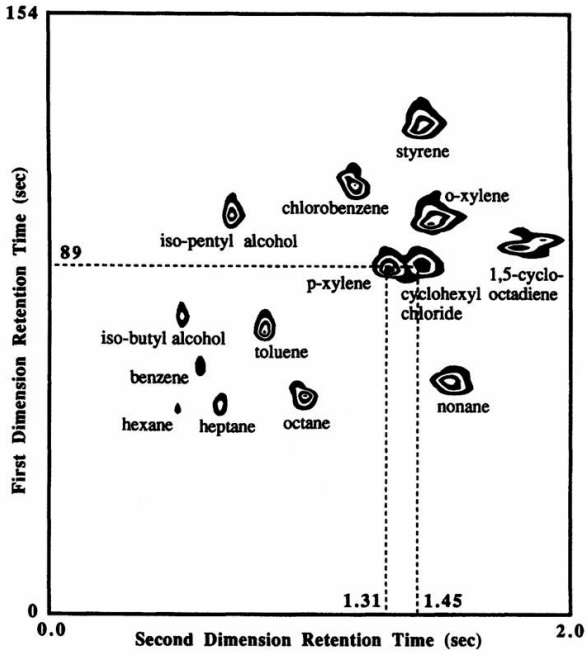


Figure 3.8. *The two-dimensional chromatogram of the mixture of fourteen components. Second dimension separation is presented on the x-axis, while first dimension separation is on the y-axis [6].*

The GC×GC system developed by Liu and Phillips employed an on-column thermal modulator based on electrical current pulses, in which a short segment of fused-silica column was rapidly heated and cooled to trap and release analytes in a periodic way. Although much more rudimentary than modern cryogenic modulators, this device operated according to the same fundamental principle of thermal trapping followed by rapid desorption used nowadays [38].

To fully understand the mechanism behind this type of separation, the by now didactic way to explain it is to consider sample analytes as features of different shapes, colors, and sizes,

as indicated in Figure 3.9. Arguably, those geometrical shapes represent three distinct sample dimensionality types. As indicated by Giddings [15], the usefulness of a comprehensive system for the separation of a complex mixture is given essentially when the dimensionality of the separation system matches the dimensionality of the sample. In Figure 3.9, when only one dimension is exploited, the result corresponds to a limited resolving power and extensive peak overlapping, where essentially the analyst can select among separations based only on size, color, or shape, and accept losing the separation achieved by choosing one over the others. In contrast, comprehensive two-dimensional separations combine two orthogonal mechanisms, such as size and color, or size and shape, generating a two-dimensional distribution of components. This multidimensional space provides a much more ordered and dispersed arrangement of the sample constituents, enabling the resolution of compounds that would otherwise co-elute in any single dimension [39].

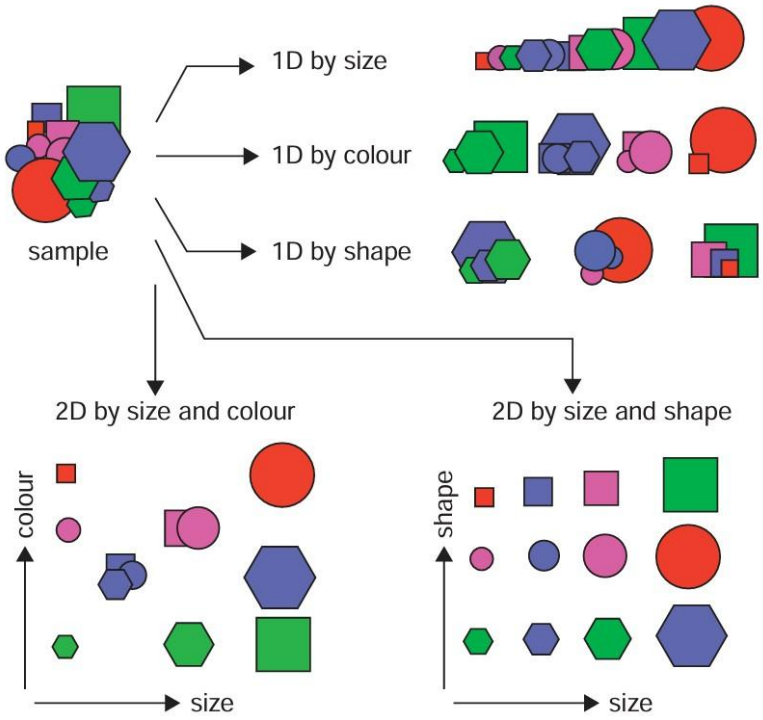


Figure 3.9. *Match between separator and sample dimensionality in GC×GC [39].*

In a GC×GC experiment, the process begins with the introduction of the sample into the injector, where it mixes with the carrier gas and undergoes the first separation step, exactly as in conventional one-dimensional GC. Instead of flowing directly to the detector, however, the effluent exiting the first column is directed to an intermediate device, the modulator, positioned between the two columns. This component is the core of the GC×GC system, as it repeatedly traps and releases small, well-defined portions of the first dimension effluent, transferring them as focused pulses into the second column [40]. Functioning essentially as an automated online injector, the modulator divides the entire first chromatogram into short segments according to a fixed modulation period (P_M), typically on the order of a few seconds. The detector then records a succession of short secondary chromatograms, each corresponding to one modulation cycle and having a duration equal to the modulation period [41]. This schematic process is illustrated in Figure 3.10.

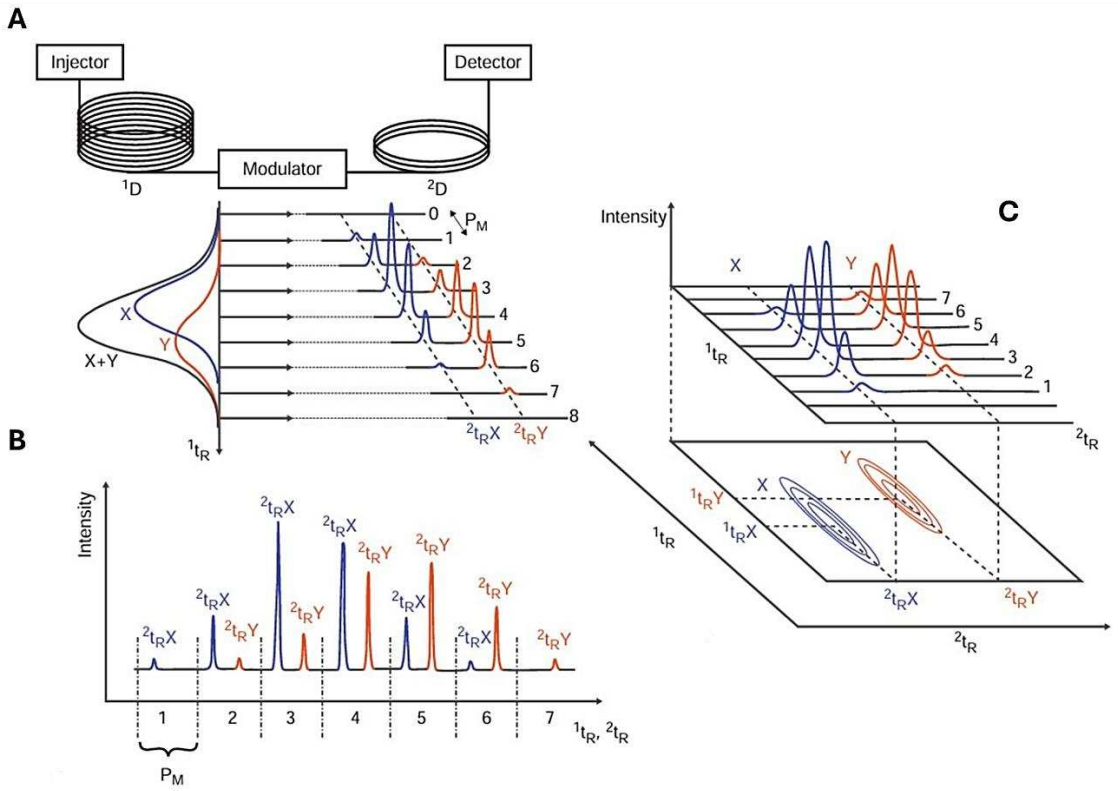


Figure 3.10. *Scheme of column coupling in a GC×GC system and of how data are handled. A) Two overlapping compounds (X and Y) coming out of the first dimension at a defined retention time (^1t_R). As the modulation process occurs during a defined P_M, narrow bands of analytes enter in the second dimension and appear to have different second dimension retention times (^2t_RX and ^2t_RY). B) Raw data signal as recorded by the detector through the entire separation process. C) Construction of the two-dimensional contour plot from the collected high-speed secondary chromatograms of in which similar signal intensities are connected by a contour line [39].*

As also indicated in Figure 3.10, the second dimension column in a GC×GC system is intentionally much shorter than the first. This design arises from the requirement that each reinjected modulation slice must be fully separated in the second dimension before the next modulation event occurs. If this does not happen, peaks originating from different cycles would overlap, leading to the so-called "wraparound" effect. Consequently, the second separation is performed at extremely high speed, typically around two orders of magnitude faster than the first dimension. Because these rapid separations happen while the first process is still running, the overall analysis time of a GC×GC experiment remains comparable to that of a conventional one-dimensional gas chromatographic (1D-GC) run [42, 43].

3.3. Advantages of GC×GC

The ability to provide a substantially larger peak capacity and to increase the selectivity of the separations are not the only advantages of GC×GC over 1D-GC; in fact, other invaluable features of this technique are listed as follows:

- Speed

In this sense, GC×GC increases the "speed" of analysis not by shortening the chromatographic runtime, but instead by maximizing the number of separated components accessible within a fixed retention window. Basically, by slicing the first dimension effluent and subjecting each portion to a fast second dimension separation, GC×GC produces a much greater number of resolved peaks per unit of time [44].

- Sensitivity

By distributing the sample components across two retention spaces, GC×GC removes much of the coeluting background that, in 1D-GC, artificially elevates the baseline and masks low-abundance analytes. As a result, the effective signal becomes less affected by the chemical interferences. Moreover, during the modulation, band compression further enhances the sensitivity of the analytes, as each modulation event focuses the first dimension band into a narrow reinjection pulse, thus improving the signal-to-noise ratio (S/N). This dual effect explains why GC×GC often exhibits detection limits up to 20 higher than those encountered with conventional GC [45–49]. In Figure 3.11, the process of amplifying the sensitivity is shown.

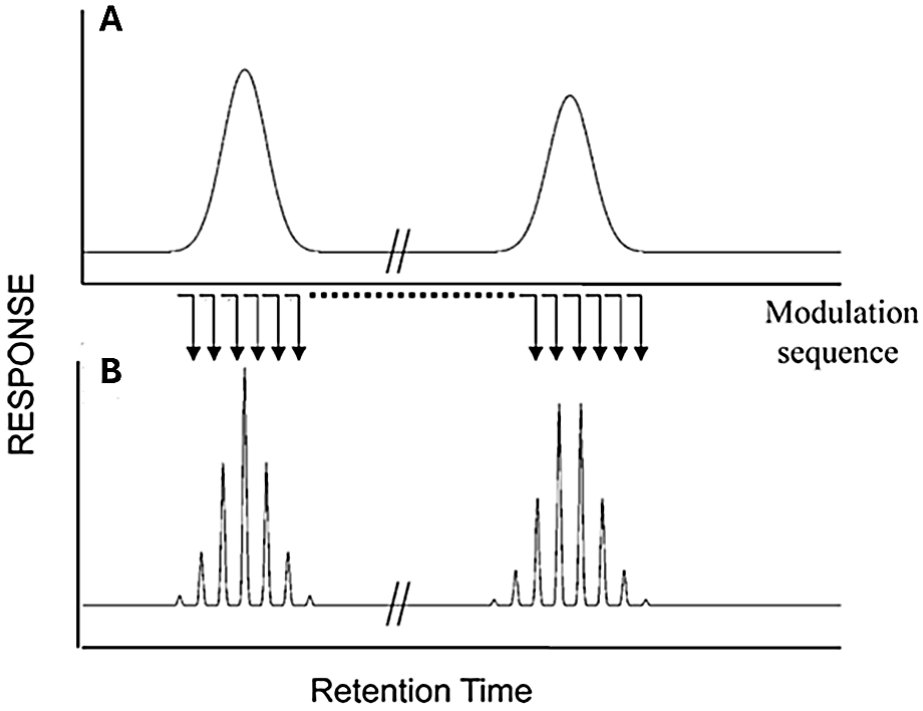


Figure 3.11. *Effect of the modulation on the signal-to-noise ratio. A) First dimension elution profiles of two analytes in the chromatographic plane. B) Corresponding modulated peaks obtained after modulation. Each 1D band is sliced into a series of narrow focused pulses, resulting in much taller peaks and consequently increasing the signal-to-noise ratio. Adapted from [44].*

- Spatial order

In contrast to 1D-GC, where compounds elute along a single axis and co-eluting phenomena are often common, GC×GC distributes components across a two-dimensional plane, producing characteristic patterns. In this way, homologous series tend to align into bands and trends that reflect polarity properties. This spatial organization greatly facilitates qualitative interpretation and class identification, enabling a rapid "visual" discrimination of compound families, especially in complex mixtures [50–54]. An example of this spatial order is given in Figure 3.12 by the work of Vendeuvre *et al.*, in their effort to characterize complex petroleum products [52].

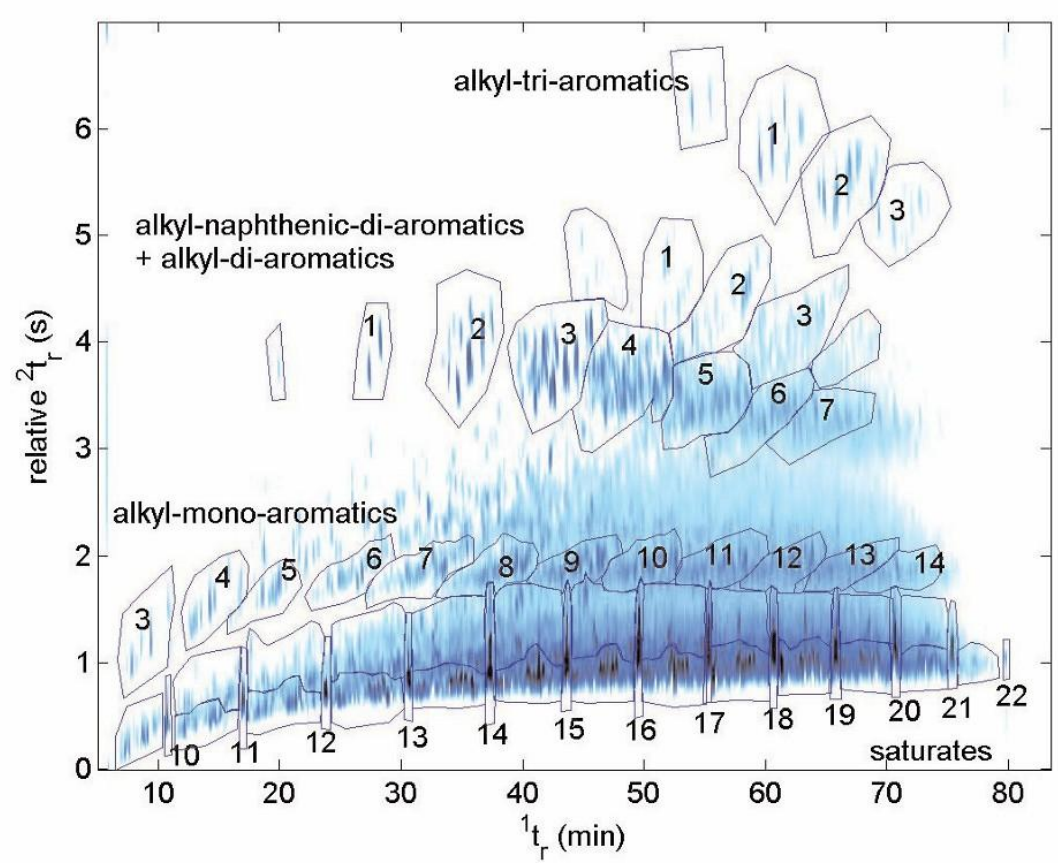


Figure 3.12. *Highly structured patterns of a diesel two-dimensional chromatogram. Analytes are arranged according to two orthogonal retention mechanisms, and distinct chemical families occupy different regions of the chromatographic plane. Saturated hydrocarbons cluster at the lower right of the plane, followed by progressively more aromatic classes: alkyl-mono-aromatics, alkyl-di- and naphthenic aromatics, and finally alkyl-tri-aromatics positioned at higher second dimension retention times [52].*

3.4. Basic instrumentation

Except for a few components, such as modulators and other temperature-control components, a GC×GC system is characterized by the same apparatus found in standard 1D-GC systems. This aspect can also be considered as another advantage of GC×GC, since from an external perspective, without opening the GC oven, 2D-GC resembles a conventional one-dimensional gas chromatograph. In other words, despite the increased complexity discussed in this section, GC×GC does not seem to be a completely different instrument for a non-expert operator, but rather an extension of the familiar 1D-GC platform.

The first part of the GC×GC is the injector, where samples are introduced; common options include split/splitless (SSL), on-column injection, and programmed-temperature vaporization injector (PTV). Then, the capillary column is connected to this unit and placed into the oven, where the first separation occurs. Moreover, at the end of the column, the effluent is trapped and reinjected through the modulator into the second capillary column, which is coated with a different stationary phase for further separation. At the end of the two chromatographic separations, the detector [55].

The data collected by the detector consist of a series of high-speed second dimension chromatograms, which are typically arranged adjacent to each other to reconstruct the two-dimensional chromatogram. Following additional processing, the peaks are visualized on a two-dimensional plot, with colors or contour lines indicating signal intensity. In Figure 3.13, the process involved in the formation of the 2D plot is presented [56].

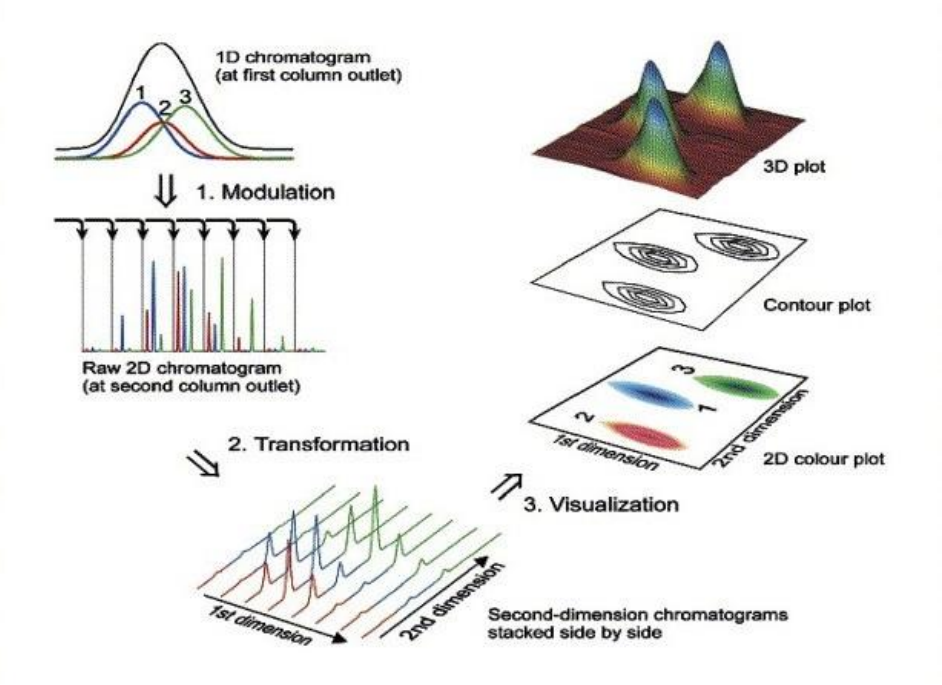


Figure 3.13. *Data flow in a comprehensive two-dimensional gas chromatography (GC×GC) analysis, from modulation to visualization. During modulation (Step 1), the effluent from the first column is periodically trapped and released in a series of narrow pulses, generating a raw second-dimension chromatogram. In the transformation stage (Step 2), these second-dimensional slices are arranged by modulation order, producing a structured two-dimensional array whose axes correspond to the retention times of the first and second dimensions. Finally, in the visualization stage (Step 3), the data can be arranged in different formats, including contour maps or 3D surface plots, allowing the analyst to observe the spatial organization characteristic of GC×GC separations [56].*

3.4.1. Injectors and injection systems

The sample introduction in the GC system is an essential step of the analytical process. Especially in quantitative analysis, correct and reproducible sample transfer is mandatory, as it often represents the greatest source of variability [57].

A. Split/Splitless injector

The split/splitless injector is a hot vaporizing injector and represents, without doubt, the most popular GC injector currently in use due to its ruggedness and ease of use [58]. This device consists of a metal chamber maintained at a constant temperature by a surrounding heater block. Inside this heated section, a glass or quartz liner is positioned and sealed around its outer edge to ensure the gas flows exclusively through the center of the liner. Then, the carrier gas enters the liner from the top and exits into the column, with the option of a split vent at the bottom, activated by a solenoid valve. Moreover, a septum seals the top of the injector, allowing a syringe to inject samples into the liner. Additionally, a septum purge line is included to prevent materials from being released from the septum or sample residue left during injection, which could lead to contamination or distorted peaks in the chromatogram [20]. Figure 3.14 shows a schematic representation of a common SSL injector.

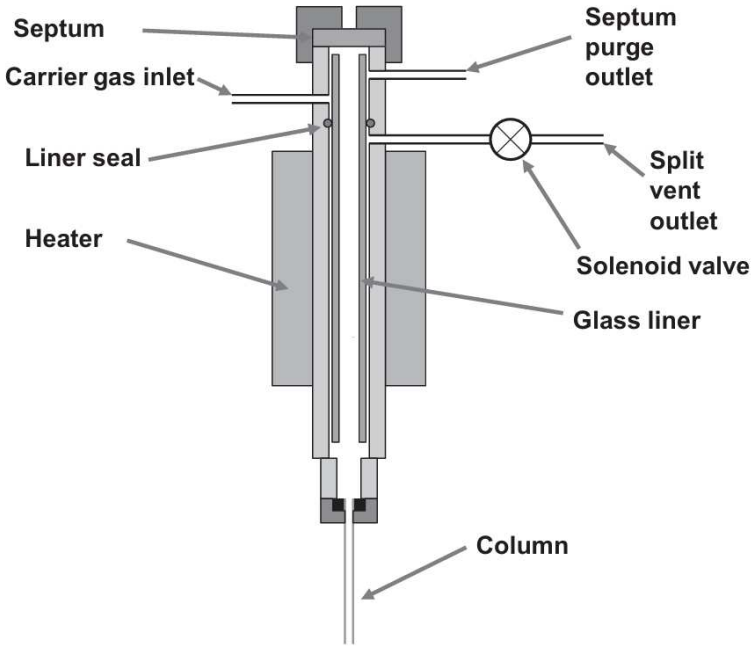


Figure 3.14. *Section of an SSL injector [39]*

As indicated by the name of this injector, it can operate in two fundamental modes: split or splitless.

In the first case, the sample is vaporized and split into two portions: a small portion is carried into the column by the carrier gas, while the majority is vented through the split vent at the bottom of the injector. Splitting is the technique of choice for highly concentrated samples, where the amount of sample entering the column can be adjusted *via* the split ratio (Eq. 3.6). With this device, samples are injected into the column over a short period, resulting in narrow, focused bands. Split injection, however, consumes a large amount of carrier gas, and a second drawback is that the splitting process sometimes discriminates against high molecular weight solutes. This leads to components entering the column that are not representative of the entire sample injected. In light of this, modern chromatographs offer the possibility to use a "gas saver" procedure by reducing carrier gas flow rate shortly after the injection, while discrimination is addressed by optimizing other parameters (such as liner geometry), but especially using a PTV injector (see next) [20, 58–60].

$$Split\ ratio\ = \frac{Flow\ rate\ of\ gas\ through\ the\ split\ vent}{Flow\ rate\ through\ the\ column} \qquad \text{Eq. 3.6}$$

Splitless injection, on the other hand, is the preferred technique for trace analysis. Contrary to split injections, it allows a fairly complete transfer of the sample vapors into the column. In this mode, the split valve is closed prior to injection, and the vaporizing chamber temporarily stores the sample vapors. The transfer of analytes into the column occurs at the column flow rate and typically requires less than a minute to complete. After this period, the split valve is reopened to vent the remaining solvent vapors, preventing the formation of an excessively broad solvent peak [60]. In conventional splitless injections, the maximum sample volume is determined by the capacity of the injector liner to store the solvent vapors. If the liner is overloaded, the gases expand out of the liner into the system lines, leading to analyte losses. To avoid this, common software included in modern GC enables the calculation of the volume of the vapors, according to linear geometry, temperature, and pressure of the injector [60].

B. Programmed temperature vaporization injector (PTV)

As indicated by the name, this injector allows programming the temperature of injection; however, modern PTVs can also modify several other parameters such as split/splitless time, split/splitless ratio, and column flows. The optimization of these features helps create a versatile device capable of handling compounds with diverse chemical properties, concentrations, and thermolabile substances [61]. This device essentially addresses the main drawbacks of the traditional SSL injector. A key feature of this is the ability to deposit samples in liquid form by maintaining the unit block at low temperature, then gradually heating it to the predefined temperature. In this way, the subsequent heating avoids flash vaporizations and promotes a simultaneous and uniform release of all components, reducing volatility dependent losses. Moreover, injection into a cold liner ensures that there is no hot metal present during sample vaporization, so breakdown of sample components is less frequent. Moreover, no component stays in the liner long enough to be exposed to any temperatures higher than necessary for its vaporization. Finally, once the sample components of interest have vaporized, the liner may be heated to a much higher temperature to bake out traces of less volatile sample residue that would otherwise build up in the liner [20, 60, 62]. An example of a method for the PTV injector is given in Figure 3.15.

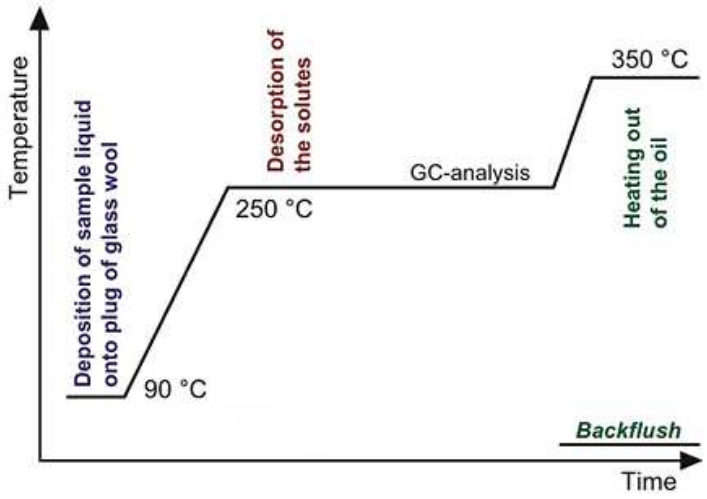


Figure 3.15. *Typical programmed temperature vaporization (PTV) injection sequence. The process begins with a low initial temperature (depending on the analytes of interest and on the solvent boiling point), during which the sample is introduced in splitless mode and deposited inside the liner, preventing immediate solvent flash vaporization. The injector is then rapidly (or gradually) heated to an intermediate temperature (approximately 250 °C), causing controlled desorption and vaporization of the analytes, which are subsequently transferred onto the GC column for the analysis. Once analyte transfer is complete, the temperature is elevated further (up to 350 °C) to remove non-volatile matrix components, minimizing carryover and contamination [62].*

The benefit of PTV injectors extends beyond addressing the main problems related to common SSL, as seen so far, and also enables options previously not possible. In this sense, for some samples, lower detection limits are needed, which can only be achieved by injecting a larger volume of sample into the column in a large-volume injection (LVI) [63]. This technique can be considered a combination of split and splitless modes applied sequentially at different temperatures. A large sample volume (typically from a few µL to 150 µL [64]) is introduced into a cold PTV liner while the split vent is open and the carrier-gas flow is set to a high flow rate (≥100 mL min^{-1}). Under these conditions, the solvent is selectively evaporated and purged through the split vent, whereas the less volatile analytes remain

trapped in the liner. After the solvent venting step, the split valve is closed, and the injector is rapidly heated to vaporize the retained analytes and transfer them, in splitless mode, into the GC column [20].

- Thermal desorption tubes

The development of low-temperature injection systems has allowed PTV units to evolve into highly versatile devices, capable not only of introducing liquid samples (by the syringe) but also of accommodating entire sampling devices placed directly inside the liner or the injector body. This versatility is exemplified using thermal desorption tubes, which transformed PTV injectors into an integrated thermal-desorption interface [65].

Thermal desorption tubes are versatile sample introduction techniques for gas chromatography, integrating sampling and injection into a single device. Essentially, by heating sorbents under a flow of gas, volatile compounds are released and transferred directly into the GC column, often achieving several enrichment factors compared to solvent extraction/liquid injection approaches [66, 67]. These solvent-free devices consist essentially of glass or metal tubes packed with different sorbent materials that should meet specific requirements, such as inertness, hydrophobicity, low tendency to artefact formation, and high-temperature stability [67]. Sorbents can be broadly categorized into carbon-based materials and porous organic polymers. The first, such as activated carbon, carbon molecular sieves, and graphitized carbon blacks, offer high surface areas and strong dispersive interactions, making them suitable for hydrocarbons and other low-polarity analytes. Porous organic polymers, in contrast, provide thermally stable, highly hydrophobic adsorbents widely used for mid-volatility compounds. If a wide volatility range of compounds is to be monitored, it is often necessary to pack a tube with more than one sorbent material, arranged in order of increasing strength from the sampling end. A schematic representation of a section of a thermal desorption tube is given in Figure 3.16.

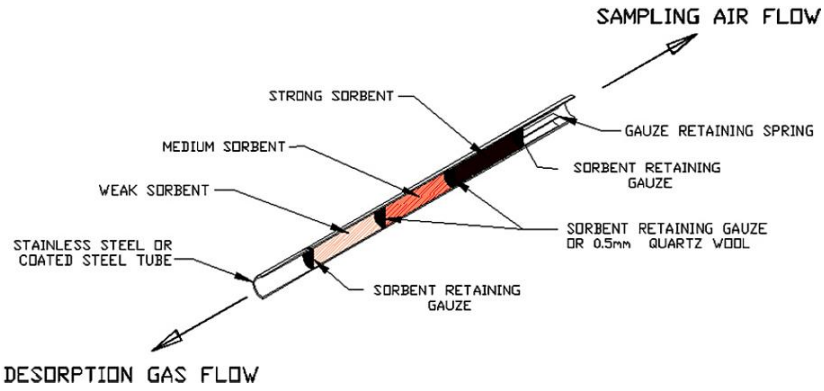


Figure 3.16. *Multi-sorbent thermal desorption tube, designed to retain analytes across a broad volatility range. Air flows from the open end through the sorbent beds, trapping analytes according to their affinity, while during thermal desorption, the carrier gas flows in the opposite direction to efficiently release the retained analytes [67].*

Standard sampling tubes often require a large volume of carrier gas for complete desorption of analytes from the sorbent, and the "injection" (or transfer) of these volumes is often incompatible with capillary chromatography, compromising analytical resolution and sensitivity. In this sense, a dual-stage thermal desorption is a key advancement over single-stage systems, enabling far more efficient analyte concentration. In the single-stage method (Figure 3.17A), the sorbent tube is simply heated, and trapped compounds are released into the high-flow carrier gas. In the dual-stage system, instead, analytes after the first desorption are captured on a secondary, cryogenically or thermoelectrically cooled focusing trap, which

concentrates them into a narrow band. The trap is then rapidly heated, re-injecting the analytes as a sharp band [65, 68–71] (Figure 3.17B).

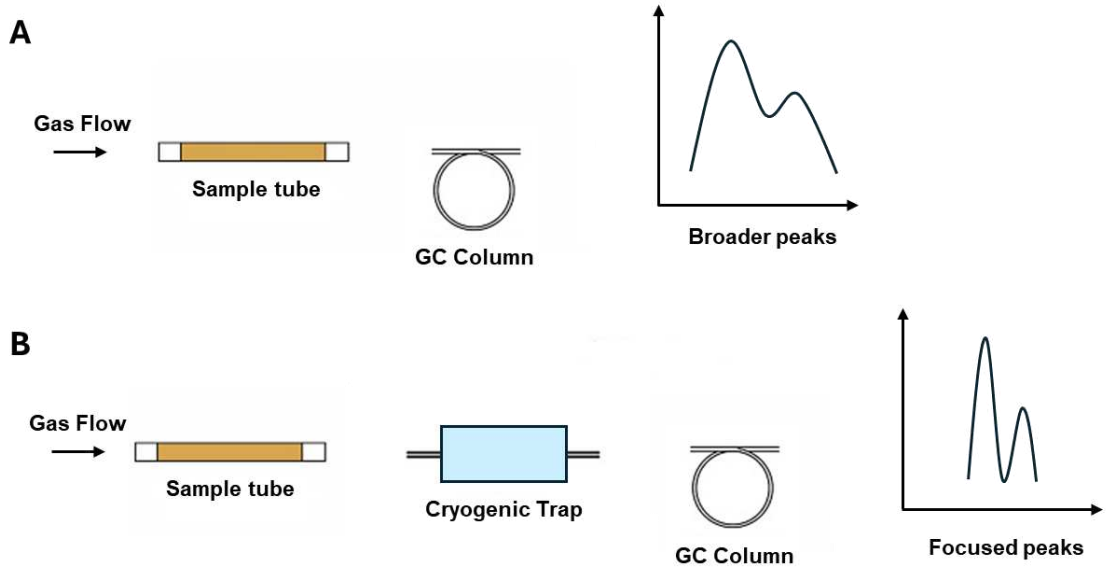


Figure 3.17. *Single-stage and dual-stage thermal desorption. A) Single-stage TD; analytes released from the sample tube are transferred directly to the GC column through the gas stream, resulting in wide injection bands and poor peak shapes. B) Dual-stage TD; after the first stage, analytes are trapped in a short section of a capillary column by a focusing trap; the trap is then rapidly heated, releasing the analytes in a narrow, concentrated plug. This refocusing effect generates sharper, more resolved chromatographic peaks. Adapted from [65].*

3.5. Column combinations

To satisfy the orthogonality principle, the two GC columns connected in series must rely on different and independent separation mechanisms. In most cases, GC×GC separations involved a first non-polar column (*e.g.*, 100% dimethyl polysiloxane or 5% diphenyl/95% dimethyl polysiloxane) and a more polar or (semi-)polar second column (50% phenyl/50% methyl polysiloxane) [72, 73]. In this configuration, analytes are separated as a function of increasing boiling point in the first dimension and according to their polarity in the second [72]. This configuration is usually referred to as normal orthogonality in the literature; conversely, a reverse orthogonality system associates a first (semi-)polar column with a second non-polar column. Independent of the setup used, following the Giddings principle [15], the more orthogonal the system is, the more independent the retention mechanisms of the two columns will be, and thus, theoretically, the more efficient the separation of the compounds will be in the 2D chromatographic plane. A visual representation of how the 2D plot varies with the column configurations is shown in Figure 3.18.

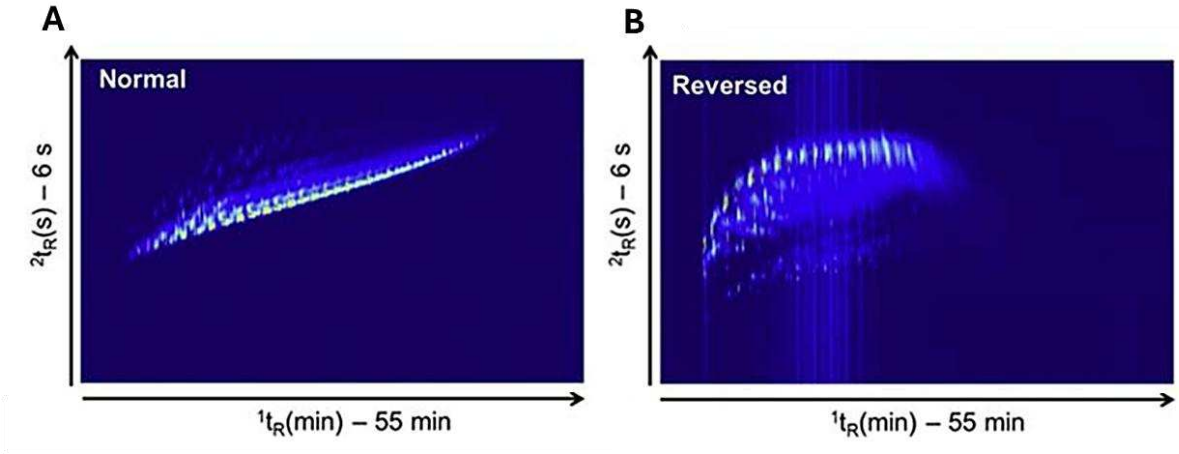


Figure 3.18. *GC×GC chromatographic patterns obtained using normal and reversed column configuration on a diesel sample. A) Normal configuration. Analytes exhibit a structured band where retention increases with volatility in the first dimension and polarity in the second, producing ordered distributions. B) Reversed configuration. Analytes lose their structured alignment through the horizontal separation and produce a more diffuse pattern in the vertical separation [72].*

Regarding column dimensions, the first column is identical to those used in conventional 1D-GC. This is because it is preferable to have relatively wide peak widths at the modulator inlet to ensure proper sampling of potentially co-eluting analytes before their separation in the second dimension. First dimension columns are commonly 15 m to 60 m long, with an inner diameter in the range of 0.25-0.53 mm and a film thickness of the order of 0.25-1 µm. These columns provide conventional peak widths of 5-30 s. Conversely, the second separation must be quick enough to complete in a time shorter than the modulation period to prevent any wraparound effect (the elution of a compound in the next modulation cycle, appearing incorrectly in the following slice). Therefore, short, narrow-bore columns are used in the second dimension: commonly 0.5-1.5 m long, 0.1-0.25 mm in inner diameter, and with a film thickness of 0.1-0.25 µm [27, 39].

3.6. Modulators

To achieve a high-quality two-dimensional gas chromatography separation, it is important to ensure the effective transfer of the first dimension (^{1}D) column effluent to the second dimension (^{2}D). To preserve the separation obtained in the first dimension, a modulator interface is required. Creating a universal modulator is challenging, and over the years, many different types of devices have been developed [74–76].

A modulator must perform three primary functions: trapping, focusing, and reinjecting (^{1}D column effluent to the ^{2}D column). Specifically, sample analytes subjected to first-dimension separation enter the modulator interface *via* the column. Here, they are then collected for a predetermined period before being reinjected as a narrow pulse onto the head of the ^{2}D column for further separation prior to detection [74].

Since its inception, various modulators have been developed, each employing different physical properties to facilitate modulation. Modulators are generally classified into two main categories: thermal modulators and flow modulators. The first type uses temperature control to trap analytes from the ^{1}D effluent and release them in the ^{2}D separation. The second category utilizes gas flow to control and isolate fractions of the ^{1}D eluate, redirecting them into the second dimension [77]. Then, modulators are also categorized as “low duty cycle” and “high (or unit) duty cycle”, where only small portions or all the ^{1}D eluate is directed to the ^{2}D head [78].

In order to preserve the separation performed in the first dimension, the effluent coming from the first dimension must be sampled many times to avoid any interference of already separated components. A general rule to meet this requirement is that each chromatographic peak must be sampled within its peak width at the base (ωb), three or four times (Mr) according to its respective modulation period (P_M) [79]. The illustration of this simple arithmetical manipulation is given in Eq. 3.6.

$$M_r = \frac{\omega_b}{P_M} \qquad \text{Eq. 3.6}$$

Furthermore, another important parameter is the modulation phase, defined as the difference between the center of the ^{1}D peak and the mean of the peak region sampled by the modulator. This attribute directly influences the resolution and the reconstructed peak width of the analyte eluting from the ^{1}D column [80]. Two extreme situations can be seen in Figure 3.19:

A. In-phase modulation; symmetrical pulse sequence with a single maximum peak is observed.

B. 180 degrees out-of-phase modulation; pulse sequence corresponds to two equal maximum peaks.

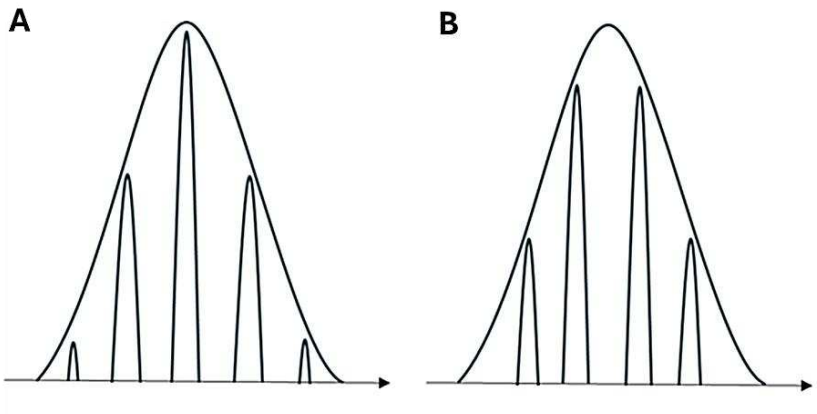


Figure 3.19. *Phase modulation features. A) In-phase modulation. B)180 degrees out-of-phase modulation.*

Among the two modulation classes previously reported, thermal modulation is the most popular. The concept of thermal modulation in GC×GC was first introduced in 1985 by Phillips *et al.* [38], who created (as indicated in Section 3.2) a simple thermal desorption device consisting of a short section of coated capillary that could be rapidly heated by an electrical current through a conductive film. Although pioneering, this early device produced broad, unfocused injection plugs because analytes entering the modulator were not sufficiently concentrated before being released into the second column. Subsequent efforts by the same group led to the development of the "sweeper" modulator, the first commercially available thermal modulator [81]. In this design, a slotted heater, maintained roughly 100 °C above the oven temperature, was mechanically swept over a thick-film capillary segment acting as a trapping medium. By alternating heating and cooling, this device generated sharper pulses and significantly improved reinjection focusing.

Towards the late 1990s, a major conceptual shift occurred with the introduction of cryogenic modulation, where cooling rather than heating produced the required focusing effect. Early cryogenic modulators used a mobile liquid CO_2 jet that moved along the column to trap analytes at the head of the second dimension [75]. Once cooled, the analytes were immobilized until the trap moved away, releasing them as a short plug. Although this system represented the first reliable cryogenic modulator suitable for routine use, its performance was limited by the relatively warm temperature of liquid CO_2 (approximately -70 °C), which was insufficient for efficient trapping of very volatile compounds, associated also with the mechanical fragility of the moving shuttle near the capillary [75].

To address these issues, Beens *et al.* [82], developed an alternative CO_2-based cryogenic modulator without moving parts. In this configuration, two fixed jets alternately cooled adjacent segments of the capillary, trapping sequential fractions of the effluent which were then refocused and released by the heat of the surrounding oven air. This dual-jet approach produced highly reproducible modulation cycles and eliminated the mechanical vulnerabilities of earlier systems. Later, Adahchour *et al.* [83] simplified the design further by introducing a single-jet version, reducing instrumental complexity but requiring meticulous flow optimization due to a single focusing zone.

In 2000, Ledford *et al.* described an improved CO_2-cooled modulator that incorporated a continuously flowing cold jet and hot gas pulses controlled by solenoid valves to create dual-stage thermal modulation without moving parts [84]. Then, the concept was further refined using liquid nitrogen, enabling efficient modulation of even the most volatile analytes. A schematic illustration of the quad-jet mechanism is given in Figure 3.20.

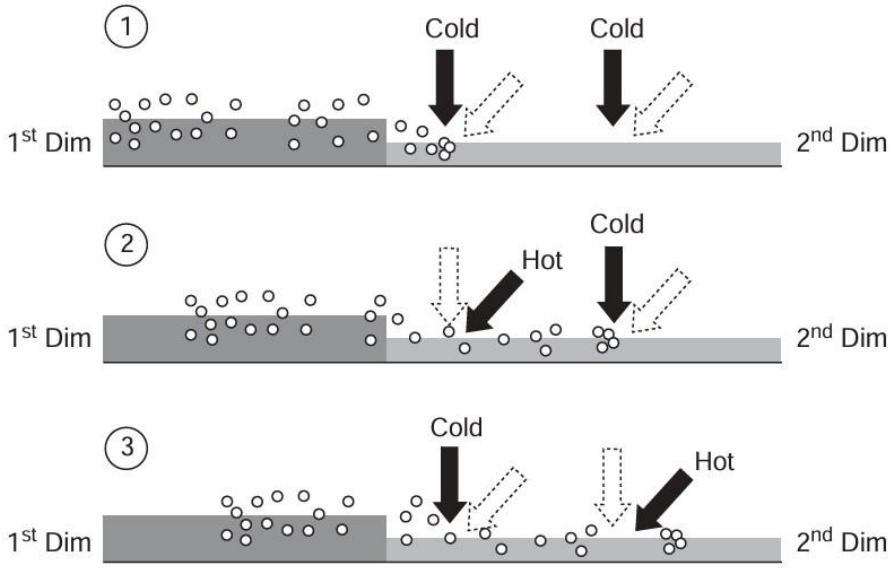


Figure 3.20. *Sequence of events in a quad-jet dual-stage cryogenic modulator. In step 1, analytes eluting from the first dimension column are trapped at the head of the second dimension by two cold jets, immobilizing the effluent in a narrow cryogenic zone. In step 2, one cold jet remains active while a hot jet briefly heats the trapped segment, releasing the focused analytes as a sharp injection plug into the second dimension column. Meanwhile, the continuous effluent from the first dimension accumulates in the adjacent cooled zone. In step 3, the cycle repeats: the newly trapped material is cooled while the previously accumulated analytes are remobilized by a hot jet pulse. This alternating cold and hot sequence produces periodic focused injections that define the modulation process in comprehensive two-dimensional gas chromatography [39].*

Most recently, commercial developments have introduced closed-cycle cryogenic refrigeration systems that eliminate the need for consumable cryogens while still achieving

effective modulation of volatile and semi-volatile compounds. These systems represent the latest evolution in thermal modulation technology, offering sustainability and operational simplicity for modern GC×GC workflows [85].

3.7. Mass spectrometry detection

Extremely narrow peaks are obtained in comprehensive two-dimensional gas chromatography, typically between 100 and 500 ms. Those values place stringent performance requirements on the detector. To accurately reconstruct such narrow peaks, the detector must have a very high acquisition frequency. At least ten data points are required for a reliable quantification and consequently, a 200 ms peak demands acquisition frequencies of at least 50 Hz, with higher rates preferred for complex samples [86]. If this condition is not met, the chromatographic peak cannot be properly reconstructed (Figure 3.21), resulting in a loss of chemical information and reduced sensitivity. In practice, when the detector acquisition rate is not adequately optimized, the outcome is a distorted peak shape. Another important aspect to consider is how the detector responds to an incoming chemical signal (*i.e.*, the ability to produce a quantifiable signal when an analyte enters the detection region). This parameter is not related to how often the signal is sampled, but by how strongly the detector reacts to the passage of the analyte. Such behavior depends on various factors, such as the physio-chemical mechanism behind the detection, as well as on factors like detector geometry and electronic components [87].

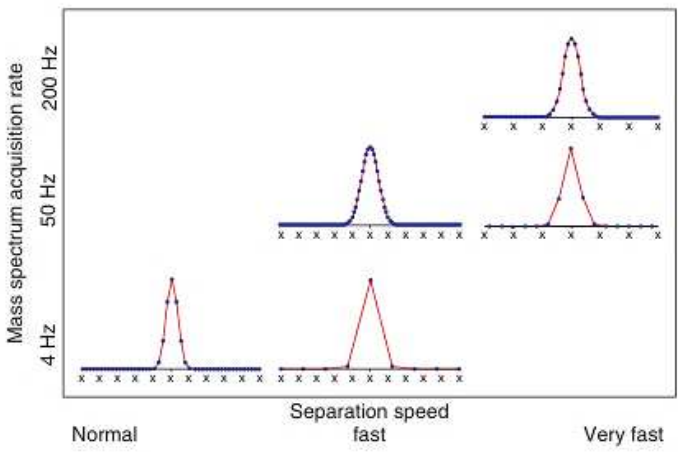


Figure 3.21. *Relationship between chromatographic separation speed and mass-spectral acquisition rate [87].*

Historically, the flame ionization detector was first used in GC×GC due to its fast sampling rate, whereas mass spectrometers, in contrast, initially struggled with the rapid acquisition requirements of this technique. Advances in electronics and system architecture only then made mass spectrometry (MS) fully compatible with GC×GC, enabling highly informative detection on the chromatographic separation [88].

One of the major strengths of GC×GC-MS lies in its multidimensional selectivity. Compounds that co-elute in the first dimension often share similar vapor pressures and therefore retain similar retention times on low-polarity stationary phases. In the second dimension, analytes undergo rapid separation driven by different interactions. Moreover, if co-elution persists, the mass analyzer provides an additional orthogonal separation based on mass-to-charge ratio (*m/z*). The combination of those mechanisms creates a "cubic separation space" in which identification relies on three independent parameters: first dimension retention time, second dimension retention time and mass spectral fragmentation

patterns [89]. A schematic representation of a high-order data structure is given in Figure 3.22.

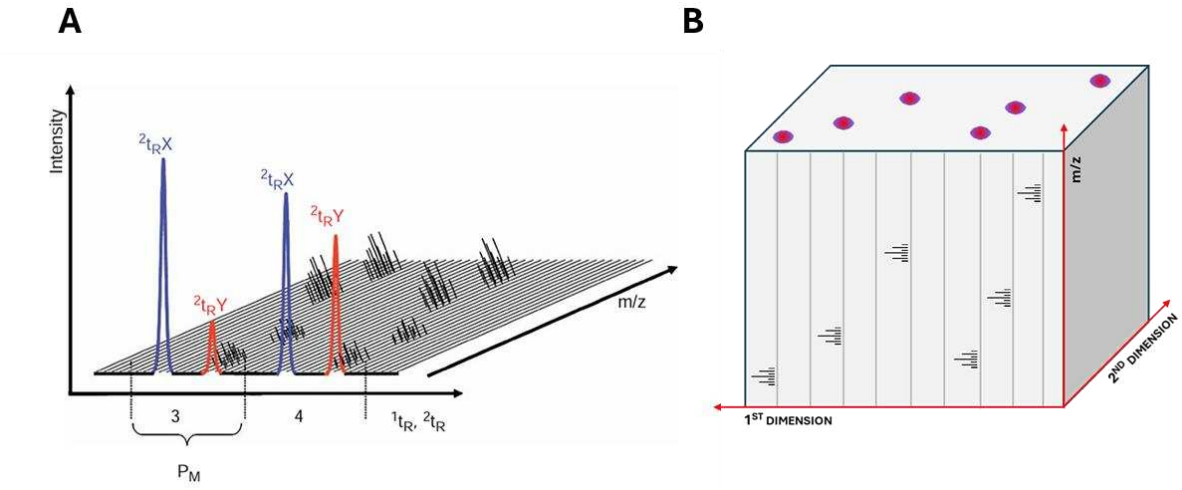


Figure 3.22: *Generation of a three-dimensional data structure. A) Each modulation slice produces a second-dimension peak accompanied by a full mass spectrum. B) Resulting "data cube" combining chromatographic and spectral information. Adapted from [39, 89].*

In GC×GC-MS configurations, the second dimension column is interfaced directly with the mass spectrometer via a heated transfer line, which prevents gas condensation and maintains chromatographic integrity. The mass spectrometer itself consists of several core components that collectively operate under high-vacuum conditions: sample inlet, ion source, mass analyzer, and detector. Typical pressures are 10^{-4} and 10^{-8} torr, and the mass-analysis chamber is typically at the lower end of this range [90].

Ionization is usually accomplished by electron ionization (EI), the most common and robust technique for volatile organic compounds. In EI, a beam of high-energy electrons interacts with neutral molecules, generating fragmentation pathways. The resulting fragment ions are detected and represented as a mass spectrum, reporting relative ion abundance as a function of the mass-to-charge (*m/z*) ratio. The EI fragmentation behavior reflects a unique molecule's properties and, due to its high reproducibility, enables confident compound identification through spectral libraries (see Chapter 2, Section 2.4.1). Following ionization, ions are accelerated and directed into the mass analyzer, which separates them according to their *m/z* values before they are detected and converted into an electronic signal [91].

Chemical ionization (CI) is a much "softer" ionization technique than EI and involves the reaction of sample molecules with an ionized "reagent gas" within the ion source. Essentially, the gas is hit by a beam of electrons, producing reactive species that in turn interact with the sample analytes, ionizing them. Since ionization occurs through low energy reactions between ions and analyte molecules instead of direct electron interactions, fragmentation is greatly reduced compared to EI. This leads to less complex spectra with a higher abundance of the molecular ion. The ion source is built similarly to the EI source but with a much smaller exit slit, making it more gas-tight and allowing pressures of 0.1 to 1 Torr to be achieved inside (Figure 3.23). Although many suitable gases are reported, methane, the iso-butane, and ammonia are the most common [91, 92].

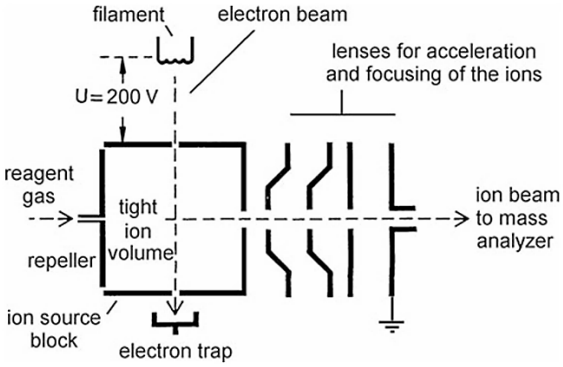


Figure 3.23. *Schematic representation of a soft ionization source. A filament emits an electron beam into a tightly sealed ionization chamber containing reagent gas. Sample ions formed in this region are then guided by acceleration lenses toward the mass analyzer [91].*

Moving to mass analyzer, time-of-flight MS (TOF-MS) are the most widely used in GC×GC. Triple quadrupole (QqQ-MS) systems also offer robustness and targeted quantification capabilities, but are limited in acquisition speed. As scanning devices, quadrupoles sequentially monitor ions across the mass range during chromatographic peak elution. If the scan speed is not sufficiently fast, due for example to geometrical or design constraints, the measured ion intensities become strongly dependent on how the mass scan is distributed along the elution profile the peak [93, 94]. Essentially, during the rising portion of the peak, ions at higher *m/z* values, that are scanned later, may appear more intense than lighter ions scanned earlier; conversely, after the peak apex, lighter ions scanned first may appear more intense than heavier ions. This effect also impacts spectral deconvolution, as distorted ion profiles resulting from poor scan timing compromise spectral consistency, complicating the reliable separation of co-eluting compounds, as generally observed for deconvolution of skewed peak profiles [95]. Acquisition frequency can be increased by narrowing the scanned mass range, however at the cost of restricting the accessible *m/z* window. Time-of-flight instruments, instead, are uniquely suited to GC×GC because they can acquire full-range mass spectra at very high acquisition rate, at hundreds of spectra per second [96].

The time-of-flight (TOF) mass analyzer consists essentially of a high-vacuum flight tube in which ions are separated by a velocity-dependent drift (Figure 3.24). In particular, ions produced in the source are accelerated into the flight tube with equal kinetic energy, and since their velocity depends on the mass-to-charge ratio (*m/z*), lighter ions arrive sooner at the detector than heavier ones. This enables all ions to be transmitted simultaneously rather than scanned sequentially, resulting in fast scans and wide mass ranges [91].

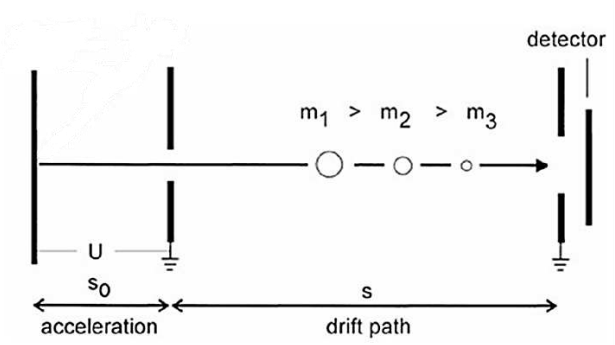


Figure 3.24. *Operating principle of a time-of-flight (TOF) mass spectrometer. After acceleration, the lightest (m3) ions move fastest and advance further along the drift tube, while the heavier (m1 & m2) ions stand behind [91].*

TOF instruments are generally classified into low-resolution (LR-TOFMS) and high-resolution (HR-TOFMS) systems. LR-TOFMS provides unit mass resolution at very high acquisition frequencies (up to 500 Hz), and quantification is commonly performed using extracted-ion chromatograms (EICs) [97].

The coupling of GC×GC with high-resolution TOFMS dramatically enhances analytical capability, although its use is generally reserved for applications requiring the highest level of selectivity and structural information. HR-TOFMS provides full-spectrum acquisition with high mass accuracy and resolution, enabling reliable molecular formula determination when the molecular ion is present. Greater mass accuracy reduces the number of plausible formula candidates for a given measured mass. For targeted analysis, extracted-ion chromatograms with narrow mass windows (≤5 ppm) substantially reduce matrix and chemical noise interferences, thereby improving analytical specificity [98]. A visual representation of this effect is provided by the author of this PhD thesis in Figure 3.25.

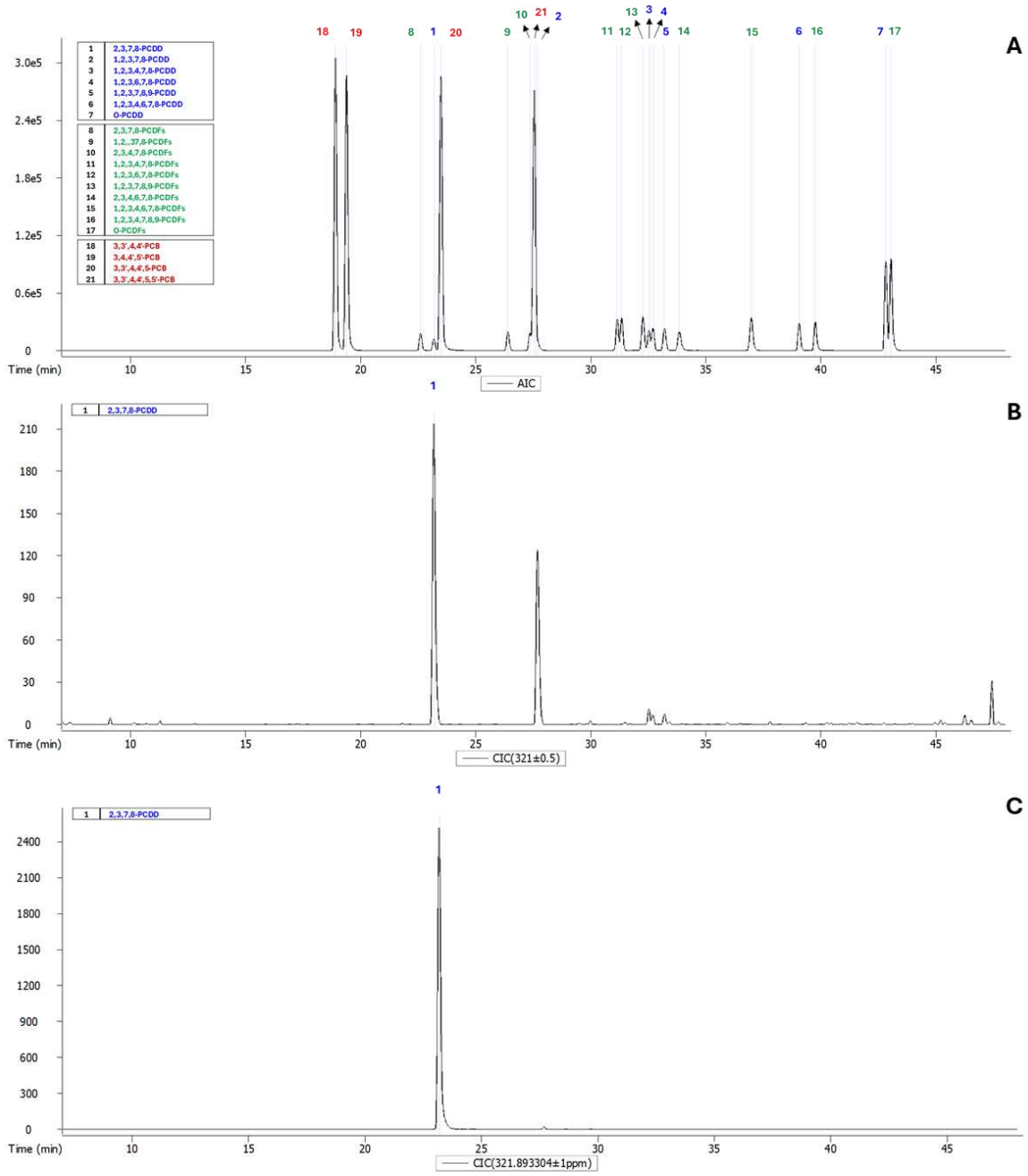


Figure 3.25. *Effect of narrowing mass extraction window on chromatographic signal intensity of a mixture of Dioxines, Furans, and Biphenyls. A) Full analytical ion chromatogram. This chromatogram is generated by summing the masses in each peak true spectrum of a series of Dioxins (PCDD), Furans (PCDF), and PCBs. B) Cluster ion chromatogram (CIC) generated with a tighter mass window (321 ± 0.5 m/z). By narrowing the mass tolerance, most of the irrelevant ions are excluded, and only a few peaks remain clearly visible. C) Extremely narrow mass window chromatogram (321.893304 m/z ± 1 ppm). Unrelated ions are excluded, and only the peak of the target analyte remains, yielding a single, sharp peak with minimal interference. Despite the drastic decrease in overall detected ions, the signal-to-noise ratio improves markedly with respect to B, as the peak of interest becomes isolated from the background chemical noise.*

References

[1] R. Aparicio and R. Aparicio-Ruız, "Authentication of vegetable oils by chromatographic techniques," *J. Chromatogr. A*, vol. 881, 2000, doi: 10.1016/S0021-9673(00)00355-1.

[2] X. Shi, H. Li, Z. Song, X. Zhang, and G. Liu, "Quantitative composition-property relationship of aviation hydrocarbon fuel based on comprehensive two-dimensional gas chromatography with mass spectrometry and flame ionization detector," *Fuel*, vol. 200, 2017, doi: 10.1016/j.fuel.2017.03.073.

[3] L. Mondello, *Comprehensive Chromatography in Combination with Mass Spectrometry*. John Wiley & Sons, Inc., 2011. doi: 10.1002/9781118003466.

[4] R. Consden, A. H. Gordon, and J. P. Martin, "Qualitative Analysis of Proteins: a Partition Chromatographic Method Using Paper," *Biochemical Journal*, vol. 12, no. 223, p. 11, 1944, doi: 10.1042/bj0380224.

[5] J. C. May and J. A. McLean, "Advanced Multidimensional Separations in Mass Spectrometry: Navigating the Big Data Deluge," *Annual Review of Analytical Chemistry*, vol. 9, Jun. 2016, doi: 10.1146/annurev-anchem-071015-041734.

[6] Z. Liu and J. B. Phillips, "Comprehensive Two-Dimensional Gas Chromatography using an On-Column Thermal Modulator Interface," *J. Chromatogr. Sci.*, vol. 29, 1991, doi: 10.1093/chromsci/29.6.227.

[7] F. Erni and R. W. Frel', "Two-dimensional column liquid chromatographic technique for resolution of complex mixture," *J. Chromatogr.*, vol. 149, 1978, doi: 10.1016/S0021-9673(00)81011-0.

[8] M. Zakaria, M.-F. Gonnord, and G. Guiochon, "Applications of two-dimensional thin-layer chromatography," *J. Chromatogr.*, vol. 271, 1983, doi: 10.1016/S0021-9673(00)80218-6.

[9] J. C. Giddings, "Concepts and Comparisons in Multidimensional Separation," *Journal of High Resolution Chromatography & Chromatography Communications*, 1987, doi: 10.1002/jhrc.1240100517.

[10] H. M. McNair, J. M. Miller, and N. H. Snow, *Basic Gas Chromatography*, Third Edition. John Wiley & Sons, Inc., 2019. doi: 10.1002/9781119450795.

[11] C. J. Venkatramani, J. Xu, and J. B. Phillips, "Separation Orthogonality in Temperature-Programmed Comprehensive Two-Dimensional Gas Chromatography," *Anal. Chem.*, vol. 68, no. 1, pp. 1486–1492, 1996, doi: 10.1021/ac951048b.

[12] D. Ryan, P. Morrison, and P. Marriott, "Orthogonality considerations in comprehensive two-dimensional gas chromatography," *J. Chromatogr. A*, vol. 1071, no. 1–2, Apr. 2005, doi: 10.1016/j.chroma.2004.09.020.

[13] A. Venter, "Comprehensive Two-Dimensional Supercritical Fluid and Gas Chromatography (SFCxGC)," University of Pretoria, 2004. Accessed: Nov. 25, 2025. [Online]. Available: http://hdl.handle.net/2263/23155

[14] R. B. Wilson, W. C. Siegler, J. C. Hoggard, B. D. Fitz, J. S. Nadeau, and R. E. Synovec, "Achieving high peak capacity production for gas chromatography and comprehensive

two-dimensional gas chromatography by minimizing off-column peak broadening," *J. Chromatogr. A*, vol. 1218, no. 21, May 2011, doi: 10.1016/j.chroma.2010.12.108.

[15] J. C. Giddings, "Sample dimensionality: a predictor of order-disorder in component peak distribution in multidimensional separation," *J. Chromatogr. A*, vol. 703, 1995, doi: https://doi.org/10.1016/0021-9673(95)00249-m.

[16] W. Engewald and K. Dettmer-Wilde, "Theory of Gas Chromatography," in *Practical Gas Chromatography*, W. Engewald and K. Dettmer-Wilde, Eds., Springer, 2017, ch. 2. doi: 10.1007/978-3-642-54640-2_2.

[17] J. C. Giddings, *Unified Separation Science*. John Wiley & Sons, Inc., 1991. doi: 10.1002/aic.690380818.

[18] R. L. Grob and E. F. Barry, *Modern Practice of Gas Chromatography*. John Wiley & Sons, Inc., 2013. doi: 10.1002/0471651141.

[19] L. M. Blumberg, *Temperature-Programmed Gas Chromatography*. WILEY-VCH Verlag GmbH & Co., 2010. doi: 10.1002/9783527632145.

[20] C. F. Poole, "Chapter 3: Conventional Gas Chromatography: Basic Principles and Instrumental Aspects," in *Advanced Gas Chromatography in Food Analysis*, P. Q. Tranchida, Ed., The Royal Society of Chemistry 2020, 2019. doi: https://doi.org/10.1039/9781788015752.

[21] E. A. Higgins Keppler, C. L. Jenkins, T. J. Davis, and H. D. Bean, "Advances in the application of comprehensive two-dimensional gas chromatography in metabolomics," *TrAC - Trends in Analytical Chemistry*, vol. 109, p. 28, Dec. 2018, doi: 10.1016/j.trac.2018.10.015.

[22] Y. Izadmanesh, E. Garreta-Lara, J. B. Ghasemi, S. Lacorte, V. Matamoros, and R. Tauler, "Chemometric analysis of comprehensive two dimensional gas chromatography–mass spectrometry metabolomics data," *J. Chromatogr. A*, vol. 1488, Mar. 2017, doi: 10.1016/j.chroma.2017.01.052.

[23] A. M. Muscalu and T. Górecki, "Comprehensive two-dimensional gas chromatography in environmental analysis," *TrAC - Trends in Analytical Chemistry*, vol. 106, Sep. 2018, doi: 10.1016/j.trac.2018.07.001.

[24] T. J. Trinklein, C. N. Cain, G. S. Ochoa, S. Schöneich, L. Mikaliunaite, and R. E. Synovec, "Recent Advances in GC×GC and Chemometrics to Address Emerging Challenges in Non-targeted Analysis," *Anal. Chem.*, vol. 95, no. 1, Jan. 2023, doi: 10.1021/acs.analchem.2c04235.

[25] P. H. Stefanuto, A. Smolinska, and J. F. Focant, "Advanced chemometric and data handling tools for GC×GC-TOF-MS: Application of chemometrics and related advanced data handling in chemical separations," *TrAC - Trends in Analytical Chemistry*, vol. 139, Jun. 2021, doi: 10.1016/j.trac.2021.116251.

[26] D. Zanella, J. F. Focant, and F. A. Franchina, "30th Anniversary of comprehensive two-dimensional gas chromatography: Latest advances," *Analytical Science Advances*, vol. 2, Apr. 2021, doi: 10.1002/ansa.202000142.

[27] F. Bertoncini, M. Courtiade-Tholance, D. Thiébaut, and T. Jones, *Gas Chromatography and 2D-Gas Chromatography for Petroleum Industry: The Race for Selectivity*. TECHNIP, 2013. ISBN: 978-2-7108-0992-0

[28] T. Górecki, J. Harynuk, and O. Panić, "The evolution of comprehensive two-dimensional gas chromatography (GC x GC)," *J. Sep. Sci.*, vol. 27, no. 5–6, Apr. 2004, doi: 10.1002/jssc.200301650.

[29] P. Q. Tranchida, D. Sciarrone, P. Dugo, and L. Mondello, "Heart-cutting multidimensional gas chromatography: A review of recent evolution, applications, and future prospects," *Anal. Chim. Acta*, vol. 716, Feb. 2012, doi: 10.1016/j.aca.2011.12.015.

[30] M. Catani, S. Felletti, and F. A. Franchina, "Separation techniques," in *Metabolomics Perspectives: From Theory to Practical Application*, Elsevier, 2022. doi: 10.1016/B978-0-323-85062-9.00003-9.

[31] B. W. J. Pirok, D. R. Stoll, and P. J. Schoenmakers, "Recent Developments in Two-Dimensional Liquid Chromatography: Fundamental Improvements for Practical Applications," *Anal. Chem.*, vol. 91, no. 1, Jan. 2019, doi: 10.1021/acs.analchem.8b04841.

[32] M. C. Simmons and L. R. Snyder', "Two-Stage Gas-Liquid Chromatography," *Anal. Chem.*, vol. 30, 1958, doi: https://doi.org/10.1021/ac60133a007.

[33] H. Schmidt-Traub, *Preparative Chromatography of Fine Chemicals and Pharmaceutical Agents*. Wiley-VCH, 2005. doi: 10.1002/3527603484.

[34] L. Ramos and U. A. T. Brinkman, "Multidimensionality in Gas Chromatography: General Concepts," in *Comprehensive Analytical Chemistry*, vol. 55, 2009, ch. Chapter 1, p. 14. doi: 10.1016/S0166-526X(09)05501-9.

[35] L. M. Blumberg, F. David, M. S. Klee, and P. Sandra, "Comparison of one-dimensional and comprehensive two-dimensional separations by gas chromatography," *J. Chromatogr. A*, vol. 1188, no. 1, Apr. 2008, doi: 10.1016/j.chroma.2008.02.044.

[36] M. S. Klee, J. Cochran, M. Merrick, and L. M. Blumberg, "Evaluation of conditions of comprehensive two-dimensional gas chromatography that yield a near-theoretical maximum in peak capacity gain," *J. Chromatogr. A*, vol. 1383, pp. 151–159, Feb. 2015, doi: 10.1016/j.chroma.2015.01.031.

[37] J. B. Phillips and Z. Liu, "United States Patent: 5,135,549," 1992

[38] J. B. Phillips, D. Luu, J. B. Pawliszyn, and G. C. Carle, "Multiplex Gas Chromatography by Thermal Modulation of a Fused Silica Capillary Column," *Anal. Chem.*, vol. 57, 1985, doi: https://doi.org/10.1021/ac00291a010.

[39] G. Semard, M. Adahchour, and J. F. Focant, "Basic Instrumentation for GC×GC," in *Comprehensive Analytical Chemistry*, vol. 55, 2009, ch. Chapter 2. doi: 10.1016/S0166-526X(09)05502-0.

[40] J. F. Focant *et al.*, "Measurement of PCDDs, PCDFs, and non-ortho-PCBs by comprehensive two-dimensional gas chromatography-isotope dilution time-of-flight mass spectrometry (GC x GC-IDTOFMS)," *Talanta*, vol. 63, no. 5, Aug. 2004, doi: 10.1016/j.talanta.2004.05.043.

[41] J. Beens, *Comprehensive Two-Dimensional Gas Chromatography: The State of Seperation Arts*. www.Chromedia.org, 2009.

[42] J. V. Seeley and S. K. Seeley, "Multidimensional gas chromatography: Fundamental advances and new applications," *Anal. Chem.*, vol. 85, no. 2, Jan. 2013, doi: 10.1021/ac303195u.

[43] M. Adahchour, J. Beens, and U. A. T. Brinkman, "Recent developments in the application of comprehensive two-dimensional gas chromatography," *J. Chromatogr. A*, vol. 1186, no. 1–2, Apr. 2008, doi: 10.1016/j.chroma.2008.01.002.

[44] P. Marriott and R. Shellie, "Principles and applications of comprehensive two-dimensional gas chromatography," *TrAC - Trends in Analytical Chemistry*, vol. 21, 2002, doi: https://doi.org/10.1016/S0165-9936(02)00814-2.

[45] T. Górecki and A. Mostafa, "Sensitivity of Comprehensive Two-dimensional Gas Chromatography (GCXGC) Versus One-dimensional Gas Chromatography (1D GC)." [Online]. Available: https://www.chromatographyonline.com/view/sensitivity-comprehensive-two-dimensional-gas-chromatography-gcxgc-versus-one-dimensional-ga

[46] J. Beens and U. A. T. Brinkman, "Comprehensive two-dimensional gas chromatography - A powerful and versatile technique," *Analyst*, vol. 130, no. 2, 2005, doi: 10.1039/b407372j.

[47] W. Welthagen, R. A. Shellie, J. Spranger, M. Ristow, R. Zimmermann, and O. Fiehn, "Comprehensive two-dimensional gas chromatography-time-of-flight mass spectrometry (GC × GC-TOF) for high resolution metabolomics: Biomarker discovery on spleen tissue extracts of obese NZO compared to lean C57BL/6 mice," *Metabolomics*, vol. 1, no. 1, p. 65, Mar. 2005, doi: 10.1007/s11306-005-1108-2.

[48] M. Adahchour, J. Beens, R. J. J. Vreuls, and U. A. T. Brinkman, "Recent developments in comprehensive two-dimensional gas chromatography (GC × GC). I. Introduction and instrumental set-up," *TrAC - Trends in Analytical Chemistry*, vol. 25, no. 5, May 2006, doi: 10.1016/j.trac.2006.03.002.

[49] D. Ryan and P. Marriott, "Comprehensive two-dimensional gas chromatography," *Anal. Bioanal. Chem.*, vol. 376, no. 3, Jun. 2003, doi: 10.1007/s00216-003-1934-x.

[50] J. S. Arey, R. K. Nelson, L. Xu, and C. M. Reddy, "Using comprehensive two-dimensional gas chromatography retention indices to estimate environmental partitioning properties for a complete set of diesel fuel hydrocarbons," *Anal. Chem.*, vol. 77, no. 22, Nov. 2005, doi: 10.1021/ac051051n.

[51] J. Beens, J. Blomberg, and P. J. Schoenmakers, "Proper Tuning of Comprehensive Two-Dimensional Gas Chromatography (GCxGC) to Optimize the Separation of Complex Oil Fractions," *Journal of High Resolution Chromatography*, vol. 23, no. 3, 2000, doi: 10.1002/(SICI)1521-4168(20000301)23:3%3C182::AID-JHRC182%3E3.0.CO;2-E.

[52] C. Vendeuvre, R. Ruiz-Guerrero, F. Bertoncini, L. Duval, and D. Thiébaut, "Comprehensive two-dimensional gas chromatography for detailed characterisation of petroleum products," *Oil and Gas Science and Technology*, vol. 62, no. 1, Jan. 2007, doi: 10.2516/ogst:2007004.

[53] R. Van Der Westhuizen, M. Ajam, P. De Coning, J. Beens, A. de Villiers, and P. Sandra, "Comprehensive two-dimensional gas chromatography for the analysis of synthetic and crude-derived jet fuels," *J. Chromatogr. A*, vol. 1218, no. 28, Jul. 2011, doi: 10.1016/j.chroma.2011.05.009.

[54] M. E. Machado, “Comprehensive two-dimensional gas chromatography for the analysis of nitrogen-containing compounds in fossil fuels: A review,” *Talanta*, vol. 198, Jun. 2019, doi: 10.1016/j.talanta.2019.02.031.

[55] P. De Coning and J. Swinley, *A Practical Guide to Gas Analysis by Gas Chromatography*. Elsevier, 2019. ISBN: 978-0-12-818889-7

[56] J. Dallüge, J. Beens, and U. A. T. Brinkman, “Comprehensive two-dimensional gas chromatography: A powerful and versatile analytical tool,” *J. Chromatogr. A*, vol. 1000, no. 1–2, Jun. 2003, doi: 10.1016/S0021-9673(03)00242-5.

[57] O. D. Sparkman, Z. Penton, and F. G. Kitson, *Gas Chromatography and Mass Spectrometry: A Practical Guide*. Academic Press, 2011.

[58] K. Grob, *Split and Splitless Injection for Quantitative Gas Chromatography*, Fourth Edition. WILEY-VCH Verlag GmbH, 2001. ISBN: 3-527-29879-7

[59] T. E. Beesley, B. Buglio, and R. P. W. Scott, *Quantitative Chromatographic Analysis*. Marcel Dekker, Inc., 2001. ISBN:0-8247-0503-3

[60] M. Biedermann, “Inlet Systems and Sample Introduction,” in *Practical Gas Chromatography*, Springer Berlin Heidelberg, 2014, ch. Chapter 5, pp. 161–203. doi: 10.1007/978-3-642-54640-2_5.

[61] W. Engewald, J. Teske, and J. Efer, “Programmed temperature vaporiser-based injection in capillary gas chromatography,” *J. Chromatogr. A*, vol. 856, pp. 259–278, 1999, doi: 10.1016/S0021-9673(99)00723-2.

[62] A. Frankhauser-Noti and K. Grob, “Injector-internal thermal desorption from edible oils performed by programmed temperature vaporizing (PTV) injection,” *J. Sep. Sci.*, vol. 29, no. 15, Oct. 2006, doi: 10.1002/jssc.200600064.

[63] E. Hoh and K. Mastovska, “Large volume injection techniques in capillary gas chromatography,” *J. Chromatogr. A*, vol. 1186, no. 1–2, Apr. 2008, doi: 10.1016/j.chroma.2007.12.001.

[64] K. Grob and M. Biedermann, “Vaporising systems for large volume injection or on-line transfer into gas chromatography: classification, critical remarks and suggestions,” *J. Chromatogr. A*, vol. 750, 1996, doi: 10.1016/0021-9673(96)00345-7.

[65] E. Woolfenden, “Thermal Desorption Gas Chromatography,” in *Gas Chromatography*, Second Edition., C. F. Poole, Ed., in Handbooks in Separation Science. , Amsterdam: Elsevier, 2021, ch. 10. doi: 10.1016/B978-0-12-820675-1.00009-5.

[66] E. Woolfenden, “Sorbent-based sampling methods for volatile and semi-volatile organic compounds in air. Part 1: Sorbent-based air monitoring options,” *J. Chromatogr. A*, vol. 1217, no. 16, Apr. 2010, doi: 10.1016/j.chroma.2009.12.042.

[67] E. Woolfenden, “Sorbent-based sampling methods for volatile and semi-volatile organic compounds in air. Part 2. Sorbent selection and other aspects of optimizing air monitoring methods,” *J. Chromatogr. A*, vol. 1217, no. 16, p. 2694, Apr. 2010, doi: 10.1016/j.chroma.2010.01.015.

[68] J.-L. Wang, S.-W. Chen, and C. Chew, “Automated gas chromatography with cryogenic / sorbent trap for the measurement of volatile organic compounds in the atmosphere,” *J. Chromatogr. A*, vol. 863, 1999, doi: https://doi.org/10.1016/S0021-9673(99)00965-6.

[69] J. V Hinshaw, "Thermal Desorption Sampling," *LCGC North America*, vol. 34, no. 1, pp. 26–32, 2016. Accessed: Nov. 25, 2025. [Online]. Available: https://www.chromatographyonline.com/view/thermal-desorption-sampling

[70] F. J. M. Novaes and P. J. Marriott, "Cryogenic trapping as a versatile approach for sample handling, enrichment and multidimensional analysis in gas chromatography," *J. Chromatogr. A*, vol. 1644, May 2021, doi: 10.1016/j.chroma.2021.462135.

[71] P. J. Marriott and R. M. Kinghorn, "Longitudinally Modulated Cryogenic System. A Generally Applicable Approach to Solute Trapping and Mobilization in Gas Chromatography," *Anal. Chem.*, vol. 69, no. 13, 1997, doi: https://doi.org/10.1021/ac961310w.

[72] P. H. Stefanuto and J. F. Focant, "Columns and column configurations," in *Basic Multidimensional Gas Chromatography*, Elsevier, 2020, ch. Chapter 3. doi: 10.1016/B978-0-12-813745-1.00003-9.

[73] A. Mostafa, M. Edwards, and T. Górecki, "Optimization aspects of comprehensive two-dimensional gas chromatography," *J. Chromatogr. A*, vol. 1255, pp. 38–55, Sep. 2012, doi: https://doi.org/10.1016/j.trac.2011.06.010.

[74] H. Boswell, H. Y. Chow, and T. Gorecki, "Modulators," in *Basic Multidimensional Gas Chromatography*, Elsevier, 2020, pp. 89–140. doi: 10.1016/B978-0-12-813745-1.00004-0.

[75] P. Q. Tranchida, G. Purcaro, P. Dugo, and L. Mondello, "Modulators for comprehensive two-dimensional gas chromatography," *TrAC - Trends in Analytical Chemistry*, vol. 30, no. 9, Oct. 2011, doi: 10.1016/j.trac.2011.06.010.

[76] H. D. Bahaghighat, C. E. Freye, and R. E. Synovec, "Recent advances in modulator technology for comprehensive two dimensional gas chromatography," *TrAC - Trends in Analytical Chemistry*, vol. 113, Apr. 2019, doi: 10.1016/j.trac.2018.04.016.

[77] M. Edwards, A. Mostafa, and T. Górecki, "Modulation in comprehensive two-dimensional gas chromatography: 20 years of innovation," *Anal. Bioanal. Chem.*, vol. 401, no. 8, Nov. 2011, doi: 10.1007/s00216-011-5100-6.

[78] J. V. Seeley, N. J. Micyus, S. V. Bandurski, S. K. Seeley, and J. D. McCurry, "Microfluidic deans switch for comprehensive two-dimensional gas chromatography," *Anal. Chem.*, vol. 79, no. 5, pp. 1840–1847, Mar. 2007, doi: 10.1021/ac061881g.

[79] W. Khummueng, J. Harynuk, and P. J. Marriott, "Modulation ratio in comprehensive two-dimensional gas chromatography," *Anal. Chem.*, vol. 78, no. 13, p. 4587, Jul. 2006, doi: 10.1021/ac052270b.

[80] R. C. Y. Ong and P. J. Marriott, "A Review of Basic Concepts in Comprehensive Two-Dimensional Gas Chromatography," *J. Chromatogr. Sci.*, vol. 40, 2002, doi: 10.1093/chromsci/40.5.276.

[81] P. J. Marriott, R. M. Kinghorn, R. Ong, P. Morrison, P. Haglund, and M. Harju, "Comparison of thermal sweeper and cryogenic modulator technology for comprehensive gas chromatography," *Journal of High Resolution Chromatography*, vol. 23, no. 3, 2000, doi: 10.1002/(SICI)1521-4168(20000301)23:3%3C253::AID-JHRC253%3E3.0.CO;2-I.

[82] J. Beens, M. Adahchour, R. J. Vreuls, K. van Altena, and U. A. Th Brinkman, "Simple, non-moving modulation interface for comprehensive two-dimensional gas chromatography," *J. Chromatogr. A*, 2001, doi: 10.1016/S0021-9673(01)00785-3.

[83] M. Adahchour, J. Beens, and U. A. T. Brinkman, "Single-jet, single-stage cryogenic modulator for comprehensive two-dimensional gas chromatography (GC x GC)," *Analyst*, vol. 128, no. 3, Mar. 2003, doi: 10.1039/b212657p.

[84] E. B. Ledford and C. Billesbach, "Jet-cooled thermal modulator for comprehensive multidimensional gas chromatography," *Journal of High Resolution Chromatography*, vol. 23, no. 3, 2000, doi: 10.1002/(SICI)1521-4168(20000301)23:3<202::AID-JHRC202>3.0.CO;2-5.

[85] R. K. Nelson *et al.*, "Applications of Comprehensive Two-Dimensional Gas Chromatography (GC × GC) in Studying the Source, Transport, and Fate of Petroleum Hydrocarbons in the Environment," in *Standard Handbook Oil Spill Environmental Forensics: Fingerprinting and Source Identification*, Elsevier, 2016. doi: http://dx.doi.org/10.1016/B978-0-12-809659-8.00008-5.

[86] P. Q. Tranchida and L. Mondello, "Detectors and basic data analysis," in *Basic Multidimensional Gas Chromatography*, vol. 12, Elsevier, 2020, p. 227. doi: 10.1016/B978-0-12-813745-1.00006-4.

[87] P. Q. Tranchida, L. Mondello, S. D. H. Poynter, and R. A. Shellie, "Comprehensive Two-Dimensional Gas Chromatography Combined with Mass Spectrometry," in *Comprehensive Chromatography in Combination with Mass Spectrometry*, John Wiley & Sons, 2011, ch. 6, pp. 171–242. doi: https://doi.org/10.1002/9781118003466.ch6.

[88] P. Q. Tranchida and L. Mondello, "Comprehensive 2D Gas Chromatography," in *Hyphenations of Capillary Chromatography with Mass Spectrometry*, Elsevier, 2019, ch. 2.2. doi: 10.1016/B978-0-12-809638-3.00005-3.

[89] M. S. Eschner, W. Welthagen, T. M. Gröger, M. Gonin, K. Fuhrer, and R. Zimmermann, "Comprehensive multidimensional separation methods by hyphenation of single-photon ionization time-of-flight mass spectrometry (SPI-TOF-MS) with GC and GC×GC," *Anal. Bioanal. Chem.*, vol. 398, no. 3, Oct. 2010, doi: 10.1007/s00216-010-4021-0.

[90] E. Hoffmann and V. Stroobant, *Mass Spectrometry - Principles and Applications*, Third Edition. John Wiley & Sons Inc., 2007. ISBN: 978-0-470-03310-4

[91] J. H. Gross, *Mass Spectrometry*, Second Edition. Springer, 2017. doi: 10.1007/978-3-642-10711-5.

[92] D. J. Harvey, "Mass spectrometric detectors for gas chromatography," in *Gas Chromatography*, Second Edition., C. F. Poole, Ed., in Handbooks in Separation Science. , Amsterdam: Elsevier, 2021, ch. Chapter 14, pp. 399–424. doi: https://doi.org/10.1016/B978-0-12-820675-1.00022-8.

[93] A. Samokhin, "Spectral skewing in gas chromatography–mass spectrometry: Misconceptions and realities," *J. Chromatogr. A*, vol. 1576, Nov. 2018, doi: 10.1016/j.chroma.2018.09.033.

[94] M. Kirchner, E. Matisová, S. Hrouzková, and J. De Zeeuw, "Possibilities and limitations of quadrupole mass spectrometric detector in fast gas chromatography," *J. Chromatogr. A*, vol. 1090, no. 1–2, Oct. 2005, doi: 10.1016/j.chroma.2005.06.090.

[95] A. Fellinger, "Deconvolution of Overlapping Skewed Peaks," *Anal. Chem.*, vol. 66, no. 4, 1994, doi: https://doi.org/10.1021/ac00091a013.

[96] M. P. B. Mourão, A. H. J. Kolk, and H. G. Janssen, "General principles and history," in *Hyphenations of Capillary Chromatography with Mass Spectrometry*, Elsevier, 2019, ch. Chapter 1.1. doi: 10.1016/B978-0-12-809638-3.00001-6.

[97] P. Q. Tranchida, F. A. Franchina, P. Dugo, and L. Mondello, "Comprehensive two-dimensional gas chromatography-mass spectrometry: Recent evolution and current trends," *Mass Spectrom. Rev.*, vol. 35, no. 4, pp. 524–534, Jul. 2016, doi: 10.1002/mas.21443.

[98] I. Špánik and A. Machyňáková, "Recent applications of gas chromatography with high-resolution mass spectrometry," *J. Sep. Sci.*, vol. 41, no. 1, Jan. 2018, doi: 10.1002/jssc.201701016.

This page was intentionally left blank

Chapter 4

Non-targeted terpenoid profiling in Cannabis inflorescences by sorbent-based sampling with two-stage thermal desorption coupled to GC×GC-TOFMS*

4.1. Introduction

Cannabis sativa L. (Cannabis) is an extensively cultivated species belonging to the *Cannabaceae* family. Historically, in addition to representing a source of psychoactive substances, which makes it one of the world's most widely used recreational drugs, Cannabis has been, and still is, extensively used not only as a raw material for the textile and flavor industries [1, 2], but also in various applications as food products and medicinal formulations [3].

Similar to other natural products, many constituents of interest are included, owing to their flavor and biological activity. Terpenes or isoprenoids represent the largest class of secondary metabolites in Cannabis and are mostly responsible for its characteristic aroma. These volatile molecules are additionally involved in various bio-mechanisms, such as plant defense or plant-to-plant communication, also exhibiting various therapeutic properties [4].

Chemically, this class of hydrocarbons derives from the fusion of several isoprene units of C_5H_8 and is classified based on the number of fundamental units, grouping them in monoterpenes (*e.g.*, camphene, pinene, and myrcene), sesquiterpenes (*e.g.*, caryophyllene, humulene, and bisabolene), and diterpenes (*e.g.*, phytane and taxadiene) [5]. Terpenes can also be referred to as terpenoids when containing heteroatoms (*i.e.,* oxygen).

A wide spectrum of biological activities can be attributed to the class of terpenes and terpenoids, among which the viricidal, antidepressant, antidiabetic, analgesic, and antiplasmodial activities are attributed to β-pinene, carvone, limonene, caryophyllene, and linalool [6–9]. Even if not fully understood, these pharmacological effects depend on the synergistic interactions (also known as an "entourage effect") with other terpenoids as well as cannabinoids [10, 11]**.**

Studying and profiling terpenes and terpenoids in Cannabis, as well as in other natural sources, could lead to the discovery of new therapeutic molecules, either directly or through the "entourage effect". Furthermore, as a raw material for the medical, fragrance, and food

*This section has been adapted from: **De Poli M.**, Chenet T., Felletti S., Spadafora M., Cavazzini A., Franchina F.A.; "Sorbent-Based Sampling With Two-Stage Trapping/Desorption Coupled to Comprehensive Two Dimensional Gas Chromatography and Mass Spectrometry for Terpenoids Profiling in Cannabis"; Analytical Science Advances 6 (2025), e202400044.

chemical elucidation becomes important to verifying feedstock, identifying potential adulteration, and ensuring product safety. The determination of terpenoids in Cannabis, as in other plant raw materials, is commonly carried out with one-dimensional gas chromatography (GC) with a flame ionization detector [12]. However, conventional GC methods often lack separation power for the full molecular separation of the sample constituents, especially when considering the complex composition of natural samples like Cannabis and the variety of terpenoids known [13].

The use of comprehensive multidimensional GC (GC×GC), thanks to its increased separation power, selectivity, and sensitivity (see Chapter 3, Section 3.3), can be considered the technique of choice for detailed molecular sample characterization, known as "profiling" (see Chapter 2, Section 2.2.3).

In the context of non-targeted analysis, exploring chemical space requires analytical workflows that preserve the inherent complexity of the sample while minimizing bias introduced during preparation. In the case of Cannabis inflorescences, whose volatile fraction comprises a heterogeneous mixture of compounds, the choice of sampling strategy plays a critical role, determining which portion of this chemical space becomes effectively accessible [14]. In this sense, besides the separation technique, different extraction methods exist, and the choice depends on the sample type and objective(s) of the study [15–18]. As for many other plant products, headspace extraction (HS) is the logical and preferred choice for flavor profiling. Opposite to liquid extraction, HS is based on solvent-free extractions, which can be exploited through static or dynamic mode and mediated by a solid sorbent or not [19].

Thermal desorption tubes represent an advantageous solution in this context, as they integrate sampling, analyte enrichment, and subsequent thermal release into a single device that interfaces with a GC×GC system. This capability enhances analytical information continuity (AIC; see Chapter 2, Section 2.5), ensuring the chemical diversity present in the sample is directly transferred into the chromatographic system. Moreover, the reusability of these devices supports the inclusiveness required by modern non-targeted analytical workflows [20, 21].

In the present research, the terpenes and terpenoids profile in Cannabis inflorescences was addressed by developing and optimizing a dynamic headspace method using adsorption tubes thermally desorbed into a GC×GC-MS system. Initially, the extraction efficiency of the groups of interest was assessed using probe standards, employing traps packed with different adsorbents (Tenax-TA, Carbotrap T420, Carbotrap 202) and a two-stage thermal desorption setup. Once the optimal experimental conditions were identified, the terpenoid profiles of commercial samples were assessed, highlighting differences in chemical composition and the advantages of employing multidimensional GC separation.

4.2. Materials and methods

4.2.1. Chemicals, standards, and samples

A test standard mixture of 12 compounds (Supelco, Bellefonte, PA, USA) was initially used to evaluate the secondary trapping/release stage with a cold trap. To evaluate the extraction performance of the different thermal desorption tubes, a mixture of 37 probe compounds was used (ID and GC×GC chromatograms shown in Figure 4.2).

Regarding the real-world samples, 4 CBD-dominant chemovar samples of *Cannabis sativa* inflorescences were obtained from a local store (Ferrara, Italy) and transferred to a brown glass jar for storage at room temperature in dark conditions.

4.2.2. Sampling and extraction

Dynamic headspace extraction was used for both the standard mix and Cannabis inflorescences. Briefly, 2 µL of the probe mix was spiked into a 20 mL air-tight headspace glass vial and conditioned for 5 minutes at 40 °C. A syringe was then used to manually draw 30 mL of room air through the vial headspace and onto the thermal desorption tubes at approximately 60 mL/min.

Regarding the samples, 20 mg of homogenized Cannabis inflorescences were placed into a 20 mL air-tight headspace glass vial. Sample conditioning and dynamic extraction conditions were the same as for the standard mix. Three and four technical replicates were performed on separate thermal desorption tubes (TDT) for the standard mix and the Cannabis samples, respectively. For chemovar #2 sampled on Tenax-TA, one replicate was excluded due to a clear technical failure (abnormally low total ion current), and the reported data for this chemovar are based on three valid replicates.

After each cycle of extraction/analysis, the tubes were conditioned using a Phoenix T220 (REDshift s.r.l., San Giorgio in Bosco, Padua, Italy) at 30 °C/min to 320 °C (held for 60 min) under pure N_2 flow (100 mL/min). Tube blanks were conducted periodically to exclude carry-over.

Regarding the TDT, 3 different sorbent combinations, packed as single or multibed, were evaluated. Specifically, these were Tenax-TA (poly-(2,6-diphenyl-p-phenylenoxide)), Carbotrap T420 (Carbotrap F + Carbotrap Y), Carbotrap 202 (Carbopack C + Carbopack B); for brevity, in the text and figures, these thermal desorption tubes are labelled as TA, CT420, and C202, respectively. All TDTs were obtained from Supelco (Merck Life Science, Darmstadt, Germany).

Initially, the extraction performance of each tube was investigated using the fragrance standard mix; subsequently, only the tube types with higher extraction capacity and reproducibility were used to evaluate the terpenoid profile in real samples (*i.e.*, Cannabis inflorescences).

4.2.3. Instrumental and experimental conditions

The development of the GC×GC-MS method was conducted on a Pegasus BT 4D (LECO Corporation, Mönchengladbach, Germany) equipped with an Agilent 8890 GC and an automated tube handling device PAL System (CTC Analytics AG, Zwingen, Switzerland). The chromatographic columns were a 30 m × 0.25 mm × 0.25 µm df Rxi-5ms as the first dimension and a 2 m × 0.25 mm × 0.25 µm df Rxi-17SilMS as the second dimension (both from Restek Corporation, Bellefonte, PA, USA). Injection was performed using an OPTIC-4 multimode injector (GL Sciences B.V., Eindhoven, Netherlands) equipped with an inlet Peltier cooler and a liner exchanger (GL Sciences B.V.). The GC was also equipped with a cryogenic trap (GL Sciences B.V.) as a secondary trapping/release stage. The carrier gas used was helium at 1.5 mL/min column flow throughout the run. The vent time (60 sec) before thermal desorption consisted of a split flow of 50 mL/min and a column flow of 0.5 mL/min, with the inlet body held at 20°C. During thermal desorption, the inlet was heated from 20 °C to 300 °C at 20 °C/s, using a split flow of 150 mL/min and restoring the initial 1.5 mL/min column flow. The cryogenic trap was set at -20°C throughout the vent time (60 s) and the thermal desorption (45 s), after which its temperature was raised to 250 °C at 50 °C/s to allow the secondary release of the analytes.

The oven temperature program was 40°C (held for 1 min), then ramped at 4 °C/min to 190 °C, and finally ramped at 30 °C/min to 320 °C (held for 3 min). Temperature offsets for the secondary oven and modulator oven were set to +5 °C and +15 °C, respectively. A 3 s modulation period was used (cold and hot jets were 0.7 and 0.8 s, respectively). A mass range

from 40 to 500 *m/z* was collected with an acquisition frequency of 150 Hz. An acquisition delay of 180 s was used. The transfer line and ion source temperatures were set at 280 °C and 250 °C, respectively.

For the cryogenic trap evaluation, 2 µL of Grob mixture (10 ppm in hexane) was spiked on a TA tube. Then, flushed at room temperature at 100 mL/min for 60 s. Chromatographic parameters were assessed under single and dual-stage thermal desorption conditions in one-dimensional gas chromatography.

Data were collected and analyzed using ChromaTOF software version 5.56 (LECO Corporation), and NIST23 was used as the mass spectral library. The signal-to-noise threshold for peak integration was set at 10, and putative identification was based on spectral similarity (≥70%) and linear retention indices (RI) with a tolerance of ±30 units.

4.3. Results and discussion

4.3.1. Optimization of thermal desorption and sorbent selection

The instrumental configuration for thermal desorption used a multimode injector with an automated tube exchanger. Such a configuration avoids the use of an external unit, making the system more compact and requiring fewer hardware components, providing a direct, more efficient injection into the GC column head. Nevertheless, the overall chromatography can benefit from an additional trapping/release stage using a cryotrap. Therefore, the first step in the method development was to evaluate the cryotrap conditions.

For such a purpose, a test standard mixture (*i.e.*, Grob mix) was used, and full width at half height, tailing factor, peak height, and resolution were evaluated with ("trap active") and without the trap ("trap not active"). During thermal desorption, analytes are ideally transferred from the tube into the GC column as a narrow band, and the presence of an additional focusing system (dual-stage thermal desorption) reduces both band broadening and breakthrough of highly volatile compounds, thereby improving the final separation [20]. Not surprisingly, the secondary trapping/release stage reduced the average full width at half-height (FWHM) from 2.9 to 2.1 s. At the same time, the tailing factor dropped from an average of 1.6 to 1.5, and an average increase of 28% in peak intensity was also observed.

Figure 4.1 shows the overlapped 1D chromatogram with and without the trap, in which the chromatographic improvement can be observed in terms of peak shape, peak height, and resolution between some critical pairs. The use of the secondary trapping/release stage increased the chromatographic resolution between the undecane/nonanal and 2,6-dimethyl-phenol/2-ethyl-hexanoic acid considerably by a factor of ×1.7 and ×2.5, respectively. Such advantages observed in one-dimensional chromatography are also beneficial in two-dimensional separations.

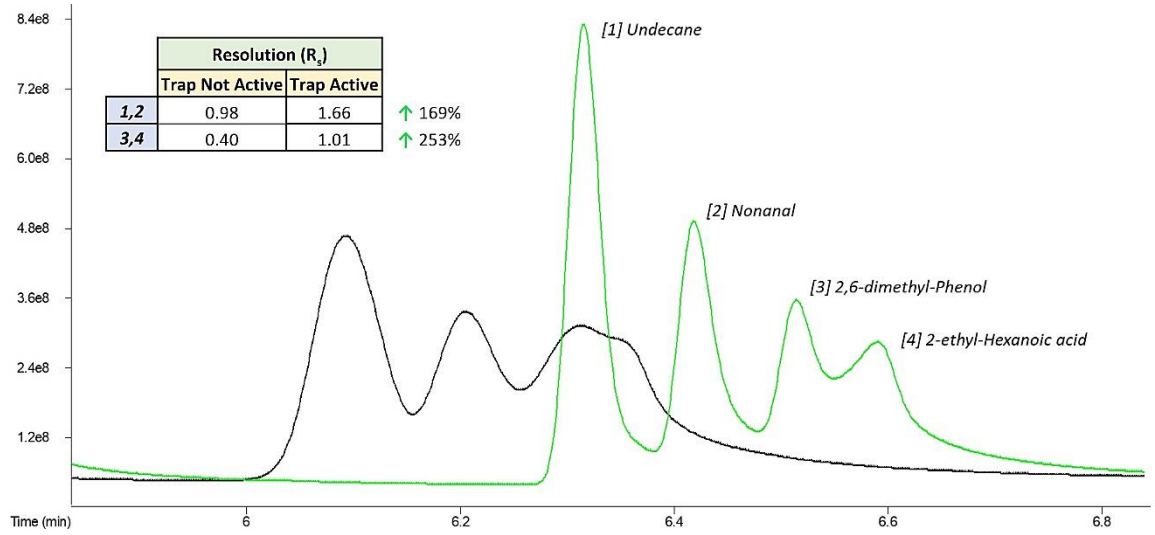
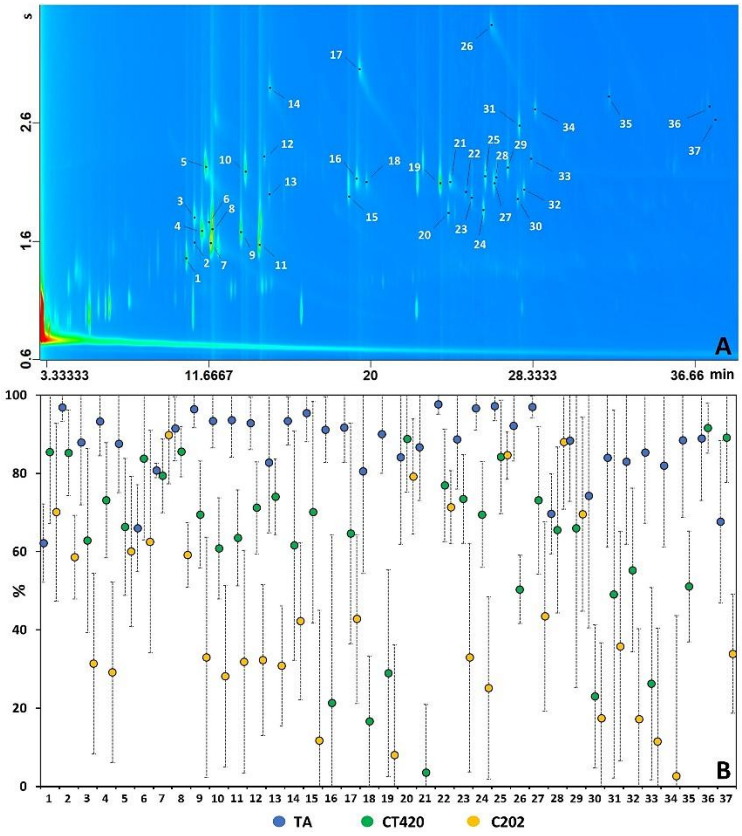


Figure 4.1. *Zoomed GC chromatogram of critical pairs obtained with single-stage (black trace) and dual-stage thermal desorption (green trace).*

Before sampling real Cannabis inflorescences, we evaluated the extraction efficiencies of 3 differently packed thermal desorption tubes (TDTs) on probe analytes. Such a probe mixture was selected to match the range and class of compounds expected in the Cannabis inflorescences, covering different degrees of analytes' polarity and boiling point. In addition to terpenes (mono- and sesqui-), it contained other aldehydes, ketones, and alcohols (listed in Figure 4.2), with varying polarities and boiling points. Dynamic extraction was performed under the same conditions (see Materials and Methods) while varying the TDT types. The separation of the 37 analytes resulting from the optimized thermal desorption (TD-)GC×GC conditions is shown in Figure 4.2A.

The response observed with the different tube types is variable, and it is related to the properties, thus the selectivity of the adsorbent materials [22–24].

In terms of analyte signals, C202 tubes showed substantially lower signals than TA and CT420, indicating lower extraction efficiency (Figure 2B). Even though TA tubes showed a better overall response, the CT420 tubes were comparable for some terpenes (*e.g.*, β-myrcene, *o*-cymene, eucalyptol, limonene). It is also worth noting that the TA's extraction performance was 33.2% higher than the CT420's, based on the average of the 37 probe analytes. Therefore, TA and CT420 tubes were used to further investigate real-world samples.



Peak	ID	Characteristic *m/z*
1	β-Myrcene	93
2	Pseudolimonen	93
3	Octanal	43
4	Acetic acid, hexyl ester	43
5	Methyl p-cresyl ether	122
6	o-Cymene	119
7	D-Limonene	68
8	Eucalyptol	43
9	Dihydromyrcenol	59
10	Cyclal C	67
11	Tetrahydro linalool	73
12	Tricyclal	67
13	Fenchol	81
14	Phenylethyl Alcohol	91
15	Citronellol	69
16	Allyl (3-methylbutoxy)acetate	43
17	Methyl phenethyl ketone	43
18	Nerol	69
19	Vertenex	57
20	β-Patchoulene	161
21	Lavandulyl acetate	69
22	Cycloseychellene	123
23	Caryophyllene	93
24	α-Guaiene	105
25	Seychellene	122
26	Yara yara	115
27	α-Patchoulene	93
28	γ-Patchoulene	121
29	α Isomethyl ionone	135
30	β-Bisabolene	69
31	Cashmeran	191
32	α-Bulnesene	93
33	α-N-Methyl ionone	121
34	Lilial	189
35	α-pentyl-Cinnamaldehyde	91
36	Versalide	243
37	Tonalid	243

Figure 4.2. *A) GC×GC plot (TIC) of probe fragrance mixture using TA as trapping material. B) Analyte response and reproducibility (error bars, n=3) with thermal desorption tubes (blue dots = TA; green dots =CT420; yellow dots =C202. C) Y-scale is normalized for each analyte to the highest signal among the three tube types. For peak numbering, refer to the peak list on the right.*

4.3.2. Volatile organic compounds profiling of different cannabis chemovars

The volatile profiles of four different Cannabis chemovars' inflorescences were extracted using TA and CT420 tubes, thanks to their satisfactory extraction performance, as demonstrated by the probe mixture and discussed in the previous paragraph.

Figure 4.3A shows a representative two-dimensional chromatogram obtained using TA tubes. The use of a non-polar primary column combined with a polar secondary column results in a structured separation in which the two main terpene classes, *i.e.*, monoterpenes and sesquiterpenes, appear clustered. This pattern is also observed in conventional one-dimensional GC analyses; however, GC×GC enhances class separation and peak organization of complex terpene mixtures. For the rapid classification of these terpenoids, two elution regions were drawn based on the RI information (considering both experimental and literature values) and the characteristic fragments (*e.g.*, 69, 93, 105, 121, 136 *m/z*, etc.). These two elution regions are circled and spanned from an RI of 877 (*i.e.*, the monoterpenes cyclofenchene) to 1810 (*i.e.*, the sesquiterpenoid nootkatone [25]).

Considering the large number of compounds present in Cannabis inflorescences, GC×GC certainly proves to be a suitable separation technique. Some examples of peak pairs that would be otherwise co-eluted in conventional one-dimensional separation are shown in Figure 4.3B and 4.3C. In the case of the monoterpenes Terpinolene and Fenchone (Figure 4.3B), the resulting chromatographic resolution improved from 0.24 to 1.51, respectively in 1D and 2D. Similarly, the sesquiterpenes 3,7(11)-Selinadiene and α-Bisabolene gained a resolution of 1.14 using GC×GC, in contrast to 0.19 in 1D (Figure 4.3B). One more benefit

arising from the improved GC×GC separation is evident in Figures 4.3B and 4.3C, where the purity of mass spectra results in higher spectral matches with reference databases.

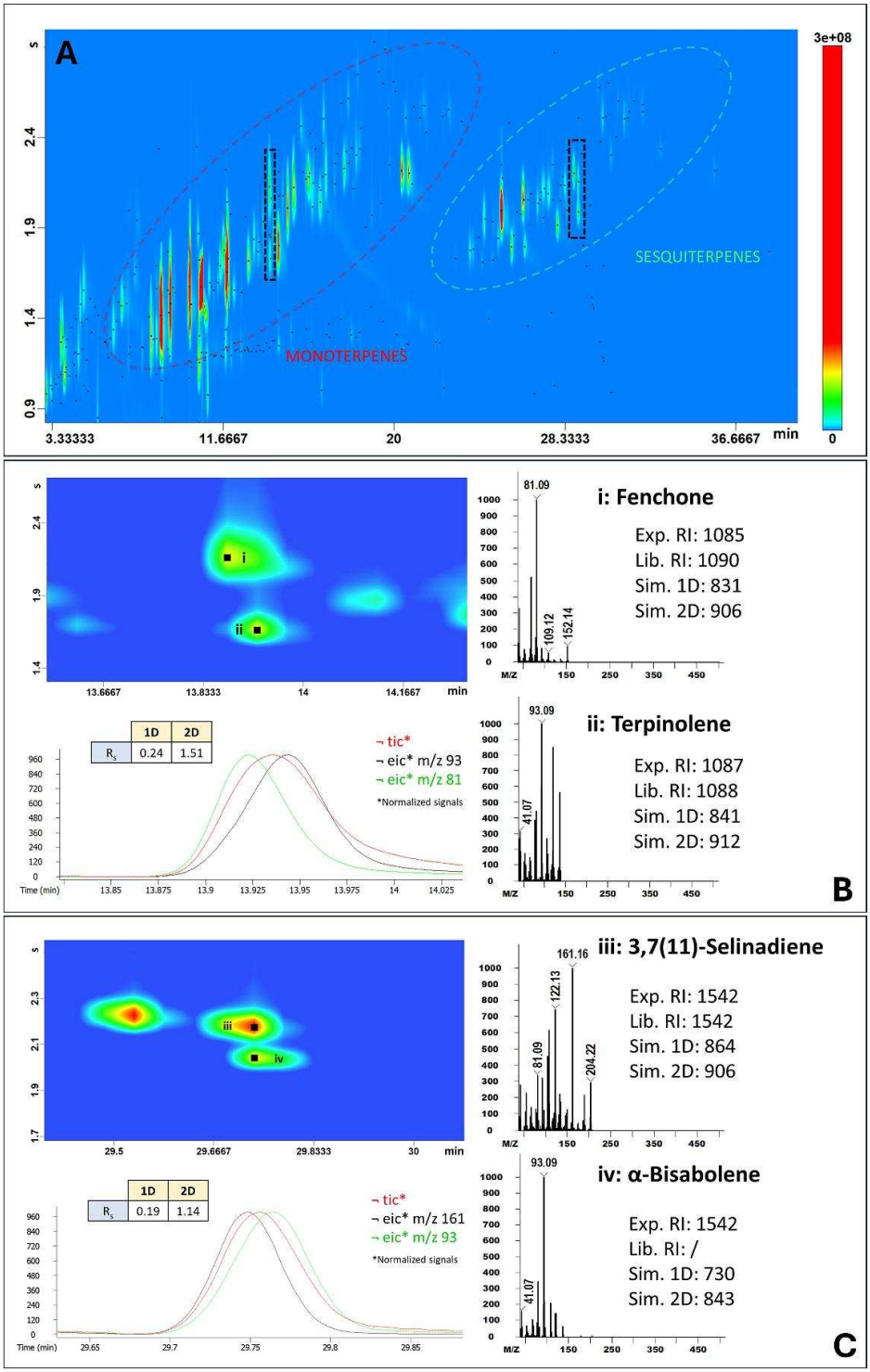


Figure 4.3. *A) 2D elution regions of terpene classes. B &C) Chromatographic resolution comparison of selected terpenes between 2D and 1D, with relative spectral and RI information.*

On average, considering the two terpenoid elution regions, 183 analytes were detected in the sample headspace using TA.

To compare the performance of the two selected TDTs, 75 putatively identified terpenes were selected from 4 different Cannabis chemovars (spectral similarity (Sim) > 70%; linear

retention indices ±30 units). These compounds are listed in Table 4.1 and visually highlighted in Figure 4.4.

Table 4.1. *List of the 75 putatively identified terpenes in Cannabis. [1]tr (min): Retention time in the first dimension; [2]tr (s): Retention time in the second dimension; Similarity: Spectral similarity; Exp. RI: Experimental Retention Index; Lib. RI: Library Retention Index.*

#	Compounds	$^{1}t_r$ (min)	$^{2}t_r$ (s)	Similarity	Exp. RI	Lib. RI
1	Cyclofenchene	7.05	1.20	906	877	890
2	Bornene <2->	7.75	1.24	892	903	908
3	Santolinatriene	8.15	1.31	795	915	902
4	Thujene <alpha->	8.40	1.29	904	923	929
5	Pinene <alpha->	8.60	1.42	910	929	937
6	Camphene	9.10	1.49	943	945	952
7	Dehydrosabinene	9.30	1.54	805	951	956
8	Pinene <beta->	10.00	1.58	855	972	979
9	Myrcene <beta->	10.55	1.58	862	989	991
10	Phellandrene <alpha->	11.00	1.60	863	1003	1005
11	Carene <3->	11.20	1.56	736	1009	1011
12	Terpinene <alpha->	11.41	1.60	794	1014	1017
13	Cymene <ortho->	11.51	1.69	703	1017	1024
14	Cymene <para->	11.71	1.78	892	1023	1025
15	Limonene	11.86	1.66	811	1027	1029
16	Ocimene <cis-, beta->	12.21	1.56	882	1037	1035
17	Ocimene <trans-, beta->	12.56	1.54	726	1047	1046
18	Terpinene <gamma->	12.91	1.72	809	1057	1058
19	Sabinene hydrate <cis->	13.16	1.84	727	1064	1069
20	Fenchone	13.91	2.13	848	1086	1090
21	Terpinolene	13.96	1.73	876	1087	1086
22	Epoxymyrcene <6,7->	14.16	1.86	807	1093	1096
23	Linalool	14.36	1.80	883	1099	1099
24	Fenchol <endo->	14.81	2.01	882	1111	1119
25	Linalool <alpha->	15.11	2.12	760	1120	1099
26	Pinocarveol <trans->	15.71	2.17	841	1137	1139
27	Isocitral <exo->	15.81	2.18	713	1140	1144
28	Ipsdienol	16.01	2.02	741	1146	1146

29	Terpineol <beta-, trans->	16.06	2.15	753	1147	1169
30	Pinene oxide <beta->	16.41	2.04	762	1157	1156
31	Pinocarvone	16.56	2.50	877	1161	1164
32	Isoborneol	16.66	2.19	850	1164	1165
33	Terpinenol <4->	17.11	2.14	747	1177	1177
34	Terpineol <alpha->	17.56	2.23	876	1190	1189
35	Myrtenol	17.76	2.26	840	1196	1195
36	Myrtenal	17.76	2.61	913	1196	1193
37	Verbenone	18.21	2.75	909	1209	1228
38	Carveol	18.56	2.35	758	1219	1219
39	Fenchyl acetate	18.61	1.91	827	1220	1219
40	Citronellol	18.91	1.97	875	1229	1228
41	Carvone	19.46	2.70	753	1245	1242
42	Pinocarvyl acetate <trans->	21.26	2.51	771	1297	1297
43	Myrtenyl acetate	21.81	2.39	710	1314	1327
44	Ylangene	23.71	1.78	901	1373	1372
45	Copaene	23.86	1.78	920	1378	1376
46	Bourbonene <beta->	24.11	1.85	779	1386	1384
47	Sesquithujene <7-epi->	24.31	1.71	913	1392	1394
48	Isocomene <beta->	24.41	1.95	849	1395	1412
49	Sesquithujene	24.81	1.73	872	1408	1402
50	Caryophyllene <cis->	24.81	1.94	929	1408	1406
51	cis-α-Bergamotene <cis-, alpha->	25.06	1.85	900	1416	1415
52	Caryophyllene	25.21	2.00	948	1421	1419
53	Bergamotene <trans-, alpha->	25.71	1.80	913	1437	1435
54	Guaiene <alpha->	25.81	1.86	874	1440	1439
55	Aromandendrene	26.16	1.98	856	1452	1440
56	Humulene	26.31	2.06	901	1457	1454
57	Farnesene <trans-, beta->	26.36	1.80	910	1458	1452
58	Alloaromadendrene	26.51	2.06	895	1463	1461
59	Acoradiene <beta->	26.61	2.06	801	1466	1471
60	Muurolene <gamma->	27.01	2.04	893	1479	1478
61	Bergamotene <beta-, trans->	27.26	1.96	857	1487	1483
62	Eudesmene <beta->	27.31	2.12	906	1489	1486

63	Valencene	27.51	2.09	908	1495	1492
64	Muurolene <alpha->	27.71	2.06	874	1502	1497
65	Bisabolene <beta->	27.96	1.91	935	1510	1509
66	Cadinene <gamma->	28.11	2.14	865	1516	1512
67	Guaiene <cis-beta->	28.32	2.14	848	1522	1498
68	Calamenene	28.37	2.34	857	1524	1523
69	Sesquiphellandrene <beta->	28.42	1.99	803	1526	1524
70	Selinadiene <4(15),7(11)>	28.77	2.21	874	1538	1540
71	Selinadiene <3,7(11)>	28.97	2.16	903	1545	1542
72	Caryophyllene oxide	30.17	2.51	840	1586	1587
73	Acorenol <beta->	31.22	2.50	815	1623	1649
74	Bisabolol <alpha->	32.97	2.34	913	1686	1684
75	Isovalencenol	35.92	2.77	787	1796	1788

Of the 75 putatively identified compounds, 43 were classified as monoterpenes and 32 as sesquiterpenes.

It must be highlighted that additional compounds were detected, eluting within these regions (*e.g.*, on sample#1, 45 compounds within the monoterpene and 45 compounds within the sesquiterpenes region), and are marked with red and green dots on the 2D plot in Figure 4.4. Considering the spectral fragmentation patterns, these can reasonably be other terpenoids that were not included in the comparison due to lower identification levels (Sim<70% and ΔRI>30 or absent).

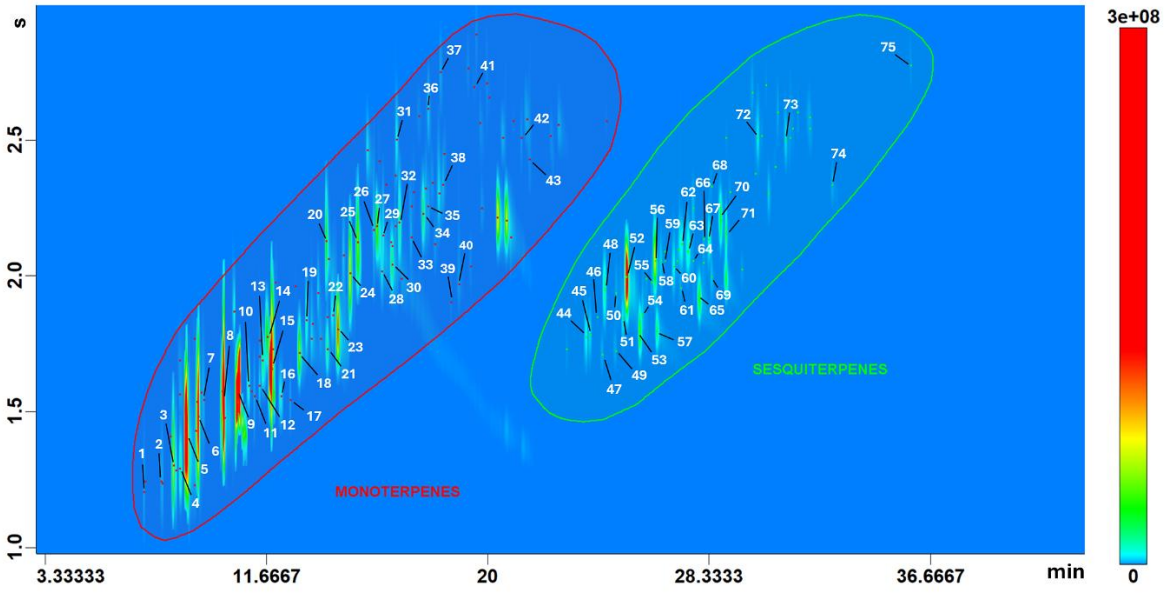


Figure 4.4. *Putatively identified compounds on cannabis inflorescence (sample#1) with Tenax-TA; S/N>100.*

The overall extraction performance of the tubes, in terms of total response and the number of detected peaks, for the terpenoids across the four chemovars, is shown in Figure 4.5. On

the one hand, this response indicates the extraction efficiency of the trapping materials; on the other hand, the number of detected peaks highlights the selectivity and analyte coverage of the trapping materials towards terpenes. In addition to differences in the distribution of monoterpenes and sesquiterpenes among chemovars, a consistent trend in extraction performance is evident. Focusing on the difference between TA and CT420 in Figure 4.5A, the former slightly outperforms in terms of signal, and more importantly, in terms of reproducibility (on average 5.4 and 7.5 for monoterpenes and sesquiterpenes for TA *vs* 7.9 and 9.7, respectively, for CT420). Another observation is highlighted in Figure 4B: the number of detected peaks (S/N > 10) is shown for the two different extraction TD tubes. In this sense, tubes packed with Tenax yielded more peaks than CT420, confirming greater universality and broader selectivity.

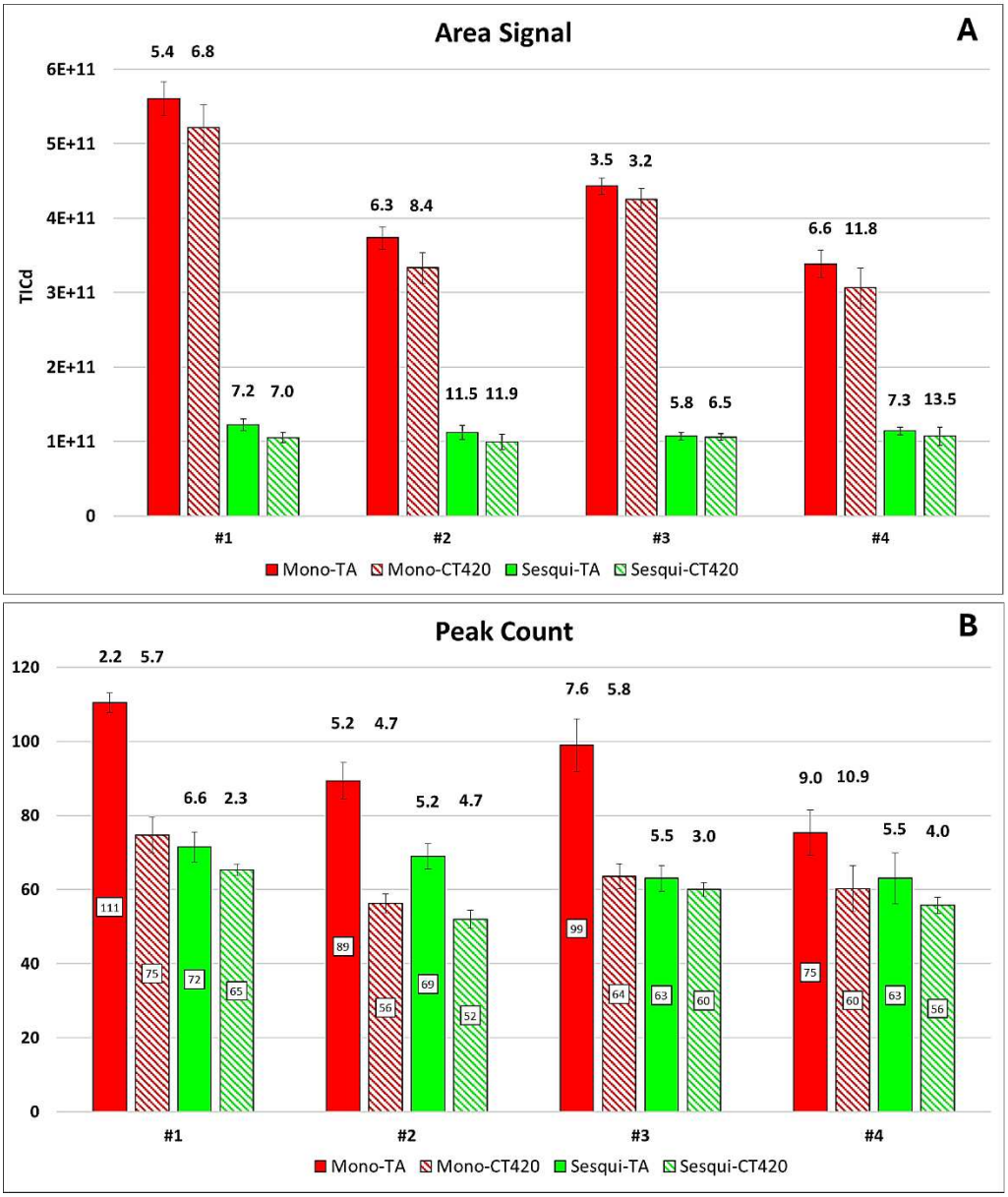


Figure 4.5. *A) Total signal area comparison on monoterpenes and sesquiterpenes between Tenax-TA and Carbotrap T420. B) Tube collection efficiency on peak count, where error bars represent the standard deviation (n = 3), which is also reported numerically on top of each bar.*

Considering only Tenax-TA extractions, the detailed distribution of the 75 terpenoids across the four chemovars is shown in the heatmap in Figure 4.6A. In this representation, missing values were replaced with the dataset's minimum value, and features were normalized using

an auto-scaling procedure (*i.e.*, mean-centered and divided by each variable's standard deviation). Each column corresponds to a putatively identified substance, while each row represents a sample within a chemovar group. The color scale (from dark blue to yellow) reflects low-to-high relative abundances after auto-scaling, and the dendrograms highlight natural similarity patterns among both chemovars and compounds.

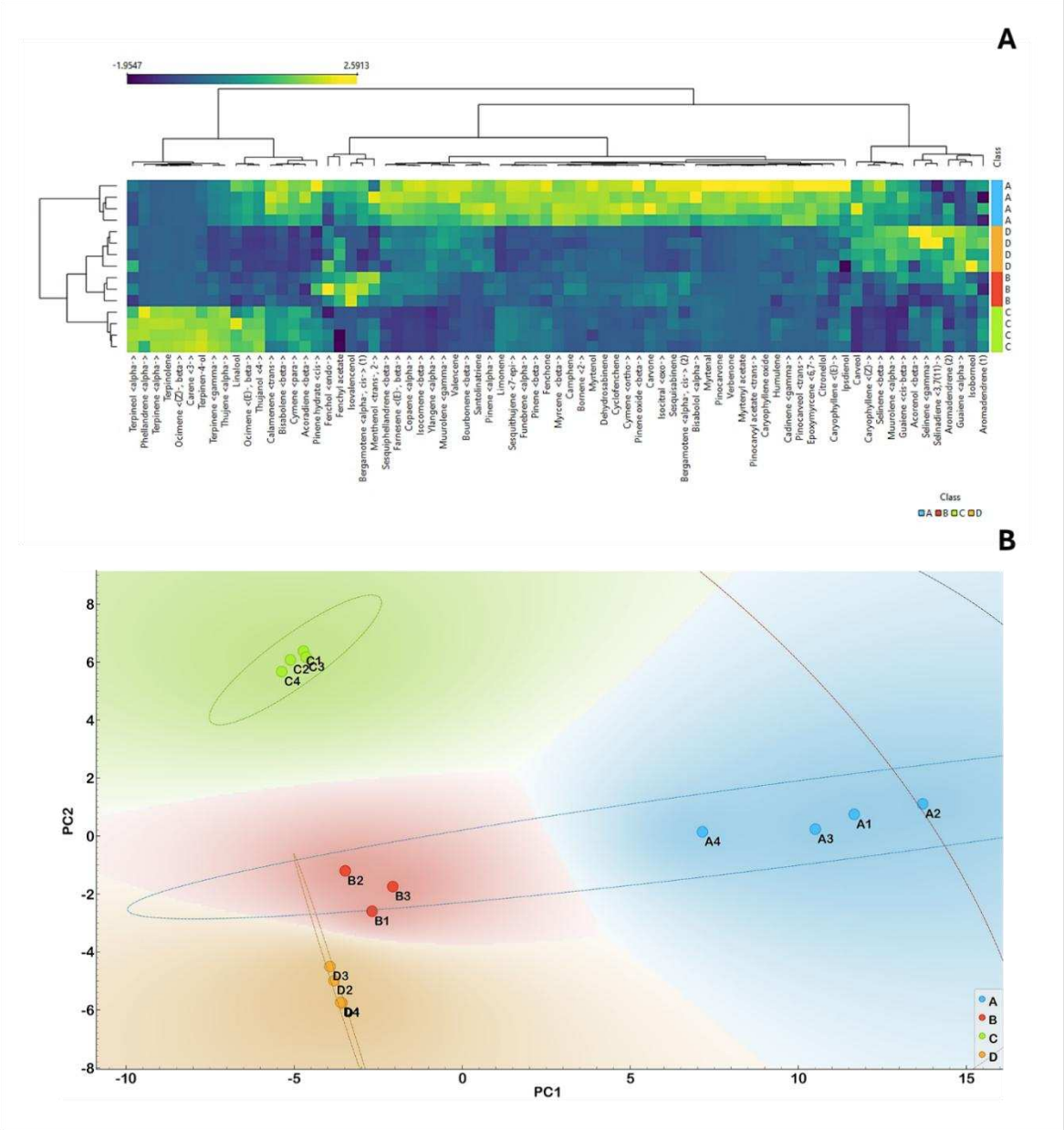


Figure 4.6. *A) Heatmap with respective hierarchical cluster analysis of the 75 putative compounds using Tenax-TA. B) Score plot made by principal component analysis of the tentatively identified substances using Tenax-TA.*

In Figure 4.6A, a clear separation among chemovars is evident, with samples within the same group clustering closely, indicating consistent group volatilomic profiles. Differences in the abundance and distribution of mono- and sesquiterpenes across chemovars produce distinct color patterns, underscoring Tenax-TA's ability to capture a broad portion of the volatilome. As highlighted by the intense yellow color, chemovar A represents the most flavoring species, followed by chemovar C and then B, D. Also, explorative principal component analysis (PCA) was performed, and the score plot is presented in Figure 4.6B. Here, each sample is plotted in the space of the first two principal components (PCs), and confidence

ellipses emphasize the within-group consistency. Four different group clusters can be seen, highlighting how the volatile components allows a clear distinction between chemovars.

Moreover, the broader selectivity of TA tubes discussed earlier was also observed through the detection of compounds not extracted by CT420 tubes. An example of compounds uniquely extracted and detected using TA is shown in Figure 4.7. These two compounds elute between the monoterpene and sesquiterpene regions, are absent in the blanks, and do not meet reliable identification criteria (Sim>70% and ΔRI<30). The mass spectrum of these analytes, which are probably not included in the MS database, is also presented in Figure 4.7.

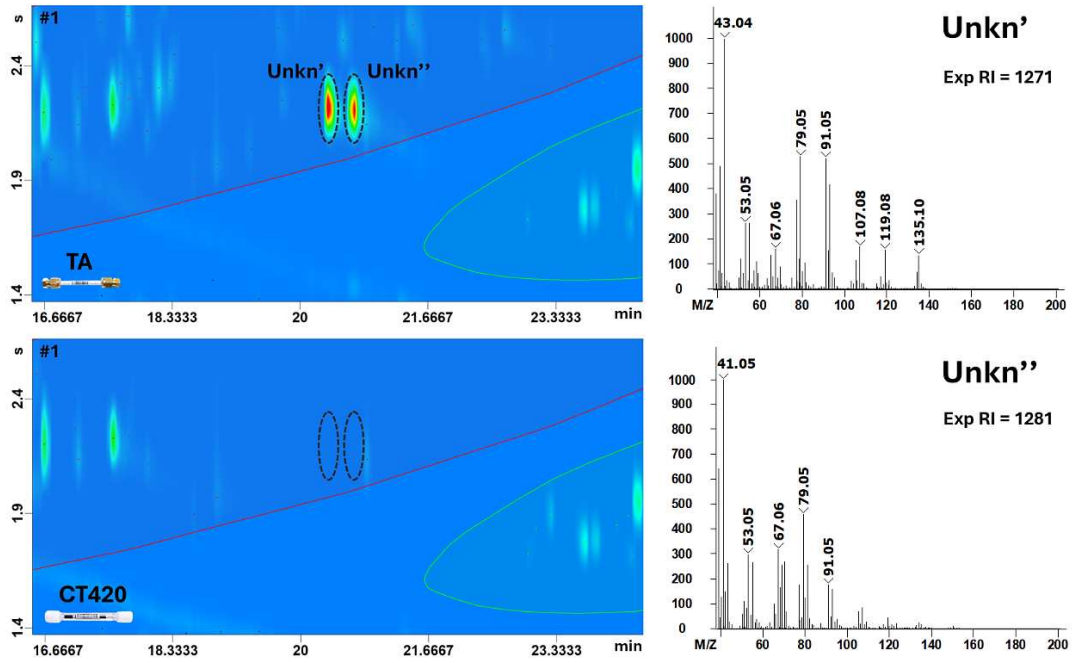


Figure 4.7. *Elution region and mass spectra of two unidentified analytes (Unkn' and Unkn''), uniquely extracted and detected when using Tenax-TA tubes.*

4.4. Conclusions

This study revealed how the use of dynamic headspace extraction followed by thermal desorption represents an effective and non-invasive approach for the investigation of volatile profiles of Cannabis inflorescences, which can be advantageous for on-field applications. Using both probe standard analytes and real-world samples, we showed that the secondary trapping/release stage implemented herein greatly improves chromatographic performance (*i.e.*, peak quality and resolution).

For the choice of sampling material, we observed that either the Tenax-TA (porous polymer) or the Carbotrap T420 (a combination of two graphitized carbon blacks) tubes are suitable for terpenoid profiling in plants, although the former provides better reproducibility. However, when the scope of the investigation extends beyond a targeted or profiling approach, the use of the porous polymer provides broader analyte coverage and is thus more suitable for non-targeted purposes.

Regarding the separation technique, we showed that comprehensive two-dimensional gas chromatography coupled to mass spectrometry enables deep molecular characterization of Cannabis with enhanced resolution and identification reliability, highlighting flavor differences across chemovars.

References

[1] J. Fike, "The history of hemp," in *Industrial Hemp as a Modern Commodity Crop*, Wiley, 2019, pp. 1–25. doi: 10.2134/industrialhemp.c1.

[2] A. W. Zuardi, "History of cannabis as a medicine: a review," *Braz J Psychiatry*, pp. 153–7, 2006.

[3] E. B. Russo, "History of cannabis and its preparations in saga, science, and sobriquet," 2007. doi: 10.1002/cbdv.200790144.

[4] D. Cox-Georgian, N. Ramadoss, C. Dona, and C. Basu, "Therapeutic and medicinal uses of terpenes," in *Medicinal Plants: From Farm to Pharmacy*, Springer International Publishing, 2019, pp. 333–359. doi: 10.1007/978-3-030-31269-5_15.

[5] S. R. Sommano, C. Chittasupho, W. Ruksiriwanich, and P. Jantrawut, "The Cannabis Terpenes," Dec. 01, 2020, *MDPI*. doi: 10.3390/MOLECULES25245792.

[6] C. B. Duschatzky *et al.*, "Evaluation of Chemical and Antiviral Properties of Essential Oils from South American Plants," *Antivir Chem Chemother*, vol. 16, no. 4, pp. 247–251, Aug. 2005, doi: 10.1177/095632020501600404.

[7] S. L. Guzmán-Gutiérrez, H. Bonilla-Jaime, R. Gómez-Cansino, and R. Reyes-Chilpa, "Linalool and β-pinene exert their antidepressant-like activity through the monoaminergic pathway," *Life Sci*, vol. 128, pp. 24–29, May 2015, doi: 10.1016/j.lfs.2015.02.021.

[8] M. Jung, M. Park, H. Lee, Y.-H. Kang, E. Kang, and S. Kim, "Antidiabetic Agents from Medicinal Plants," *Curr Med Chem*, vol. 13, no. 10, pp. 1203–1218, Apr. 2006, doi: 10.2174/092986706776360860.

[9] C. R. Nogueira and L. M. X. Lopes, "Antiplasmodial Natural Products," *Molecules*, vol. 16, no. 3, pp. 2146–2190, Mar. 2011, doi: 10.3390/molecules16032146.

[10] E. B. Russo, "Taming THC: potential cannabis synergy and phytocannabinoid-terpenoid entourage effects," *Br J Pharmacol*, vol. 163, no. 7, pp. 1344–1364, Aug. 2011, doi: 10.1111/j.1476-5381.2011.01238.x.

[11] D. Namdar *et al.*, "Terpenoids and Phytocannabinoids Co-Produced in Cannabis Sativa Strains Show Specific Interaction for Cell Cytotoxic Activity," *Molecules*, vol. 24, no. 17, p. 3031, Aug. 2019, doi: 10.3390/molecules24173031.

[12] A. Leghissa, Z. L. Hildenbrand, and K. A. Schug, "A review of methods for the chemical characterization of cannabis natural products," Jan. 01, 2018, *Wiley-VCH Verlag*. doi: 10.1002/jssc.201701003.

[13] E. M. Humston-Fulmer, D. E. Alonso, and J. E. Binkley, "Improving cannabis differentiation by expanding coverage of the chemical profile with GCxGC-TOFMS," in *Comprehensive Analytical Chemistry*, vol. 90, Elsevier B.V., 2020, pp. 169–196. doi: 10.1016/bs.coac.2020.04.007.

[14] G. Black *et al.*, "Exploring chemical space in non-targeted analysis: a proposed ChemSpace tool," *Anal Bioanal Chem*, vol. 415, no. 1, pp. 35–44, Jan. 2023, doi: 10.1007/s00216-022-04434-4.

[15] G. Micalizzi, F. Vento, F. Alibrando, D. Donnarumma, P. Dugo, and L. Mondello, "Cannabis Sativa L.: a comprehensive review on the analytical methodologies for

cannabinoids and terpenes characterization," *J Chromatogr A*, vol. 1637, p. 461864, Jan. 2021, doi: 10.1016/j.chroma.2020.461864.

[16] F. A. Franchina, D. Zanella, L. M. Dubois, and J. F. Focant, "The role of sample preparation in multidimensional gas chromatographic separations for non-targeted analysis with the focus on recent biomedical, food, and plant applications," *J Sep Sci*, p. jssc.202000855, Oct. 2020, doi: 10.1002/jssc.202000855.

[17] F. A. Franchina, L. M. Dubois, and J. F. Focant, "In-Depth Cannabis Multiclass Metabolite Profiling Using Sorptive Extraction and Multidimensional Gas Chromatography with Low- A nd High-Resolution Mass Spectrometry," *Anal Chem*, vol. 92, no. 15, pp. 10512–10520, Aug. 2020, doi: 10.1021/acs.analchem.0c01301.

[18] S. Louw, "Recent trends in the chromatographic analysis of volatile flavor and fragrance compounds: Annual review 2020," *Analytical Science Advances*, vol. 2, no. 3–4, pp. 157–170, Apr. 2021, doi: 10.1002/ansa.202000158.

[19] E. Liberto, C. Bicchi, C. Cagliero, C. Cordero, P. Rubiolo, and B. Sgorbini, "Chapter 1: Headspace sampling: An 'evergreen' method in constant evolution to characterize food flavors through their volatile fraction," in *Food Chemistry, Function and Analysis*, vol. 2020-Janua, no. 17, Royal Society of Chemistry, 2020, pp. 3–37. doi: 10.1039/9781788015752-00001.

[20] E. Woolfenden, "Sorbent-based sampling methods for volatile and semi-volatile organic compounds in air. Part 1: Sorbent-based air monitoring options," *J Chromatogr A*, vol. 1217, no. 16, pp. 2674–2684, Apr. 2010, doi: 10.1016/j.chroma.2009.12.042.

[21] E. Woolfenden, "Sorbent-based sampling methods for volatile and semi-volatile organic compounds in air. Part 2. Sorbent selection and other aspects of optimizing air monitoring methods," *J Chromatogr A*, vol. 1217, no. 16, pp. 2685–2694, Apr. 2010, doi: 10.1016/j.chroma.2010.01.015.

[22] K. Dettmer and W. Engewald, "Adsorbent materials commonly used in air analysis for adsorptive enrichment and thermal desorption of volatile organic compounds," *Anal Bioanal Chem*, vol. 373, no. 6, pp. 490–500, Jul. 2002, doi: 10.1007/s00216-002-1352-5.

[23] F. A. Franchina, G. Purcaro, A. Burklund, M. Beccaria, and J. E. Hill, "Evaluation of different adsorbent materials for the untargeted and targeted bacterial VOC analysis using GC×GC-MS," *Anal Chim Acta*, vol. 1066, pp. 146–153, Mar. 2019, doi: 10.1016/j.aca.2019.03.027.

[24] F. A. Franchina, D. Zanella, T. Dejong, and J.-F. Focant, "Impact of the adsorbent material on volatile metabolites during in vitro and in vivo bio-sampling," *Talanta*, vol. 222, p. 121569, Jan. 2021, doi: 10.1016/j.talanta.2020.121569.

[25] V. I. Babushok, P. J. Linstrom, and I. G. Zenkevich, "Retention Indices for Frequently Reported Compounds of Plant Essential Oils," *J Phys Chem Ref Data*, vol. 40, no. 4, Dec. 2011, doi: 10.1063/1.3653552.

This page was intentionally left blank

Chapter 5

Investigation of solvent-dependent metabolomic profiles in human serum *via* comprehensive two-dimensional gas chromatography and mass spectrometry

5.1. Introduction

Metabolomics is a research domain that aims to comprehensively identify small metabolites (*i.e.,* molecules with MW < 1000 Da) in biological systems, providing a snapshot of the physiological status of an organism of interest [1]. This discipline, when applied to biological fluids (*e.g.*, serum, plasma, urine), enables the simultaneous analysis of hundreds of compounds, capturing both endogenous metabolites produced by the inherent metabolic pathways and exogenous compounds absorbed primarily through the diet [2]. In this way, metabolomics provides a detailed picture of how biochemical pathways shift in response to genetic inputs and how environmental and lifestyle factors influence the biochemical equilibrium of an organism [3, 4]. Moreover, given that sample collection is relatively non-invasive and relies on straightforward procedures, this domain is particularly suitable for large-scale studies aimed at characterizing diseases in their early stages [5, 6].

As the downstream product of gene expression and protein activity, metabolomics is a valuable tool for identifying diagnostic or predictive biomarkers of various disorders [7]. For example, numerous studies present in the literature have examined correlations between metabolic profiles and multiple types of cancer [8–14]. Beyond oncology, metabolomics has also proven effective in metabolomic profiling of Alzheimer's disease, childhood obesity, COVID-19 infection, maternally inherited diabetes, and rheumatoid arthritis [15–19]. All this scientific evidence highlights the capacity of this discipline to capture specific biochemical perturbations.

In the context of human blood, the Human Metabolome Database (HMDB) is currently the most comprehensive repository of metabolites, reporting thousands of low-molecular-weight compounds detected in this matrix [20]. The full chemical diversity of human blood, however, remains far from being completely characterized [21]. In general, biological fluids derived from blood (*i.e.*, serum and plasma) contain metabolites that can be broadly classified into two main categories: water-soluble and lipid-soluble compounds, with the latter accounting for approximately 88% of the total metabolome [22, 23]. Lipid molecules satisfy numerous biological functions, acting as membrane constituents, signaling mediators, and long-term energy reservoirs. In contrast, non-lipid metabolites account for only about 12% of the blood metabolome and span a wider variety of chemical classes, including amino acids, carbohydrates, organic acids, and other small molecules. Blood metabolites also span

an extraordinary concentration range, covering several orders of magnitude, with highly abundant species such as cholesterol and urea present at millimolar levels, while certain phospholipids occur at picomolar concentrations [22]. Moreover, metabolic concentrations vary substantially among healthy individuals, reflecting both internal factors (*e.g.*, age, sex) and external influences (*e.g.*, diet, physical activity) [24]. Despite extensive characterization efforts, current databases represent only an estimated 20% of the complete human metabolome, highlighting that the complexity of serum composition remains to be revealed [1]. This diversity presents challenges across many aspects of metabolomic workflows, from extraction to data processing and, finally, biological interpretation.

Analytical techniques in metabolomics rely on high-throughput platforms capable of profiling large numbers of metabolites within complex biological matrices [25, 26]. Gas chromatography (GC) and liquid chromatography (LC), particularly when hyphenated to mass spectrometry (MS), are among the most widely used techniques for exploring metabolomes [27]. However, GC-based analysis is inherently limited to volatile compounds, as many metabolites possess polar functional groups and exhibit insufficient volatility. To overcome this limitation, chemical derivatization is typically employed to modify functionalities such as -OH, -NH_2, -COOH, and -SH, thereby reducing polarity, breaking intramolecular hydrogen bonding, and increasing volatility [28]. The most common procedure involves a two-step process of methoximation followed by trimethylsilylation, although alternative derivatization strategies, such as acylation or esterification, exist for specific metabolite classes [29, 30]. Trimethylsilylation, however, remains a benchmark for non-targeted GC-MS metabolomics due to its broad applicability across diverse chemical classes [31]. Further, considering the separation capacity for complex mixtures, comprehensive two-dimensional gas chromatography (GC×GC) has been increasingly adopted in metabolomics [32]. Compared with conventional one-dimensional GC, GC×GC typically enables the detection of several-fold more chromatographic features, enhancing the chemical space of the metabolome coverage [33, 34].

One of the main challenges in metabolomics analysis is the lack of standardized procedures across sample collection, chemical analysis, and data interpretation. Every step in the analytical workflow directly impacts the result of the study, shaping the accessible chemical space of the biological sample through multiple analytical choices [35]. In this context, sample preparation represents the main bottleneck in metabolomics analysis, which typically involves three stages: sample collection, metabolite extraction, and preparation for instrumental analysis [28]. Concerning the second step, selecting the extraction solvent is particularly critical, as it must efficiently transfer metabolites from the serum matrix into the solvent phase. Additionally, this process not only depends on the solvent's ability to solubilize compounds across a broad polarity spectrum but should also be as straightforward and rapid as possible to facilitate large-scale high-throughput studies and avoid degradation during sample handling [36]. Ideally, the extraction solvent should act as a non-selective extractant capable of recovering metabolites across diverse chemical classes [37]. Several studies and established protocols have suggested different solvent systems for serum/plasma metabolite extraction, with much of the heterogeneity arising from the use of sequential extraction steps, binary and ternary solvent combinations, as well as the number of internal standards or solvent volumes [38].

Dunn *et al.* proposed in 2011 a detailed protocol for large-scale metabolic profiling of serum/plasma using methanol, while Fiehn subsequently proposed in 2016 a workflow for gas chromatography-mass spectrometry of mammalian samples using a ternary mixture of isopropanol, acetonitrile, and water [39, 40]. Moreover, Stephens *et al.* and He *et al.* both found that ternary mixtures of methanol, acetonitrile, and acetone provide a strong balance between feature numbers, analytical stability, and signal intensity [41, 42]. Other solvent

combinations have also been explored, including mixtures of methanol, chloroform, methyl tert-butyl ether, ethanol, and water [38]. However, given the wide variety of chemicals available, there is currently no solvent system that is more widely used than others.

In light of this consideration, the aim of this study was to compare the extraction performance of three solvent systems, *i.e.,* methanol, isopropanol:acetonitrile:water 3:3:2 (v/v/v), and methanol:acetonitrile:acetone 1:1:1 (v/v/v) by GC×GC-TOFMS. A non-targeted metabolomic workflow was developed and optimized to maximize analytical accuracy while maintaining practical efficiency and time productivity. The different solvent mixtures were evaluated for reproducibility, the number of detected peaks, and extraction efficiency across the identified metabolite classes. This work aims to demonstrate how variations in solvent partitioning shift the extracted metabolic classes, highlighting that the overall composition of the extraction solvent directly affects the scientific outcome.

5.2. Materials and methods

5.2.1. Chemicals, standards, and solvents

Three solvent mixtures were investigated based on previous literature research [39, 40, 42]: Methanol (CH_3OH), Isopropanol + Acetonitrile + Water (C_3H_8O:CH_3CN:H_2O, 3:3:2; v/v/v), Methanol + Acetonitrile + Acetone (CH_3OH:CH_3CN:$(CH_3)_2CO$, 1:1:1; v/v/v). For concision, these mixtures are reported as solvent A, B, and C along with the text and figures. All solvents were purchased from Merck Life Science (Darmstadt, Germany).

Derivatization was performed using methoximation followed by silylation. The first step involved methoxyamine hydrochloride (MeOX, 20 mg/mL in pyridine), followed by silylation with N-methyl-N-(trimethylsilyl)trifluoroacetamide (MSTFA). Both reagents and pyridine were purchased at high purity and derivatization grade from Supelco (Merck Life Science, Darmstadt, Germany). Marked Glucose-d7 was also employed as an internal standard supplied by Sigma-Aldrich (Merck Life Science, Darmstadt, Germany).

The metabolomics profile of a human serum derived from male AB donors was investigated with the selected solvent mixture purchased from Sigma-Aldrich (Merck Life Science, Darmstadt, Germany). Method blanks were prepared to assess background signals arising from solvents, reagents, and consumables. For each blank, an empty 1.5 mL Eppendorf tube (containing no human serum) was subjected to the complete extraction and derivatization workflow, following the same steps as for the real samples, starting with the addition of the relative extraction solvent mixture.

Extractions for each solvent mixture were analyzed in triplicate. In order to monitor instrumental stability and potential analytical drift throughout the batch, after completing sample preparation, 10 µL from each sample replicate was combined into a single homogenized quality control (QC) serum mixture. This pooled QC sample was then injected at regular intervals during the analytical sequence.

5.2.2. Extraction and derivatization

Extraction and derivatization were carried out using the following non-targeted GC×GC-MS workflow for metabolomics analysis. Serum samples were stored at −80 °C until processed. After thawing on ice at 4 °C for 1 hour, 50 µL were transferred into 1.5 mL tubes, followed by the addition of 10 µL of glucose-d7 solution and 150 µL of cold extraction solvent. Samples were then vortexed for 1 min, shaken for 5 min, and centrifuged at 14,000 rpm for 5 min to promote protein precipitation and metabolite extraction. A total of 150 µL of the resulting supernatant was transferred into clean microtubes and evaporated to complete dryness using a SpeedVac concentrator. Dried extracts were then subjected to the two-step derivatization. First, 20 µL of MeOX solution was added, and samples were shaken at

maximum speed in an orbital plate device system at 25°C for 1.5 hours. Then, 60 µL of MSTFA was added, and tubes were agitated at maximum speed for 0.5 hour at 25°C. Finally, derivatized solutions were transferred to vials with glass inserts and analyzed by GC×GC-MS.

5.2.3. Instrumental experimental conditions

GC×GC-MS analyses were conducted on a Pegasus BT 4D (LECO Corporation, Mönchengladbach, Germany) equipped with an Agilent 8890 GC and an automated tube handling device PAL System (CTC Analytics AG, Zwingen, Switzerland). The chromatographic columns were a 30 m × 0.25 mm × 0.25 µm df Rxi-5ms as the first dimension and a 2 m × 0.25 mm × 0.25 µm df Rxi-17SilMS as the second dimension (both from Restek Corporation, Bellefonte, PA, USA). Injection was performed using an OPTIC-4 multimode injector (GL Sciences B.V., Eindhoven, Netherlands). The carrier gas used was helium at 1.5 mL/min column flow throughout the run. The inlet temperature was set to 280 °C, and the septum purge flow was set to 3 mL/min. 0.5 µL of the final derivatization solution was injected into the chromatographic system in split mode with a split ratio of 1:10.

The oven temperature program was 50 °C (held for 1 min), then ramped at 4 °C/min to 330 °C (held for 5 min). Temperature offsets for the secondary oven and modulator oven were set to +5 °C and +15 °C, respectively. A 3.5s modulation period was used (cold and hot jets were 0.7 and 1.05 sec, respectively). A time-of-flight mass analyzer was used to acquire a mass range of 50-500 *m/z* at 150 Hz, using electron ionization (70 eV). The ion source was maintained at 250 °C, while the transfer line was set to 280 °C. Finally, an acquisition delay of 438 s was used.

Data were collected and analyzed using ChromaTOF software version 5.56 (LECO Corporation), and NIST23 was used as the mass spectral library. The signal-to-noise threshold for peak integration was set at 10, and putative identification was based on spectral similarity (≥70%), blank subtraction, and coherent position in the two-dimensional chromatogram. Finally, predicted logP values of the selected metabolites were retrieved from PubChem (XlogP3) and ChemSpider (ACD/LogP). When available, experimentally determined logP values reported in the databases were considered for comparison [43, 44].

5.3. Results and discussion

5.3.1. Putative selection of metabolomic features

Considering the three identified solvents, all extractions were performed in triplicate and subsequently injected into the GC×GC-MS system. Quality control (QC) pools and method blanks were injected periodically across the analytical sequence to monitor the reproducibility and assess the stability of the derivatization process. The internal standard signal was monitored in both pooled QCs and individual samples, showing fluctuations within the coefficient of variation of less than 30%, confirming good analytical stability and the absence of systematic drift throughout the sequence.

After data acquisition, chromatograms were processed by applying a signal-to-noise ratio threshold of 10, followed by library searches to identify the detected peaks. By performing blank subtraction, 395 metabolites were detected on average in solvent A, 338 in solvent B, and 398 metabolites in solvent C. In order to compare the extraction behavior of the selected solvents, 79 metabolites were identified using methanol (MeOH) as the reference, with spectral similarity (*i.e.,* Sim > 750), absent in the blanks, and a reliable position in the two-dimensional chromatographic plane. Selected compounds were screened in all the replicates, and reproducibility within the same solvent was assessed. Reproducibility and the number of detected components across all replicates are visualized in ring charts in Figure 5.1.

Consistent detection of chromatographic peaks was achieved, with 79 compounds identified for solvents A and 78 for solvents B and C. Three color scales are shown, each representing a different RSD% interval. As observed, all solvents exhibited good reproducibility, with solvent C showing a fraction of metabolites with coefficients of variation (CV) below10%, exceeding 70% of the putatively identified compounds.

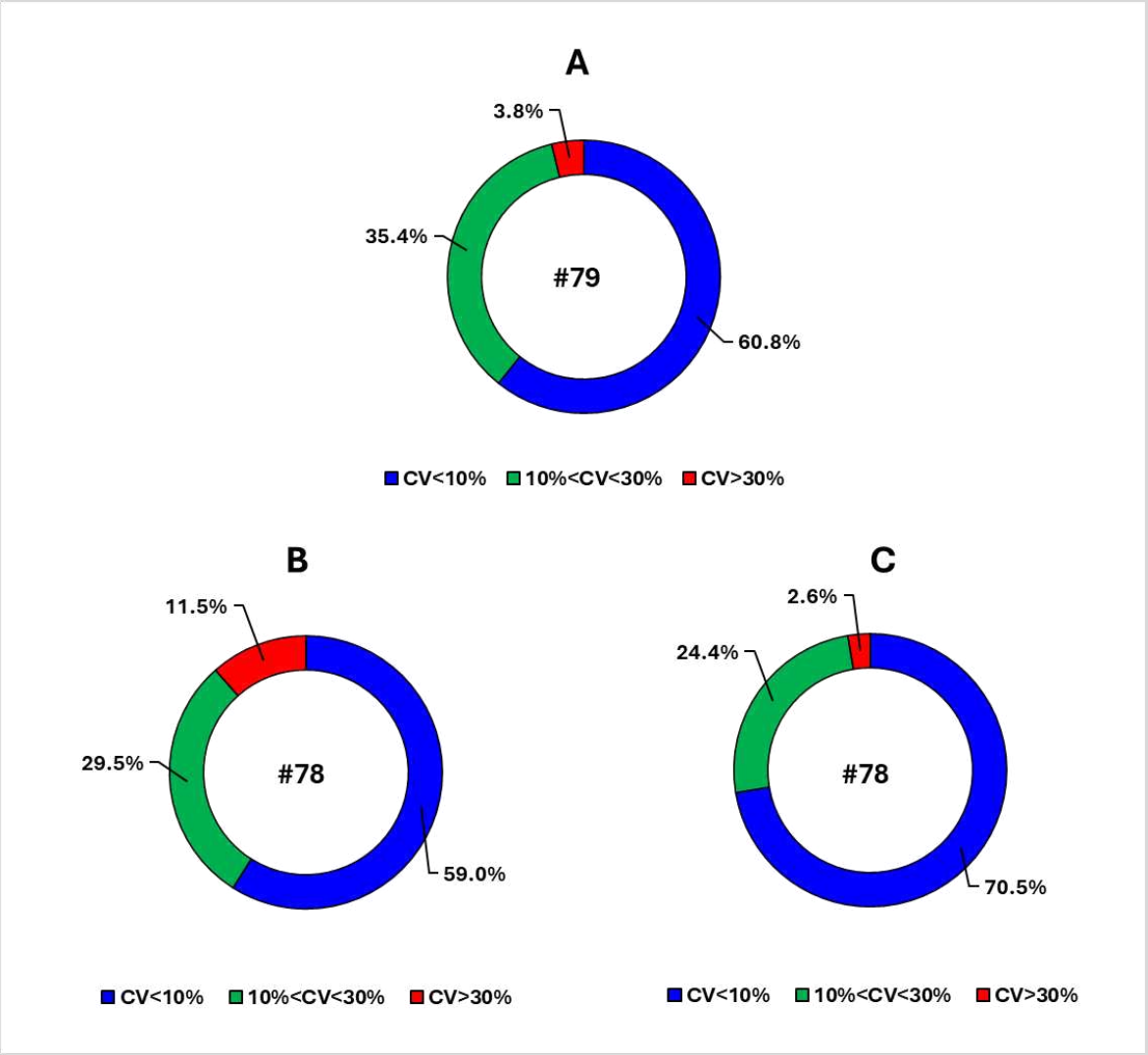


Figure 5.1. *Reproducibility ring charts of the detected features. Metabolites were selected according to spectral similarity > 75%, absence in blank samples, and consistent position in the 2D chromatographic plane (i.e., false positive subtraction). Three color scales are presented: blue = CV < 10%, green = 10%<CV<30% and red = CV>30%.*

From the initial dataset of 79 tentatively identified metabolites, compounds with CV greater than 30% in at least one extraction solvent were removed to ensure analytical robustness. This filtering step yielded a refined set of 66 metabolites reproducibly detected across all extraction conditions. In the context of non-targeted metabolomics, the detected metabolites represent only a subset of the theoretical chemical space of human serum. The spatial distribution observed in the GC×GC chromatogram reflects the portion of chemical space that becomes analytically accessible under the selected extraction and analytical conditions. The resulting compounds are displayed in the two-dimensional GC×GC contour plot in Figure 5.2, referred to as a methanol extraction replica. The use of two orthogonal separation mechanisms reduces co-elution among many metabolites, as exemplified by compounds 25 and 26, which are clearly resolved within their respective elution region, highlighted in red in Figure 5.2. Structured chromatograms also enable identification of trends and patterns

within metabolite groups. While many metabolite classes are widely spread across the chromatographic plane, amino acids, polyalcohols & sugars, and sterols are distributed into confined elution regions. Although clear, sharply defined boundaries between metabolite classes are not present, the separation power of comprehensive GC×GC is evident from these spatial trends. In one-dimensional GC, similar class-related trends may only be partially observable and are often hidden by extensive coelution and limited peak capacity.

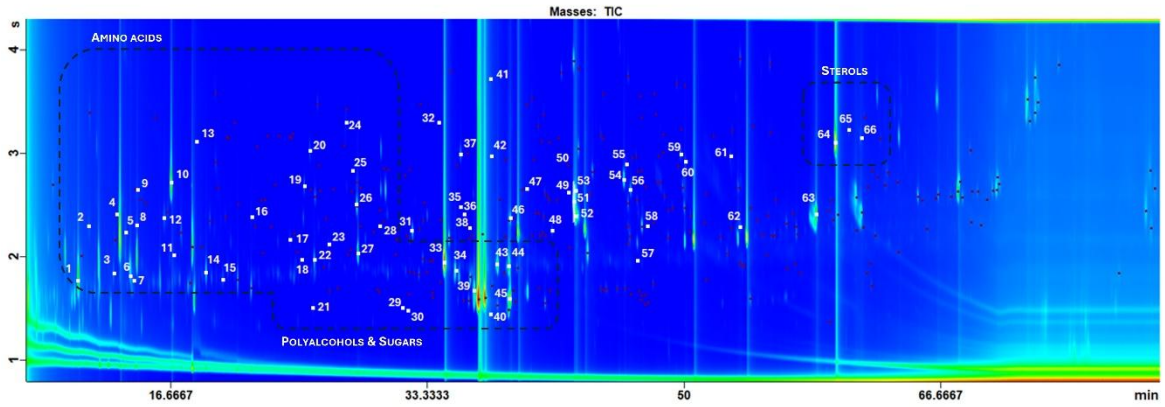


Figure 5.2. *Distribution of the selected 66 metabolites in the 2D chromatogram with a characteristic trend of metabolite classes.*

As shown in Figure 5.2, several metabolites exhibit particularly intense responses, as indicated by the red and yellow dots in the contour plot. Despite their strong signal intensity, these features did not meet the predefined confidence criteria and were therefore excluded from the final dataset.

Five distinct classes were identified among the 66 metabolites based on functional groups and chemical structures. The identified classes broadly reflect the metabolic profile of human serum and include amino acids (AA), organic acids (OA), amino-functionalized compounds (AM, i.e., structures containing an amino group but not classified as amino acids), polyols and sugars, and finally, sterols (ST). The pie chart in Figure 5.3 summarizes the distribution of compounds within identified classes, with organic acids representing more than 50% of the features (36 out of 66 compounds), followed by amino acids (17%), polyalcohols & sugars (14%), amino-functionalized compounds (11%), and finally sterols, accounting for three metabolites corresponding to approximately 5% of the overall distribution.

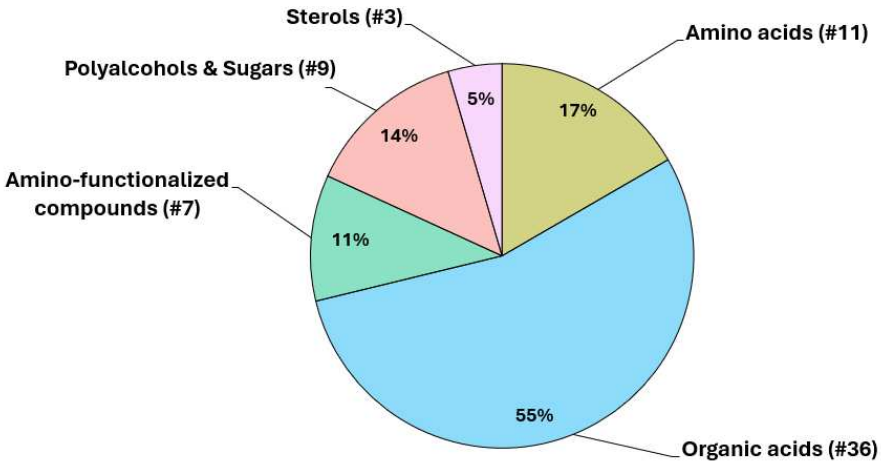


Figure 5.3. *Distribution of the 66 identified serum metabolites across five major chemical classes.*

The list of compounds highlighted in Figure 5.2 is given in Table 5.1, along with retention times (^1t_r, ^2t_r), spectral similarity (Sim.), and quantification signal (Quant. Ion).

Table 5.1. *List of the 66 putatively identified metabolites in human serum. ^{1}tr (min): Retention time in the first dimension; ^{2}tr (s): Retention time in the second dimension; Sim.: Spectral similarity; Quant. Ion: Quantification ion.*

#	Class	Compounds	^{1}tr (min)	^{2}tr (s)	Sim.	Quant. Ion
1	OA	Lactic Acid, 2TMS derivative	10.63	1.78	933	73.06
2	AA	L-Valine, TMS derivative	11.33	2.30	802	73.06
3	OA	2-Hydroxybutyric acid, 2TMS derivative	12.96	1.84	899	73.06
4	OA	Oxalic acid, 2TMS derivative	13.13	2.41	842	156.05
5	AA	L-Leucine, TMS derivative	13.72	2.23	835	73.07
6	OA	3-Hydroxybutyric acid, 2TMS derivative	14.07	1.81	812	73.07
7	OA	2-Hydroxy-3-methylbutyric acid, 2TMS derivative	14.24	1.78	834	73.07
8	AA	L-Isoleucine, TMS derivative	14.42	2.30	787	73.06
9	AA	L-Proline, TMS derivative	14.48	2.66	859	73.06
10	AM	Urea, 2TMS derivative	16.69	2.72	880	73.06
11	PA/SU	Diethylene glycol, 2TMS derivative	16.93	2.02	825	73.06
12	AA	L-Serine, 2TMS derivative	17.34	2.24	844	73.07
13	OA	Benzeneacetic acid, TMS derivative	18.39	3.11	773	75.04
14	AA	Glycine, 3TMS derivative	18.97	1.84	902	73.06
15	OA	Glyceric acid, 3TMS derivative	20.08	1.79	871	73.06
16	OA	Pentanedioic acid, 2TMS derivative	22.12	2.30	832	75.04
17	AM	2-Aminomalonic acid, N,O,O,-TMS derivative	24.40	2.17	830	73.07
18	OA	Malic acid, 3TMS derivative	25.10	1.98	779	75.04
19	OA	Salicylic acid, 2TMS derivative	25.33	2.69	936	75.04
20	AA	L-5-Oxoproline, 2TMS derivative	25.68	3.03	896	75.04
21	PA/SU	D-Threitol, 4TMS derivative	25.91	1.51	857	75.04
22	AA	L-Aspartic acid, 3TMS derivative	25.97	1.97	897	75.04
23	AA	L-Cysteine, 3TMS derivative	26.91	2.12	799	75.04
24	AM	3-Aminopiperidine-2,6-dione, 2TMS derivative	28.01	3.30	850	75.04
25	OA	Ibuprofen, TMS derivative	28.42	2.84	803	73.06
26	AA	Phenylalanine, 2TMS derivative	28.77	2.49	917	73.06
27	OA	L-Glutamic acid, 3TMS derivative	28.83	2.03	832	75.04
28	AA	Asparagine, 3TMS derivative	30.23	2.30	798	73.06

29	PA/SU	Adonitol, 5TMS derivative	31.69	1.51	823	73.06
30	PA/SU	Xylitol, 5TMS derivative	32.04	1.48	828	73.06
31	OA	(E)-Aconitic acid, 3TMS derivative	32.33	2.25	824	75.04
32	AM	4-(2-Piperidinylcarbonyl)morpholine, TMS	34.08	3.29	769	75.04
33	OA	Citric acid, 4TMS derivative	34.43	1.96	820	75.04
34	PA/SU	1,5-Anhydroglucitol, 4TMS derivative	35.13	1.87	815	73.06
35	OA	2-Hydroxyibuprofen, O,O-bis TMS derivate	35.48	2.48	783	73.06
36	OA	Sebacic acid, 2TMS derivative	35.72	2.42	788	73.06
37	AM	Isopropyl phenidate, TMS derivate	35.89	3.00	812	156.05
38	OA	4-Hydroxyphenyllactic acid, 3TMS derivative	36.12	2.28	767	73.07
39	PA/SU	β-D-Glucopyranose, 5TMS derivative	36.36	1.68	911	73.07
40	PA/SU	D-Glucitol, 6TMS derivate	37.41	1.45	833	73.07
41	OA	3-Indoleacetic acid, 2TMS derivative	37.41	3.71	868	73.06
42	AM	2-(1-Adamantyl)piperidine, TMS derivate	37.47	2.98	750	73.06
43	OA	Mannonic acid, γ-lactone, 4TMS derivate	37.88	1.92	767	73.06
44	PA/SU	2-amino-2-deoxy-1,3,4,5,6-pentakis-O-(trimethylsilyl)-D-Glucitol	38.58	1.91	751	73.06
45	PA/SU	α-D-Mannopyranose, 5TMS derivative	38.69	1.60	918	73.07
46	OA	Palmitelaidic acid, TMS derivate	38.69	2.37	839	75.04
47	AM	Dopamine, 4TMS derivative	39.74	2.66	764	73.06
48	OA	Heptadecanoic acid, TMS derivative	41.43	2.25	839	73.06
49	OA	Stearidonic acid, TMS derivate	42.43	2.62	751	75.04
50	OA	Indole-3-lactic acid, 3TMS derivative	42.43	2.89	818	73.07
51	OA	(9Z,12Z)-Octadecadienoic acid, TMS derivative	42.83	2.53	919	75.04
52	OA	(Z)-Oleic Acid, TMS derivate	42.95	2.40	921	75.04
53	OA	α-Linolenic acid, TMS derivative	42.95	2.64	848	75.04
54	OA	Arachidonic acid, TMS derivative	46.10	2.75	945	75.04
55	OA	Eicosapentaenoic Acid, TMS derivative	46.22	2.89	860	75.04
56	OA	(8Z,11Z,14Z)-Icosatrienoate, O-TMS derivate	46.51	2.66	798	75.04
57	OA	11,14-Eicosadienoic acid, TMS derivative	46.92	2.55	810	75.04
58	OA	Arachidic acid, TMS derivative	47.56	2.30	790	73.06
59	OA	Doconexent, TMS derivative	49.78	3.00	896	73.06
60	OA	(7Z,10Z,13Z,16Z,19Z)-Docosapentaenoic acid, TMS derivative	50.07	2.91	846	75.04

61	OA	1,3-Dihydroxypropan-2-yl octadec-9-enoate, 2TMS derivate	52.99	2.97	753	73.06
62	OA	1-Monooleoylglycerol, 2TMS derivative	53.63	2.29	807	73.06
63	OA	1-Palmitoyllysophosphatidic acid, 3O-TMS derivate	58.59	2.41	767	73.06
64	ST	Cholesterol, TMS derivative	59.81	3.10	903	75.04
65	ST	(3β,5α)-Cholest-7-en-3-ol, TMS derivative	60.69	3.22	800	75.04
66	ST	Campesterol, TMS derivative	61.45	3.14	764	75.04

In order to explore the polarity of the selected metabolites, predicted and experimental logP values for each metabolite class were considered, and Figure 5.4 shows these distributions through a series of boxplots. Polyalcohols and sugars (PA/SU) display strongly negative logP values, reflecting their highly polar and hydrophilic nature. A similar trend is observed for amino acids (AA), which also cluster in the negative logP region. Amino-functionalized compounds (AM) predominantly show polar behavior as well, with one notable outlier extending towards higher logP values. Organic acids (OA) represent the broadest and most chemically diverse class, ranging from highly polar metabolites (log P ≈ -3.7) to moderately lipophilic species (logP ≈ 8.5), consistent with their structural heterogeneity. Finally, sterols (ST), as expected, exhibit the highest logP values, closely clustered around 8.7, confirming their hydrophobic and non-polar characteristics.

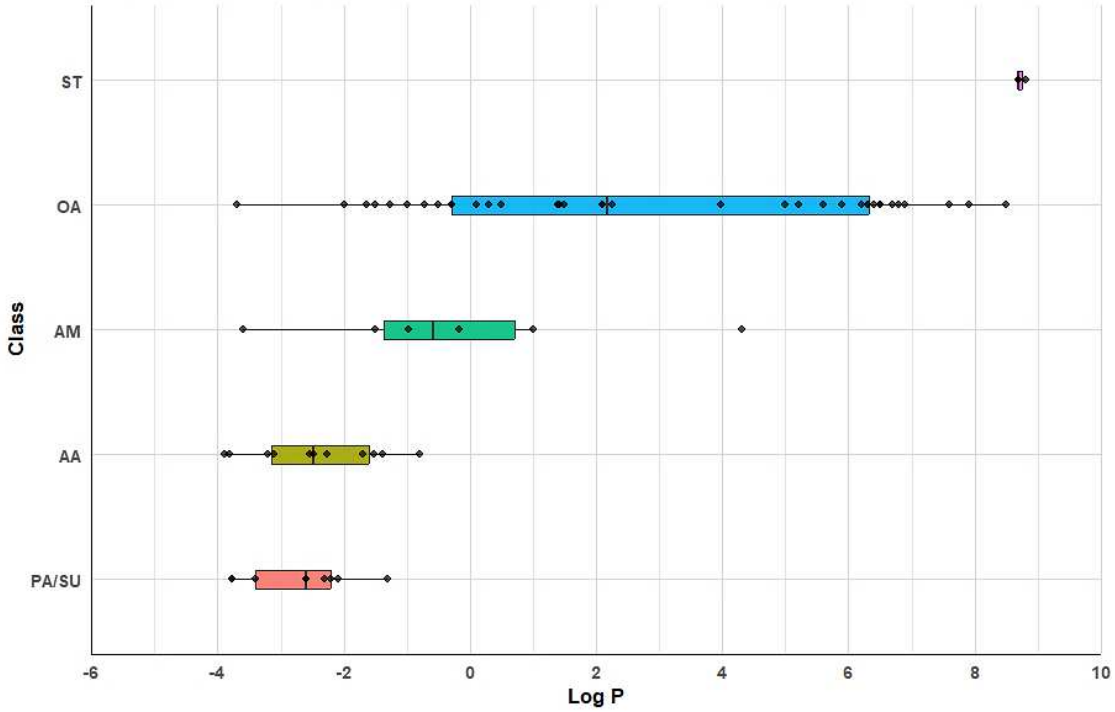


Figure 5.4. *Predicted (X-LOG-P3, Pubchem; ACD/LogP, Chemspider) and experimental (Pubchem and Chemspider) logP distributions for each metabolite class [43, 44].*

5.3.2. Evaluation of extraction performance

An exploratory hierarchical cluster analysis (Figure 5.5) was conducted to highlight differences among metabolomic profiles obtained from various solvents. The resulting dendrogram shows samples cluster according to the extraction solvent, demonstrating strong intra-group reproducibility. Among the evaluated formulations, the MeOH + ACN + acetone mixture (solvent C) generates a distinct metabolic signature, separating more clearly from the profiles obtained with solvent B (IPA:ACN:H_2O, 3:3:2) and solvent A (MeOH). Across all solvents, broad high-intensity regions (red) and low-intensity zones (blue) are visible in the heatmap, reflecting the intrinsic complexity of the serum metabolome and the limited degree of separation achievable using this global visualization alone. Nonetheless, the clustering patterns indicate that each solvent exhibits its own extraction specificity, while also sharing partial overlaps with other extraction profiles. Essentially, these results show that the choice of extraction solvent influences the fraction of chemical information retained from the original sample. In this sense, each solvent system defines a different level of analytical information continuity (AIC), selectively enhancing or suppressing specific regions of the serum chemical space.

Within the refined dataset of 66 reproducible metabolites, evaluation of extraction stability using the mean and median CV (coefficient of variation) values indicates that solvent C delivers the highest reproducibility, as shown in the ring charts in Figure 5.1. Overall, the reproducibility of all three extraction solvents can be considered satisfactory. Mean and median of the CV values of the selected metabolites are given in Table 5.1.

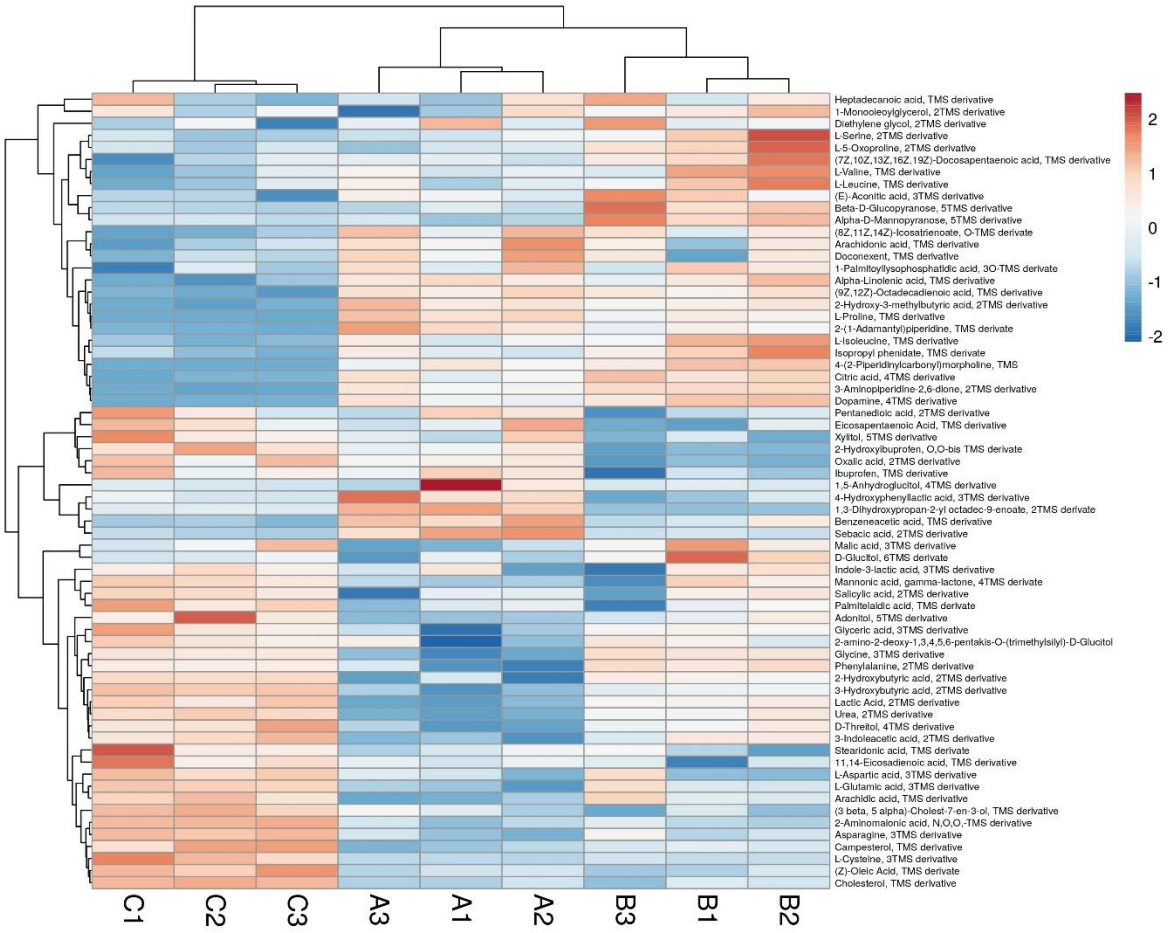


Figure 5.5. *Heatmap and hierarchical cluster analysis of the 66 metabolites putatively identified among solvent A (MeOH), B (C_3H_8O:CH_3CN:H_2O, 3:3:2) and C (CH_3OH:CH_3CN:$(CH_3)_2CO$, 1:1:1).*

Table 5.1. *Mean and median values of the coefficient of variation among solvent A (MeOH), B (C_3H_8O:CH_3CN:H_2O, 3:3:2) and C (CH_3OH:CH_3CN:$(CH_3)_2CO$, 1:1:1).*

Solvent	CV (mean)	CV (median)
A	8.46	7.21
B	8.66	7.90
C	6.26	4.26

At this point, in order to identify which metabolites contributed to the differentiation in the extraction performance of the selected solvents, a one-way analysis of variance (ANOVA) was performed. While this statistic manipulation indicates whether at least one solvent differs significantly from the others on the selected feature, it does not reveal which specific solvent pairs are responsible for these differences. Therefore, for metabolites with significant ANOVA results, a post hoc Tukey HSD test was applied to essentially determine the directionality of those changes. This comparative strategy enabled the generation of 3 volcano plots, one for each pair of solvents (A-C, A-B, and B-C). Among these, the B-C comparison is shown in Figure 5.6, since according to the author of this PhD thesis, this pair

emerged as the most informative. This observation is further supported by the heatmap, where these two solvents display the highest proportion of red values.

In Figure 5.6, although certain chemical classes (*i.e.,* organic acids) are distributed across the entire plot without exhibiting any directional trend, other metabolite groups present more defined patterns. In particular, compounds located on the left side of the plot are characterized by negative log_2FoldChange values and are more extracted by solvent B (C_3H_8O:CH_3CN:H_2O, 3:3:2), including polar and sugar-related metabolites such as PA/SU, AA, and AM. In contrast, metabolites positioned on the right side showing positive log_2FoldChange values are more extracted with solvent C (CH_3OH:CH_3CN:$(CH_3)_2CO$, 1:1:1). Here, sterols (ST), in accordance with the evaluation of Figure 5.4, highlight the affinity of solvent C for less polar and lipophilic species.

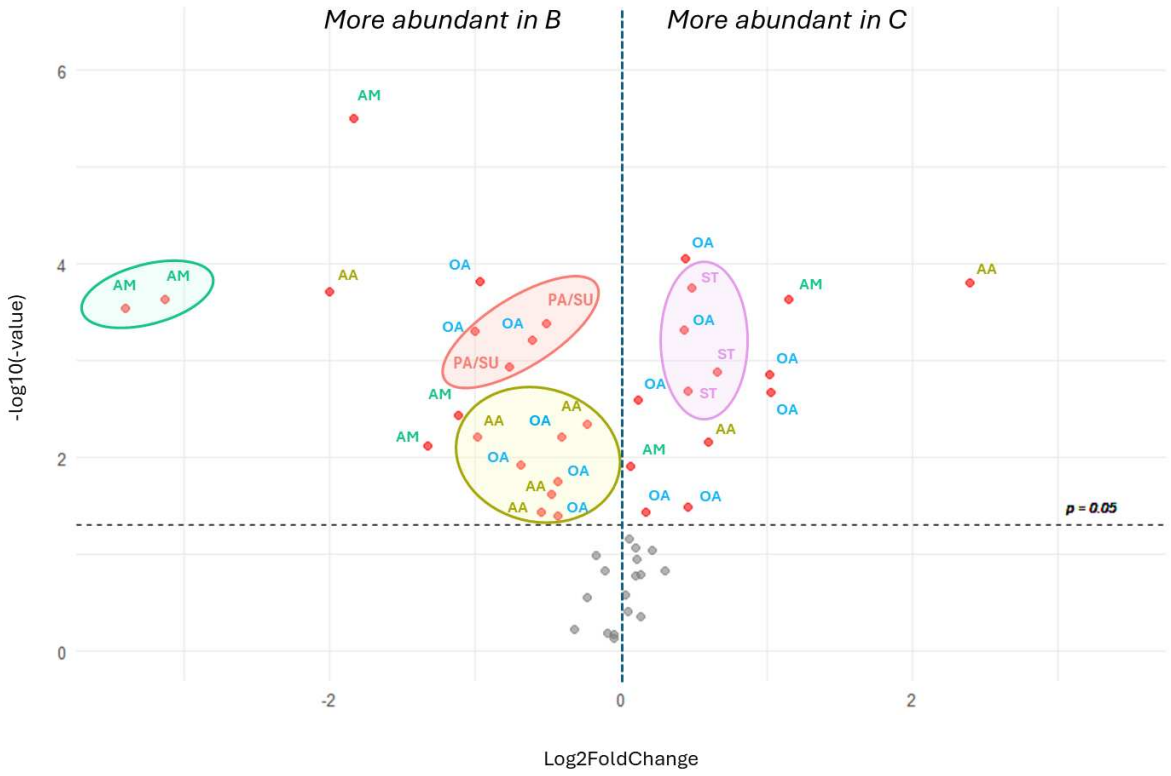


Figure 5.6. *Volcano plot comparison of solvent B (C_3H_8O:CH_3CN:H_2O, 3:3:2) and solvent C (CH_3OH:CH_3CN:$(CH_3)_2CO$, 1:1:1), highlighting statistically significant differences for metabolites with p-values below 0.05. Overall, 50 out of the 66 metabolites identified contribute to the differentiation of at least one pair of solvents. Here, 33 are highlighted with a p-value below 0.05, while 17 do not meet this criterion in this particular differentiation.*

At this stage, to further examine this aspect, we aimed to investigate the specific contribution of each extraction solvent for every identified metabolite class.

To account for the large variability in the absolute signal intensities among many metabolites, a normalization approach was applied. For each metabolite, peak areas across the three extraction solvents were normalized according to the maximum value. This was set to 100, and all remaining values were expressed as relative percentages. This procedure allows comparison of extraction efficiencies across solvents for each metabolite, independently of their absolute abundance.

Subsequently, data were grouped by chemical class. For each replicate, normalized metabolite values belonging to the same class were summed. These summed values were

then averaged across the three replicates for each solvent, yielding a single class value per solvent. Finally, average values were divided by the number of metabolites in the corresponding class and expressed on a 0-100 scale, generating a relative class extraction index for each solvent. This normalization strategy was intentionally chosen to emphasize relative extraction performance across solvents, rather than absolute abundance, which spans several orders of magnitude among metabolites and would otherwise dominate class comparisons. The results of this series of normalizations are presented in Figure 5.7.

To statistically assess differences among the three extraction solvents, a one-way ANOVA followed by Tukey HSD post hoc test was applied, with significance levels reported as $p < 0.05$ (*), $p < 0.01$ (**), and $p < 0.001$ (***). For metabolite classes, it is characterized by negative logP values, namely, amino acids (AA), amino-functionalized compounds (AM), and polyalcohols & sugars (PA&SU), solvent B, as previously highlighted consistently displayed the highest extraction yields, often with several thresholds of significant differences, reflecting its strong affinity for polar metabolites due to its water content composition. For organic acids, solvent A (methanol) yielded the highest relative extraction efficiency, although differences among solvents were less pronounced and, in some cases, did not reach statistical significance. Finally, in the case of sterols (ST), solvent C exhibited superior extraction performance, supported by highly significant p-values, confirming its enhanced ability to extract lipophilic metabolites.

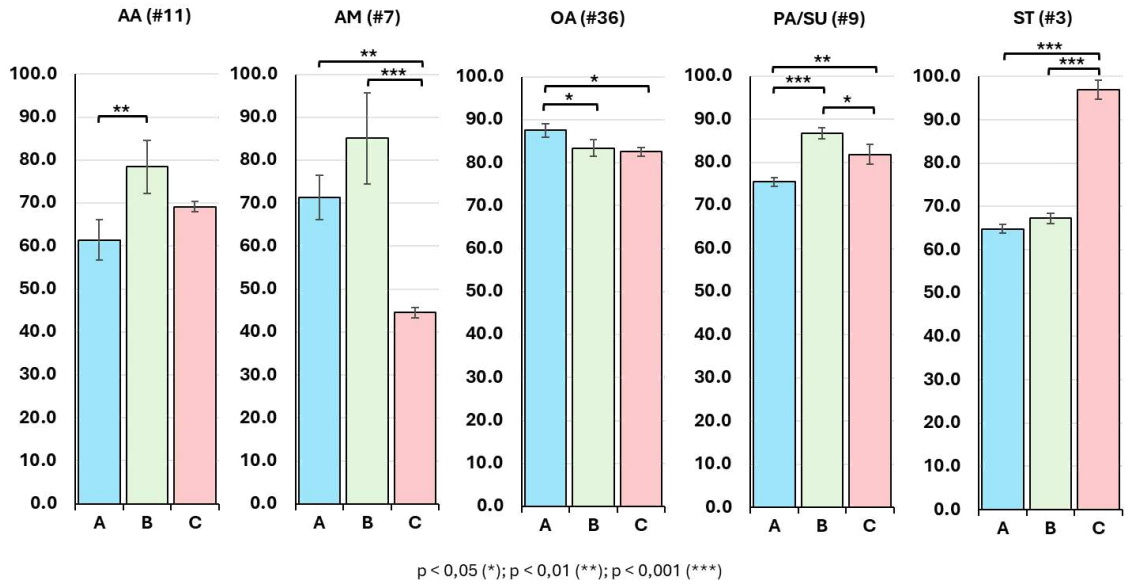


Figure 5.6. *Relative extraction yield index for each metabolite class across the three extraction solvents. A) Methanol, blue bars; B) C_3H_8O:CH_3CN:H_2O, 3:3:2, green bars; C) CH_3OH:CH_3CN:$(CH_3)_2CO$, 1:1:1, red bars.*

The considerations discussed so far highlight relative extraction capabilities and do not account for the overall contribution of each metabolite class to the serum metabolome. To address this aspect, a complementary approach based on absolute signal intensities was also considered. For each metabolite, peak areas were integrated with the total ion current (TIC) signal and averaged across the three replicates for each solvent. These average values were then summed within each metabolite class and finally expressed as percentages of the total signal intensity of the 66 selected metabolites.

The percentage contribution of each class to the total integrated area is shown in Figure 5.7. Overall, organic acids (OAs) emerged as the predominant contributors, accounting for approximately 40% of the total signal intensity. Sterols (ST) represented the second major class despite comprising only three individual metabolites. This result can be due to the

inherently high cholesterol content characteristic of human serum. Polyalcohols & sugars (PA/SU) and amino compounds (AM) also contributed appreciably to the overall chromatographic response, whereas amino acids (AA) constituted only a minor fraction (≈5%) in the dataset. Finally, solvent B showed a distinctive profile, with the PA/SU class contributing more than sterols, unlike the trends observed for solvents A and C.

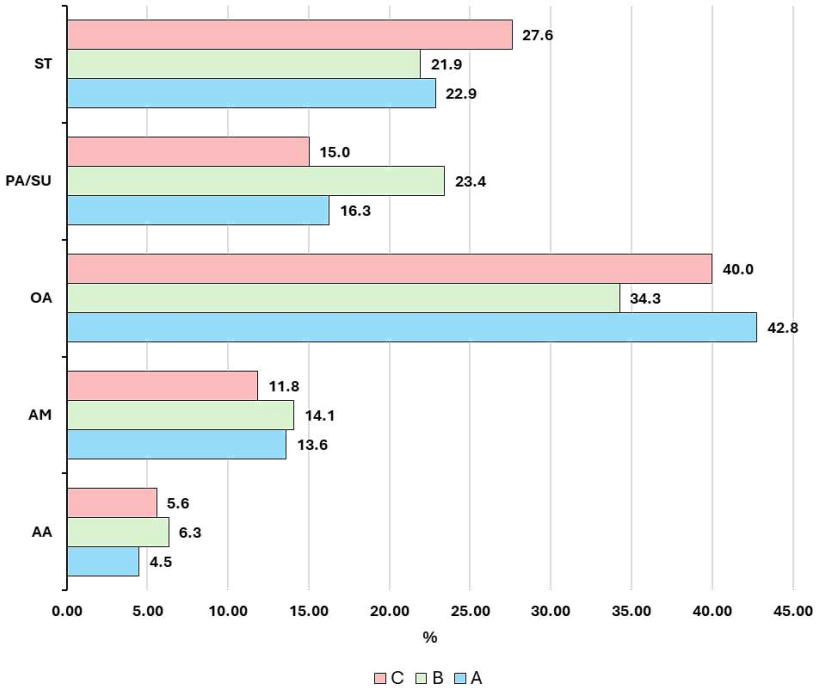


Figure 5.7. *Percentage contribution of each metabolite class to the total integrated chromatographic area (TIC) for the three extraction solvents: A) Methanol, blue bars; C) IPA:ACN:H_2O 3:3:2, green bars; D) MeOH:ACN:$(CH_3)_2CO$ 1:1:1, red bars.*

To provide an overall assessment of the extraction efficiency for each solvent system, a classification region was defined on the two-dimensional chromatogram and integrated at *m/z* 73, corresponding to the characteristic trimethylsilyl fragment $[Si(CH_3)_3]^+$. This region was delineated to exclude the column bleed zone and to ensure only sample signals rather than stationary-phase artifacts.

Extraction performance was evaluated using two complementary integration strategies: the signal- and peak-domain approaches (Figure 5.8). In the former, the total signal within the predefined chromatographic region is integrated directly from the ion trace, providing a global measure of extracted chemical content. As such, this approach is less sensitive to variability arising from chromatographic performance or data processing parameters. However, the signal domain does not differentiate chemically informative signals from non specific ones, such as an overexpression of the background noise. In the peak-domain, instead, only discrete chromatographic peaks detected within the same region are consider. Consequently, this approach is more sensitive to peak detection and integration consistency, and unresolved signals may not be captured as discrete features.

From the analytical information continuity perspective discussed in Section 2.5, these observations indicate that similar total signals do not necessarily correspond to comparable analytical information content. While the signal domain reflects the overall extracted signal, the peak domain captures how the chemical features are preserved and resolved across replicates.

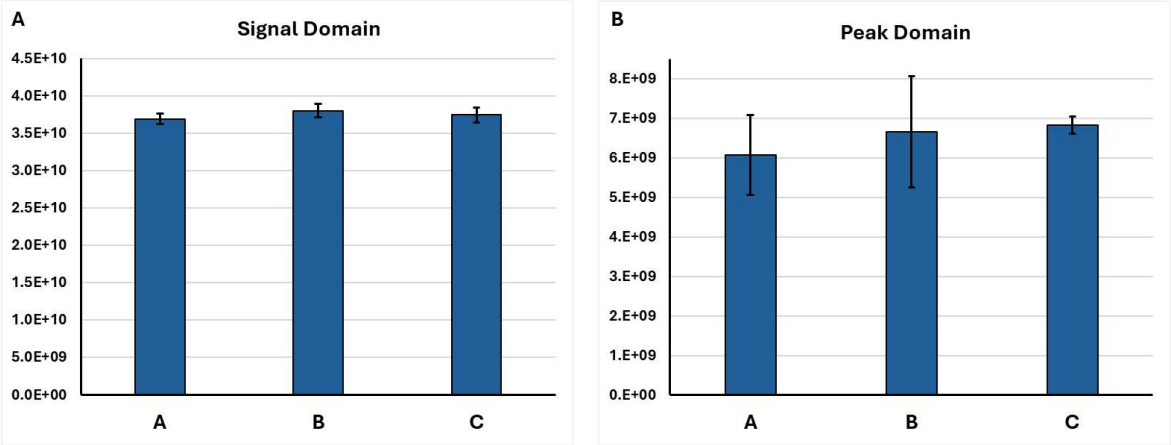


Figure 5.8. *Comparison of extraction performance across solvents A, B, and C based on an integrated classification region using the m/z 73 ion. A) Results in the signal-area domain. B) Evaluation in the peak-area domain.*

According to the classification evaluation, solvents display no pronounced differences in overall extraction yield in the signal domain, showing comparable total responses. In the peak domain, solvent C shows the least variability across replicates, suggesting improved reproducibility and more consistent extraction behavior of the metabolites. In addition, results suggest that the mixed solvent systems (B and C) tend to provide higher responses than methanol alone (A), indicating that mixtures enhance extraction efficiency relative to a single-solvent approach.

5.4. Conclusions

In this study, the proposed non-targeted workflow based on GC×GC-MS demonstrated the ability to produce structured chromatograms capable of revealing solvent-dependent trends and class-specific patterns within the human serum metabolome. The comprehensive nature of non-targeted analysis allowed broad metabolite coverage across a wide LogP range, highlighting the inherent chemical diversity of human serum. The findings showed that even minor modifications in the composition of the extraction solvent, particularly the relative proportions of water, acetonitrile, and acetone, can significantly alter the extracted metabolic profile. As a result, metabolite profiles are strongly influenced by the solvent, underscoring the importance of carefully selecting extraction conditions in non-targeted serum metabolomics.

Instead of seeking a universal optimal solvent, this research indicates that specific solvent combinations are more appropriate for different analytical goals. For instance, studies targeting highly polar metabolites, such as amino acids or saccharides, will benefit from water-rich systems, whereas the extraction of apolar analytes is better suited to mixtures containing acetonitrile and acetone. Moreover, the overall metabolic landscape of serum is shaped not only by chemical diversity but also by differences in signal intensity among metabolite classes. This fact is exemplified by sterols, which contribute disproportionately to the total signal despite being few. This complexity reflects the heterogeneous nature of serum, composed of metabolites with wide-ranging polarity, molecular weight, and relative abundance.

Overall, the results confirm that non-targeted analysis is also a powerful approach for comparative extraction studies, enabling comprehensive evaluation of solvent-dependent profiles, beyond few analytes in a conventional targeted analysis. Finally, this research demonstrated that the extraction solvent plays a critical role in defining the accessible

chemical space, effectively narrowing or broadening the analytical information continuity (AIC) of the serum metabolome.

References

[1] S. Nalbantoglu, "Metabolomics: Basic Principles and Strategies," in *Molecular Medicine*, IntechOpen, 2019. doi: 10.5772/intechopen.88563.

[2] R. Bujak, W. Struck-Lewicka, M. J. Markuszewski, and R. Kaliszan, "Metabolomics for laboratory diagnostics," Sep. 10, 2015, *Elsevier B.V.* doi: 10.1016/j.jpba.2014.12.017.

[3] W. B. Dunn *et al.*, "Molecular phenotyping of a UK population: defining the human serum metabolome," *Metabolomics*, vol. 11, no. 1, pp. 9–26, Feb. 2015, doi: 10.1007/s11306-014-0707-1.

[4] N. Bar *et al.*, "A reference map of potential determinants for the human serum metabolome," *Nature*, vol. 588, no. 7836, pp. 135–140, Dec. 2020, doi: 10.1038/s41586-020-2896-2.

[5] D. J. Beale, A. V. Karpe, and W. Ahmed, "Beyond metabolomics: A review of multi-omics-based approaches," in *Microbial Metabolomics: Applications in Clinical, Environmental, and Industrial Microbiology*, Springer International Publishing, 2016, pp. 289–312. doi: 10.1007/978-3-319-46326-1_10.

[6] M. Zampieri, K. Sekar, N. Zamboni, and U. Sauer, "Frontiers of high-throughput metabolomics," Feb. 01, 2017, *Elsevier Ltd*. doi: 10.1016/j.cbpa.2016.12.006.

[7] S. Qiu *et al.*, "Small molecule metabolites: discovery of biomarkers and therapeutic targets," Dec. 01, 2023, *Springer Nature*. doi: 10.1038/s41392-023-01399-3.

[8] S. Patel and S. Ahmed, "Emerging field of metabolomics: Big promise for cancer biomarker identification and drug discovery," Mar. 05, 2015, *Elsevier*. doi: 10.1016/j.jpba.2014.12.020.

[9] S. Ren *et al.*, "Integration of metabolomics and transcriptomics reveals major metabolic pathways and potential biomarker involved in prostate cancer," *Molecular and Cellular Proteomics*, vol. 15, no. 1, pp. 154–163, Jan. 2016, doi: 10.1074/mcp.M115.052381.

[10] G. Lubes and M. Goodarzi, "GC–MS based metabolomics used for the identification of cancer volatile organic compounds as biomarkers," Jan. 05, 2018, *Elsevier B.V.* doi: 10.1016/j.jpba.2017.07.013.

[11] S. Salciccia *et al.*, "Biomarkers in prostate cancer diagnosis: From current knowledge to the role of metabolomics and exosomes," May 01, 2021, *MDPI*. doi: 10.3390/ijms22094367.

[12] A. Wojakowska *et al.*, "Metabolic profiles of whole serum and serum-derived exosomes are different in head and neck cancer patients treated by radiotherapy," *J Pers Med*, vol. 10, no. 4, pp. 1–14, Nov. 2020, doi: 10.3390/jpm10040229.

[13] R. Xue *et al.*, "A serum metabolomic investigation on hepatocellular carcinoma patients by chemical derivatization followed by gas chromatography/mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 22, no. 19, pp. 3061–3068, Oct. 2008, doi: 10.1002/rcm.3708.

[14] J. Wu, M. Wu, and Q. Wu, "Identification of potential metabolite markers for colon cancer and rectal cancer using serum metabolomics," *J Clin Lab Anal*, vol. 34, no. 8, Aug. 2020, doi: 10.1002/jcla.23333.

[15] D. Shi *et al.*, "The serum metabolome of COVID-19 patients is distinctive and predictive," *Metabolism*, vol. 118, May 2021, doi: 10.1016/j.metabol.2021.154739.

[16] J. Zhou *et al.*, "Exploration of the serum metabolite signature in patients with rheumatoid arthritis using gas chromatography–mass spectrometry," *J Pharm Biomed Anal*, vol. 127, pp. 60–67, Aug. 2016, doi: 10.1016/j.jpba.2016.02.004.

[17] C. Sun *et al.*, "Serum metabolomic profiling in patients with Alzheimer disease and amnestic mild cognitive impairment by GC/MS," *Biomedical Chromatography*, vol. 34, no. 9, Sep. 2020, doi: 10.1002/bmc.4875.

[18] L. Szczerbinski *et al.*, "Untargeted Metabolomics Analysis of the Serum Metabolic Signature of Childhood Obesity," *Nutrients*, vol. 14, no. 1, Jan. 2022, doi: 10.3390/nu14010214.

[19] B. Cao, H. Lu, P. Liu, Y. Zhang, and C. Wang, "Serum metabolomics signature of maternally inherited diabetes and deafness by gas chromatography–time of flight mass spectrometry," *J Diabetes Investig*, vol. 16, no. 1, pp. 146–153, Jan. 2025, doi: 10.1111/jdi.14334.

[20] D. S. Wishart *et al.*, "HMDB 4.0: The human metabolome database for 2018," *Nucleic Acids Res*, vol. 46, no. D1, pp. D608–D617, Jan. 2018, doi: 10.1093/nar/gkx1089.

[21] N. Psychogios, "The Human Serum Metabolome," *PLoS One*, vol. 2, no. 6, 2011, doi: https://doi.org/10.1371/journal.pone.0016957.

[22] O. Kiseleva, I. Kurbatov, E. Ilgisonis, and E. Poverennaya, "Defining blood plasma and serum metabolome by gc-ms," Jan. 01, 2022, *MDPI*. doi: 10.3390/metabo12010015.

[23] T. Harayama and H. Riezman, "Understanding the diversity of membrane lipid composition," May 01, 2018, *Nature Publishing Group*. doi: 10.1038/nrm.2017.138.

[24] L. Brennan and H. Gibbons, "Sex matters: a focus on the impact of biological sex on metabolomic profiles and dietary interventions," *Proceedings of the Nutrition Society*, vol. 79, no. 2, pp. 205–209, May 2020, doi: 10.1017/S002966511900106X.

[25] J. L. Ren, A. H. Zhang, L. Kong, and X. J. Wang, "Advances in mass spectrometry-based metabolomics for investigation of metabolites," *RSC Adv*, vol. 8, no. 40, pp. 22335–22350, 2018, doi: 10.1039/c8ra01574k.

[26] Q. Yang *et al.*, "Metabolomics biotechnology, applications, and future trends: A systematic review," 2019, *Royal Society of Chemistry*. doi: 10.1039/c9ra06697g.

[27] M. Vinaixa, E. L. Schymanski, S. Neumann, M. Navarro, R. M. Salek, and O. Yanes, "Mass spectral databases for LC/MS- and GC/MS-based metabolomics: State of the field and future prospects," Apr. 01, 2016, *Elsevier B.V.* doi: 10.1016/j.trac.2015.09.005.

[28] Ö. C. Zeki, C. C. Eylem, T. Reçber, S. Kır, and E. Nemutlu, "Integration of GC–MS and LC–MS for untargeted metabolomics profiling," Oct. 25, 2020, *Elsevier B.V.* doi: 10.1016/j.jpba.2020.113509.

[29] M. M. Koek, R. H. Jellema, J. van der Greef, A. C. Tas, and T. Hankemeier, "Quantitative metabolomics based on gas chromatography mass spectrometry: Status and perspectives," Sep. 2011. doi: 10.1007/s11306-010-0254-3.

[30] K. Blau and J. M. Halket, *Handbook of Derivatives for Chromatography*, Second Edition. John Wiley & Sons, 1984. doi: https://doi.org/10.1007/BF02916441.

[31] J. M. Halket, D. Waterman, A. M. Przyborowska, R. K. P. Patel, P. D. Fraser, and P. M. Bramley, "Chemical derivatization and mass spectral libraries in metabolic profiling by GC/MS and LC/MS/MS," in *Journal of Experimental Botany*, Jan. 2005, pp. 219–243. doi: 10.1093/jxb/eri069.

[32] E. A. Higgins Keppler, C. L. Jenkins, T. J. Davis, and H. D. Bean, "Advances in the application of comprehensive two-dimensional gas chromatography in metabolomics," Dec. 01, 2018, *Elsevier B.V.* doi: 10.1016/j.trac.2018.10.015.

[33] B. L. Milman and I. K. Zhurkovich, "The chemical space for non-target analysis," Dec. 01, 2017, *Elsevier B.V.* doi: 10.1016/j.trac.2017.09.013.

[34] J. A. Murray, "Qualitative and quantitative approaches in comprehensive two-dimensional gas chromatography," Oct. 26, 2012. doi: 10.1016/j.chroma.2012.05.012.

[35] D. Dudzik, C. Barbas-Bernardos, A. García, and C. Barbas, "Quality assurance procedures for mass spectrometry untargeted metabolomics. a review," Jan. 05, 2018, *Elsevier B.V.* doi: 10.1016/j.jpba.2017.07.044.

[36] A. Mastrangelo, A. Ferrarini, F. Rey-Stolle, A. García, and C. Barbas, "From sample treatment to biomarker discovery: A tutorial for untargeted metabolomics based on GC-(EI)-Q-MS," 2015, *Elsevier B.V.* doi: 10.1016/j.aca.2015.10.001.

[37] F. Rey-Stolle *et al.*, "Low and high resolution gas chromatography-mass spectrometry for untargeted metabolomics: A tutorial," Jun. 01, 2022, *Elsevier B.V.* doi: 10.1016/j.aca.2021.339043.

[38] T. Cajka and O. Fiehn, "Toward Merging Untargeted and Targeted Methods in Mass Spectrometry-Based Metabolomics and Lipidomics," Jan. 05, 2016, *American Chemical Society*. doi: 10.1021/acs.analchem.5b04491.

[39] W. B. Dunn *et al.*, "Procedures for large-scale metabolic profiling of serum and plasma using gas chromatography and liquid chromatography coupled to mass spectrometry," *Nat Protoc*, vol. 6, no. 7, pp. 1060–1083, Jun. 2011, doi: 10.1038/nprot.2011.335.

[40] O. Fiehn, "Metabolomics by gas chromatography-mass spectrometry: Combined targeted and untargeted profiling," *Curr Protoc Mol Biol*, vol. 2016, Apr. 2016, doi: 10.1002/0471142727.mb3004s114.

[41] S. J. Bruce, I. Tavazzi, V. Parisod, S. Rezzi, S. Kochhar, and P. A. Guy, "Investigation of human blood plasma sample preparation for performing metabolomics using ultrahigh performance liquid chromatography/mass spectrometry," *Anal Chem*, vol. 81, no. 9, pp. 3285–3296, May 2009, doi: 10.1021/ac8024569.

[42] Z. He, Q. Luo, Z. Liu, and L. Gong, "Extensive evaluation of sample preparation workflow for gas chromatography-mass spectrometry-based plasma metabolomics and its application in rheumatoid arthritis," *Anal Chim Acta*, vol. 1131, pp. 136–145, Sep. 2020, doi: 10.1016/j.aca.2020.06.029.

[43] "PubChem." Accessed: Dec. 06, 2025. [Online]. Available: https://pubchem.ncbi.nlm.nih.gov/

[44] “ChemSpider.” Accessed: Dec. 06, 2025. [Online]. Available: http://www.chemspider.com/

Chapter 6

Hard (EI) and soft (PCI, NCI) ionization techniques as a complementary approach to expand PFAS chemical space by gas chromatography coupled to high-resolution mass spectrometry

6.1. Introduction

Per- and polyfluoroalkyl substances (PFAS) represent a class of anthropogenic compounds characterized by a fully or partially fluorinated carbon atom structure. Formal definition described PFAS as -C_nF_{2n+1}- substances, while a more recent -C_nF_{2n}-based definition has broadened this concept, substantially expanding the chemical space covered by this group of compounds [1, 2]. Today, PFAS comprises thousands of linear and branched compounds with a wide variety of terminal functional groups, resulting in more than 4000 PFAS identified for use on the global market [3, 4].

PFAS have been widely used due to their unique physicochemical properties. Combining amphiphilic behavior with thermal and chemical stability [5, 6], which has led PFAS to be commonly referred to as “forever chemicals” [7], has driven their adoption across numerous industrial and consumer applications. Typical uses include surfactants, metal coatings, electronics manufacturing, adhesives, personal care products and aqueous film-forming foams (AFFF) [8–10]. Other typical applications include water-repellent agents for textiles, carpets, and clothing, where fluorotelomer methacrylates (FTMACs) are key intermediates in the telomerization of fluorotelomer-based polymers, whereas iodine-containing PFAS, due to their labile C-I bond, play a central role as reactive precursors of fluorotelomer alcohols (FTOH) and subsequently FTMAC [11–13]. Such varied uses have led to the well-known pervasive environmental occurrence, caused merely by human activities.

In this sense, indirect sources represent a significant pathway for PFAS exposure. For example, landfills, lakes, and surface waters act as sinks for PFAS originating from waste of the industrial activities mentioned above, with concentrations strongly influenced by emission point and seasonal variability [14]. Moreover, atmospheric deposition through rain, snow, and snowmelt has emerged as a significant contributor to volatile and semi-volatile PFAS, following industrial emissions or volatilization from contaminated sites [15]. Finally, product use, disposal, and manufacturing processes can result in residual amounts of fluorotelomer-based polymers and related precursors [16].

Exposure to per- and polyfluoroalkyl substances has been increasingly associated with a wide range of adverse health outcomes, particularly during sensitive life stages (childhood and pregnancy), but not exclusively [17]. Due to their persistence, PFAS can exhibit diverse effects depending on the length of exposure and individual factors such as age, sex, or genetic predisposition [18]. Emerging evidence suggests associations between chronic PFAS contact and several types of cancer, including kidney, prostate, thyroid, and bladder cancer [19]. Moreover, metabolic disturbances have been widely reported, including insulin dysregulation, dyslipidemia (cholesterol imbalance), and hepatic steatosis [20]. Prenatal and infancy represent a critical window, as PFAS have been associated with gestational diabetes mellitus, childhood obesity, and reproductive disorders later in life [21]. Given these adverse effects and their ubiquitous presence in the environment and in everyday items, further investigation into the separation, detection, and remediation of PFAS is an urgent priority.

While PFAS research has been historically dominated by liquid chromatography (LC), gas chromatography (GC) covers a complementary chemical space that is largely inaccessible to LC, particularly for volatile and semi-volatile PFAS [22, 23]. Predictive models suggest that less than 10% of known PFAS are well suited for conventional LC, underscoring the potential of gas chromatography to expand PFAS chemical coverage [24]. Furthermore, high-resolution mass spectrometry (HRMS) enables isobaric separation and structural elucidation, while advanced configurations such as comprehensive two-dimensional gas chromatography (GC×GC) provide additional gains in selectivity and sensitivity [25, 26].

The use of multiple ionization strategies in GC-(HR)MS represents a powerful approach to expand the detectable chemical space (Figure 6.1) [27].

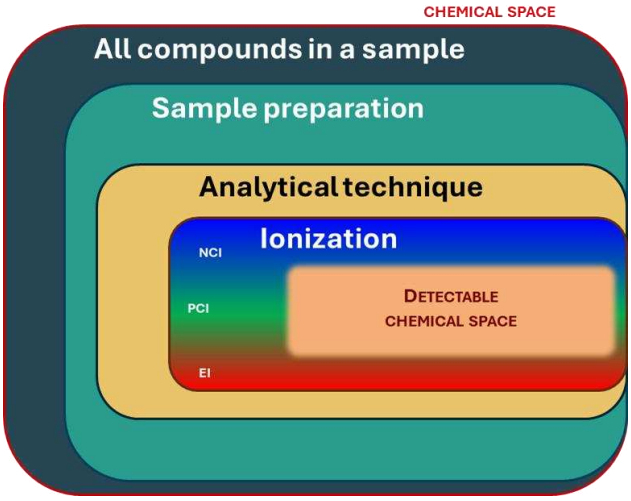


Figure 6.1. *Influence of sample preparation, analytical technique, and ionization strategy on detectable chemical space. Adapted from [28].*

Electron ionization (EI) at 70 eV remains the most widely used ionization technique due to its high reproducibility and the ability to generate highly fragmented spectra that are directly comparable across many instruments [29, 30]. However, extensive fragmentation often results in low-abundance or missing molecular ions, limiting the determination of molecular mass and related elemental compositions, especially for unknown contaminants not yet present in existing libraries [31]. In this context, soft ionization techniques such as positive chemical ionization (PCI) and negative chemical ionization (NCI) provide complementary information by preserving molecular or pseudomolecular ions and reducing spectral complexity [32, 33]. Additionally, combining hard and soft ionization extends the range of compounds that can be efficiently ionized and detected [34]. All these properties improve

selectivity for specific analyte classes and support quantitative analysis by reducing interference or background noise associated with a single ionization mode [35].

In light of these considerations, the aim of this research was to demonstrate how the complementary use of hard (EI) and soft (PCI and NCI) ionization strategies significantly improves the continuity of analytical information (AIC) and expands the detectable chemical in PFAS analysis. Fourteen volatile standards, including methacrylate, iodine-based, and fluorinated aromatic standards, were analyzed using gas chromatography coupled to high-resolution mass spectrometry using EI, PCI, and NCI. The most suitable ionization conditions were identified for each modality, and the different ionization strategies were systematically evaluated with respect to fragmentation behavior and signal intensity. Finally, analyses were also performed using comprehensive two-dimensional gas chromatography (GC×GC), highlighting the additional detectability that this technique can provide.

6.2. Materials and methods

6.2.1. Chemicals, standards, and solvents

Analytical standards of PFAS were obtained from LGC (Teddington, Middlesex, UK). Chemical classes under investigation included Fluorotelomer Methacrylates (FTMAcs), Fluorotelomer Iodine (FTIs), Fluorinated Di-Iodoalkanes (FDIAs), Fluorinated Iodine Alkanes (FIAs), and Fluoroaromatic compounds (FAc). Analytes under study included 1H,1H,2H,2H-Nonafluorohexyl Methacrylate (4:2 FTMAc), 1H,1H,2H,2H-Tridecafluorooctyl Methacrylate (6:2 FTMAc), 1H,1H,2H,2H-Heptadecafluorodecyl Methacrylate (8:2 FTMAc), 1H,1H,2H,2H-Heneicosafluorododecyl Methacrylate (10:2 FTMAc), 1H,1H,2H,2H-Perfluorooctyl Iodide (6:2 FTI), 1H,1H,2H,2H-perfluorodecyl iodide (8:2 FTI), 1H,1H,2H,2H-perfluorododecyl iodide (10:2 FTI), 1,2-Diiodoperfluorobutane (PFBuDiI), 1,2-Diiodoperfluorohexane (PFHxDiI), 1,2-Diiodoperfluorooctane (PFODiI), Perfluorodecyl Iodide (PFDeI), Perfluorododecyl Iodide (PFDoI), Bromopentafluorobenzene (BrPFB) and 1,2,4-Trifluorobenzene (1,2,4-TFB). Stock standards were solubilized and diluted in MS-grade methanol purchased from Supelco (Merck Life Science, Darmstadt, Germany). Once diluted, a mixture solution comprising all the standards was prepared: 50 µg/mL for FTMAcs, FTIs, FDIAs, and FAc, and 30 µg/mL for FIAs due to reduced solubility in methanol. This solution was analyzed by GC- and GC×GC-(HR)MS across EI, PCI, and NCI.

6.2.2. Instrumental experimental conditions

GC- and GC×GC-MS methods were performed on a Pegasus GC-HRT+ 4D (LECO Corporation, Mönchengladbach, Germany) equipped with an Agilent 8890 gas chromatograph and an Agilent 7693A autosampler. The chromatographic columns were a 30 m × 0.25 mm × 0.25 µm df Rxi-5ms as the first dimension and a 1.6 m × 0.25 mm × 0.25 µm df Rxi-17SilMS as the second dimension (both from Restek Corporation, Bellefonte, PA, USA). Injections were performed using an Agilent split/splitless (SSL) injector. The carrier gas used was helium at 1.5 mL/min column flow throughout the run. The inlet temperature was set to 250 °C, and the septum purge flow was set to 3 mL/min. One microliter of the mixture of PFAS solution was injected into the chromatographic system, and the injection was performed in split mode with a split ratio of 1:20. The oven temperature program was 40 °C (held for 5 min), then ramped at 5°C/min to 165 °C and finally raised to 280 °C at 25 °C/min (held for 1 min). Temperature offsets for the secondary oven and modulator oven were set to +5 °C and +15 °C, respectively. In GC×GC, a 2 s modulation period was used (cold and hot jets were 0.4 and 0.6 s, respectively). A 50.000 FWHM high-resolution time-of-flight mass analyzer was used to perform ion separation and to acquire a mass range from 29 to 1000 *m/z* at 10 Hz and 150 Hz, respectively, in GC- and GC×GC modes. Ion source conditions involved electron energy of 70 eV for EI, while for PCI and

NCI, 140 eV was used. Soft ionization was performed using methane gas at 0.8 mL/min in PCI mode and 1.85 mL/min in NCI mode. The ion source was maintained at 250 °C for EI, while for soft ionization (PCI and NCI), the unit was kept at 165 °C. According to the guidelines provided by the instrument manufacturer, the mass spectrometer was tuned in each ionization mode before proceeding with the analyses. Additionally, a 150 s acquisition delay was used, and the transfer line was maintained at 330°C for all analyses. Data were collected and analyzed using ChromaTOF for HRT software version 5.59 (LECO Corporation).

6.3. Results and discussion

6.3.1. PFAS fragmentation patterns

The PFAS standard mixture was analyzed using GC- and GC×GC-(HR)MS under EI, PCI, and NCI modes. All analyses were conducted under the same chromatographic conditions, with the instrument calibrated separately for each ionization source, to ensure that variations in signal intensity across the three modes reflect inherent ionization properties rather than instrumental factors.

A visualization of the elution order of target compounds is given in Figure 6.2, showing the GC-HRMS chromatogram obtained in EI mode. Additionally, chromatographic properties such as retention time and experimental retention indexes are given in the subsequent Table 6.1. The chromatographic trace corresponds to the AIC, which represents the sum of all ions in the true spectrum of each peak, integrated over the peak duration. This includes only the ions assigned after deconvolution, thereby removing noise or co-eluting contributions. Each label corresponds to a different PFAS standard, starting with 1,2,4-Trifluorobenzene (1,2,4-TFB) and ending with 1H,1H,2H,2H-Heneicosafluorododecyl Methacrylate (10:2 FTMAc). Each compound exhibits a good chromatographic profile, indicating efficient separation for most compounds with the final method conditions, and stable ionization behavior across the series was observed. Moreover, to ensure the PFAS mixture remained chemically stable throughout the sequence of measurements performed under different ionization modes, the solution was reinjected under the initial ionization condition. Analytical observations verify that no degradation, adsorption, or instrumental drift occurred during the sequence.

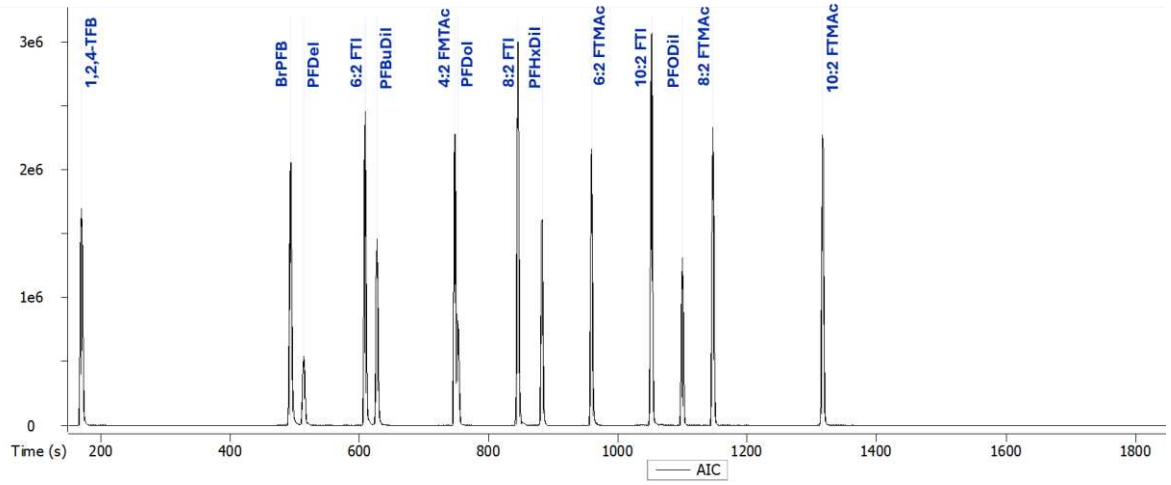


Figure 6.2. *Analytical Ion Current chromatogram of the PFAS standard mixture under EI conditions.*

Table 6.1. *Chromatographic properties of PFAS compounds ordered by chemical class (RT: Retention time; Exp. RI: Experimental Retention Index)*

Compound	Formula	Exact Mass	RT (min)	Exp. RI
1,2,4-TFB	C_6BrF_5	332.043335	2.80	685
BrPFB	$C_6H_3F_3$	432.038948	8.19	860
4:2 FTMAc	$C_{10}H_9F_9O_2$	532.032561	12.46	974
6:2 FTMAc	$C_{12}H_9F_{13}O_2$	632.026174	15.93	1077
8:2 FTMAc	$C_{14}H_9F_{17}O_2$	473.914466	19.05	1178
10:2 FTMAc	$C_{16}H_9F_{21}O_2$	573.908079	21.89	1276
6:2 FTI	$C_8H_4F_{13}I$	673.901692	10.11	911
8:2 FTI	$C_{10}H_4F_{17}I$	453.795623	14.04	1021
10:2 FTI	$C_{12}H_4F_{21}I$	553.789236	17.48	1123
PFBuDiI	$C_4F_8I_2$	653.782849	10.40	919
PFHxDiI	$C_6F_{12}I_2$	645.870392	14.65	1039
PFODiI	$C_8F_{16}I_2$	745.864005	18.27	1152
PFDeI	$C_{10}F_{21}I$	245.908789	8.52	869
PFDoI	$C_{12}F_{25}I$	132.018136	12.49	976

Electron ionization (EI)

As expected, EI generates extensive fragmentation, since the interaction between PFAS analyte and a high-energy electron (*i.e.,* 70 eV) produces a radical cation $[M]^{+\bullet}$ with substantial internal energy to exceed energetic requirements for bond dissociation, promoting rapid decomposition of $[M]^{+}\bullet$ into a variety of fragment ions [36]. In particular, this study highlights both characteristic perfluorinated ions originating from C-F bond cleavage within the perfluorinated backbone, but also compound-specific fragments related to the CF_2-based homologous series. However, it should be noted that despite the extensive fragmentation observed under EI, the molecular ion $[M]^{+\bullet}$ was still detectable for most of the PFAS chemical classes (*i.e.,* FAc, FTMAc, FTIs, FDIAs), except for the FIAs. This finding suggests that some of the remaining molecular species stay ionized without dissociating immediately. Therefore, even though EI involves high energy, retaining molecular information helps determine structure while maintaining characteristic fragmentation patterns. Moreover, mass accuracies well below 1 ppm were achieved in EI mode, outperforming those obtained under PCI and NCI conditions. Such high mass accuracy enabled the reliable assignment of molecular formulas for the observed fragmentation products.

Under EI conditions, FAcs exhibit distinct fragmentation behaviors, with no common ions observed between them, but both show characteristic fragments arising from extensive cleavage of the C_6 ring. Despite this fragmentation, the aromatic core demonstrates notable stability in the EI environment, as evidenced by the survival of benzene-derived ions in both spectra. For 1,2,4-TFB, the molecular ion is preserved and detected as $[C_6H_3F_3]^+$ at *m/z* 132, while the C_6 ring remains identifiable through the characteristic $[C_6H_2F_2]^+$ ion at *m/z* 112.

Additional fragmentation involves the opening of the aromatic ring, generating both fluorinated ions (*e.g.*, *m/z* 81, $[C_5H_2F]^+$) and non-fluorinated species (*e.g.*, *m/z* 63, $[C_5H_3]^+$). In contrast, BrPFB displays a distinctive bromine isotopic signature, with intense peaks at *m/z* 246 and 248 corresponding to $[C_6{}^{79}BrF_5]^+$ and $[C_6{}^{81}BrF_5]^+$, respectively. The presence of fragments at *m/z* 148 and 167 ($[C_6F_4]^+$ and $[C_6F_5]^+$) indicates partial preservation of the benzene backbone, while the highly abundant ions at m/z 117, 98, and 93 ($[C_5F_3]^+$, $[C_5F_2]^+$, $[C_3F_3]^+$) reflect pronounced degradation of the carbon structure.

Across the entire FTMAc series, EI produces highly comparable fragmentation spectra, characterized by a dominant base peak at *m/z* 69, assigned to the $[C_4H_5O]^+$ ion, arising from the cleavage of the C-O bond from the methacrylate moiety. This could occur as both the double bond and the carbonyl group withdraw electron density, destabilizing the adjacent carbon-oxygen bond. The resulting $[C_4H_5O]^+$ fragment could be stabilized by resonance, allowing delocalization of the positive charge through movement of electron pairs (Figure 6.3). Similar types of stabilization mechanisms by charge delocalization are given by alpha-bond cleavage sections in Gross and McLafferty-Turecek's book [32, 37].

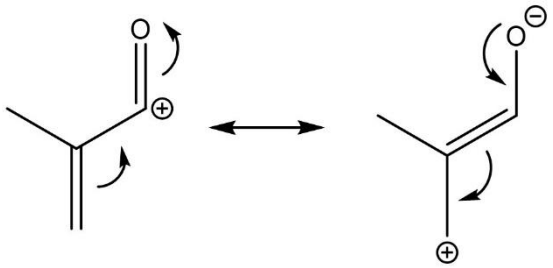


Figure 6.3. *Stabilization by resonance of the fragment* $[C_4H_5O]^+$

$[CF_3]^+$ can also be observed at *m/z* 69 for all the FTMAc and is separate from the base ion due to high-resolution acquisition. The molecular ion is consistently preserved for all congeners, appearing at *m/z* 332, 432, 532, and 632 for $[C_{10}H_9F_9O_2]^+$, $[C_{12}H_9F_{13}O_2]^+$, $[C_{14}H_9F_{17}O_2]^+$, and $[C_{16}H_9F_{21}O_2]^+$, respectively, although its relative abundance progressively decreases with increasing molecular weight. Isotopic variations corresponding to the integration of one ^{13}C atom (*e.g.*, $[C_{15}{}^{13}CH_9F_{21}O_2]^+$ of 10:2 FTMAc in Figure 6.4) are clearly resolved next to the monoisotopic molecular ion. Beyond the expected molecular series $[C_nH_9F_{2n-11}O_2]^+$ (n=10-12-14-16), whose congeners differ by $-CF_2-CF_2-$ (C_2F_4) increments, no additional homologous patterns are observed. Fragmentation pathways are largely conserved across the spectra, yielding common product ions as *m/z* 131 and 119 ($[C_3F_5]^+$ and $[C_2F_5]^+$), attributed to C-C bond cleavage at the terminal perfluorinated chain. Further consistent ions include the complete loss of the methacrylate group, producing ions at *m/z* 86 ($[C_4H_6O_2]^+$) and its higher carbon analogue at *m/z* 113 ($[C_6H_9O_2]^+$). Additional ions shared across all congeners are observed at *m/z* 77 and 95, assigned to $[C_3H_3F_2]^+$ and $[C_3H_2F_3]$+, confirming the high similarity in EI fragmentation behavior throughout the series.

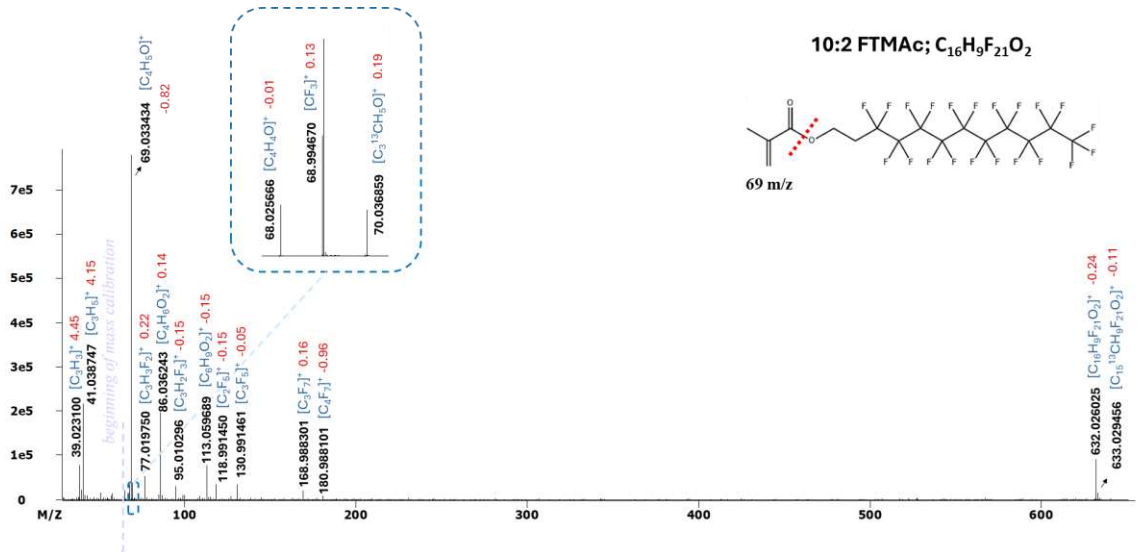


Figure 6.4. *EI-(HR)MS spectrum of the heneicosafluorododecylethyl methacrylate (10:2 FTMAc), showing proposed molecular formulas in blue and corresponding mass accuracies in red, with the dominant base ion at m/z 69 arising from C-O bond cleavage.*

For FTIs, EI produced highly reproducible fragmentation patterns across the homologous series. The molecular ion was preserved for all congeners, appearing at *m/z* 474, 574, and 674 for $[C_8H_4F_{13}I]^+$, $[C_{10}H_4F_{17}I]^+$, and $[C_{12}H_4F_{21}I]^+$, respectively, although even in this case its relative abundance consistently diminished as the molecular size increased. Base ion dominates the spectrum at *m/z* 69, arising from the loss of a $[CF_3]^+$ unit, together with a set of chain-related fluorinated ions at *m/z* 119, 131, and 169 ($[C_2F_5]^+$, $[C_3F_5]^+$, $[C_3F_7]^+$), which reflect progressive cleavage along the perfluoroalkyl chain. Iodine-containing fragments were also consistently detected, including $[CH_2I]^+$ and $[C_2H_4I]^+$ at m/z 141 and 155, as well as both the iodine radical cation $[I]^+$ at *m/z* 127 and its M+1 equivalent $[HI]^+$ at m/z 128. Homologous fragmentation patterns further emerged as a consequence of the incremental extension of the perfluoroalkyl chain, following C_2F_4 addition and generating characteristic series such as $[C_nHF_{2n-2}]^+$ (n = 7, 9, 11), $[C_nH_3F_{2n-4}]^+$ (n = 8, 10, 12), and $[C_nF_{2n-1}]^+$ (n = 6, 8, 10).

The behavior of FDIAs is dominated by highly intense iodine-related fragments, with the iodine cation at *m/z* 127 ($[I]^+$) and the $[CF_2I]^+$ ion at *m/z* 177 representing the most abundant mass peaks across all congeners. Notably, $[I]^+$ constitutes the base peak for PFODiI and PFHxDiI, whereas in PFBuDiI the base ion shifts to *m/z* 327, corresponding to the fluorinated iodine fragment $[C_4F_8I]^+$. Even in this chemical class, the molecular ion is preserved for each homologue and is observed at *m/z* 454, 554, and 654 ($[C_4F_8I_2]^+$, $[C_6F_{12}I_2]^+$, $[C_8F_{16}I_2]^+$), consistent with previous observations of decreasing intensity as molecular weight increases. Several fragments are consistently shared among all congeners, originating from both perfluoroalkyl chain scission at *m/z* 69, 93, 100, 119, 131, and 181 ($[CF_3]^+$, $[C_3F_3]^+$, $[C_2F_4]^+$, $[C_3F_5]^+$, $[C_4F_7]^+$), and from iodine related species characteristic of this compound class, such *m/z* 208, 227, and 239 ($[C_2F_3I]^+$, $[C_2F_4I]^+$, $[C_3F_4I]^+$). Moreover, homologous fragment series produced by successive C_2F_4 additions are also evident, corresponding to general formulas $[C_nF_{2n-1}]^+$ and $[C_nF_{2n}I]^+$ (n = 4, 6, 8). Finally, the iodine molecular ion $[I_2]^+$ at *m/z* 254 is consistently detected for all the compounds.

Finally, the electron-ionization spectra of FIAs reveal highly consistent fragmentation patterns across both congeners. In all cases, the base peak corresponds to the characteristic loss of a $[CF_3]^+$ fragment at *m/z* 69 while the molecular ion is absent, indicating complete fragmentation of $[M]^{+\bullet}$ under EI conditions. As expected for perfluorinated chains, both

spectra display the typical series of CF fragments at *m/z* 100, 119, 131, 169, 181, 219, and 269, corresponding respectively to $[C_2F_4]^+$, $[C_2F_5]^+$, $[C_3F_5]^+$, $[C_3F_7]^+$, $[C_4F_7]^+$, $[C_4F_9]^+$, and $[C_5F_{11}]^+$. Diagnostic iodine-containing fragments are also consistently observed, particularly $[I]^+$ at *m/z* 126 and $[CF_2I]^+$ at *m/z* 176, which confirm the presence and cleavage behavior of the terminal iodinated moiety. Notably, both congeners show intense high-mass fragments of the type $[C_nF_{2n+1}]^+$ (n = 10–12), arising from loss of the iodine atom and retention of the perfluorinated backbone, highlighting a dominant fragmentation pathway involving C-I bond cleavage followed by stabilization of the fully fluorinated carbon chain.

Positive chemical ionization (PCI)

In PCI, markedly lower fragmentation across all PFAS classes enables a more straightforward interpretation of the resulting mass spectra. The molecular ion was consistently observed, appearing not only as the protonated species (M+1) but also as characteristic adduct ions such as M+29 ($[M+C_2H_5]^+$) and M+41 ($[M+C_3H_5]^+$), reflecting the formation of collision-stabilized complexes in the ion source. Compared with EI and NCI, PCI generally showed lower ionization efficiencies, resulting in reduced signal intensity; however, at the expense of improved spectral clarity.

Under PCI conditions, both FAc exhibit limited fragmentation and spectra dominated by protonated molecular ions and characteristic adduct species. For 1,2,4-TFB, the base region is defined by the molecular cluster at *m/z* 132 ($[C_6H_3F_3]^+$) and its protonated analogue at *m/z* 133 ($[C_6H_3F_3+H]^+$), with higher fragments arising from association reactions, typical of methane in PCI mode, such as the ethyl and allyl cation adducts at *m/z* 161 and 173. In BrPFB, the PCI spectrum is dominated by the protonated molecular ion, in particular exhibiting diagnostic $^{79}Br/^{81}Br$ isotopic pattern at *m/z* 247-249, including the intense $[(C_6{}^{79}BrF_5)+H]^+$ and $[(C_6{}^{81}BrF_5)+H]^+$ ions. Additional protonated fragments such as $[C_6F_5+H]^+$ (*m/z* 167) further highlight the predominance of protonation pathways over fragmentation, in contrast to EI behavior. Finally, higher mass species at *m/z* 275-277 correspond to bromine ($[M+C_2H_5]^+$) adducts, confirming the formation of collision-stabilized complexes.

In the case of the FTMAc class, compounds exhibit spectra dominated by the intense $[M+H]^+$ molecular ions at *m/z* 333, 433, 533, and 633, indicating strong preservation of the perfluorinated backbone and limited chain scission. Even so, a characteristic feature in all spectra is the persistent fragment at *m/z* 69 ($[C_4H_5O]^+$), arising from cleavage of the methacrylate moiety and serving as a diagnostic ion for this chemical class, as can be seen from the high-resolution mass spectrum of 10:2 FTMAc in Figure 6.5. The fact that this ion can also be seen in PCI reflects the weak C-O bond due to the nearest density electronic groups. Beyond the molecular ion, fragmentation predominantly reflects homologous mass shifts, corresponding to sequential C_2F_4 additions along the fluorotelomer chain, generating series such $[(C_{n-1}H_5F_{2n-11}O)+H]^+$, $[C_nH_5F_{2n-11}O_2]^+$, and the adduct $[(C_nH_9F_{2n-11}O_2)+C_2H_5]^+$ and $[(C_nH_9F_{2n-11}O_2)+C_3H_5]^+$ (n = 10, 12, 14, 16).

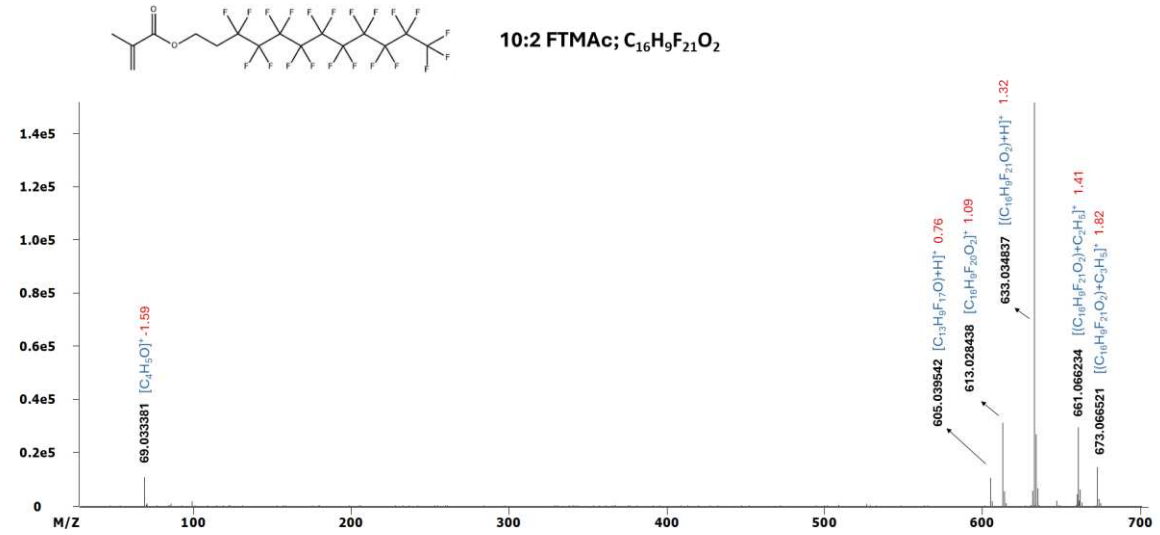


Figure 6.5. *PCI-(HR)MS spectrum of the heneicosafluorododecylethyl methacrylate (10:2 FTMAc, $C_{16}H_9F_{21}O_2$). The figure highlights, in blue, the proposed molecular formulas for the fragments based on the red mass accuracy values.*

In the case of FTIs, base peaks arise from the loss of a fluorine atom and the formation of the characteristic $[C_nH_4F_{2n-4}I]^+$ ion (n = 8, 10, 12). In all congeners, the molecular ions are strongly retained and detected at high intensity at *m/z* 473, 573, and 673 for $[C_8H_4F_{13}I]^+$, $[C_{10}H_4F_{17}I]^+$, and $[C_{12}H_4F_{21}I]^+$, consistent with limited fragmentation under PCI mode. Additional homologous series arise from the progressive addition of C_2F_4 units along the perfluoroalkyl chain, producing both $[C_nH_4F_{2n-5}]^+$ fragments and their protonated counterparts $[(C_nH_4F_{2n-4})+H]^+$. Furthermore, typical methane-PCI chemistry is reflected by the formation of adduct ions such as $[(C_nH_4F_{2n-3}I)+C_2H_5]^+$ and $[(C_nH_4F_{2n-3}I)+C_3H_5]^+$ (n = 8, 10, 12), which systematically appear across the series.

FDIAs exhibit generally low overall intensity, where base peaks for all congeners correspond to the loss of one iodine atom, yielding characteristic $[C_nF_{2n}I]^+$ (n = 4, 6, 8) fragments. At the same time, the molecular ion is consistently preserved and detected at *m/z* 453, 553, and 653 ($[C_nF_{2n}I_2]^+$ n = 4, 6, 8) for the homologues. Additional homologous series arise from sequential C_2F_4 additions, producing ions corresponding to generalized molecular formulas such as $[(C_nF_{2n-1}I)+H]^+$, $[C_nF_{2n-1}I_2]^+$, and $[(C_nF_{2n}I_2)+C_2H_5]^+$ (n = 4, 6, 8).

Finally, FIAs, like the previous class, exhibit very low ionization efficiencies, resulting in overall weak spectra. For PFDeI, the base peak corresponds to the M+29 adduct ion $[(C_{10}F_{21}I)+C_2H_5]^+$, while the typical M+41 ion is not observed. In contrast, PFDoI shows its most intense signal at the fragment $[C_{11}F_{21}]^+$, formed by the loss of iodine. Moreover, across homologues, characteristic adduct series are observed for $[(C_nF_{2n+1}I)+C_2H_5]^+$ (n = 10, 12), together with non-iodinated fragments $[C_nF_{2n-1}]^+$ (n = 9, 11).

Negative chemical ionization (NCI)

In the CI plasma, ions of both polarities are formed simultaneously (*e.g.,* $[M+H]^+$ and $[M-H]^-$), and it is solely the polarity of the acceleration voltage that determines which ions are extracted from the ion source [32]. When operating in NCI mode, this selection results in PFAS spectra with minimal fragmentation, as expected for this soft ionization mode. Unlike FAcs and FTMAcs, which display carbon-based fragmentation pathways, all iodinated PFAS classes (FIA, FTI, and FDIA) yield only the highly intense iodine anion $[I]^-$, with no significant contributions from the perfluoroalkyl chain. Although the prominent iodine fragment enables very sensitive detection, NCI spectra remain non-discriminatory among

iodine PFAS classes, as they do not provide structural information capable of differentiating one iodinated PFAS group from another.

The NCI spectrum of 1,2,4-TFB of the FAc is dominated exclusively by the deprotonated molecular ion $[M-1]^-$, observed as the characteristic fragment $[C_6H_2F_3]^-$, with no additional carbon-based fragmentation. In contrast, BrPFB displays either in NCI mode bromine isotopic pattern, with intense ions at *m/z* 246 and 248 corresponding to $[C_6{}^{79}BrF_5]^-$ and $[C_6{}^{81}BrF_5]^-$, respectively. The spectrum also shows the free halide ions $[^{79}Br]^-$ and $[^{81}Br]^-$ conditions. Additionally, BrPFB generates the highly stable $[C_6F_5]^-$ fragment at *m/z* 167, reflecting the preservation of the aromatic ring under NCI conditions.

Under NCI conditions, FTMAcs display a markedly different fragmentation behavior compared to the other PFAS classes. The spectra are dominated by *m/z* 85 anionic fragment $[C_4H_5O_2]^-$, which represents the base peak for the 4:2, 6:2, and 8:2 congeners. The resonance stabilization of this species could facilitate its elimination from the molecule, as illustrated in Figure 6.6.

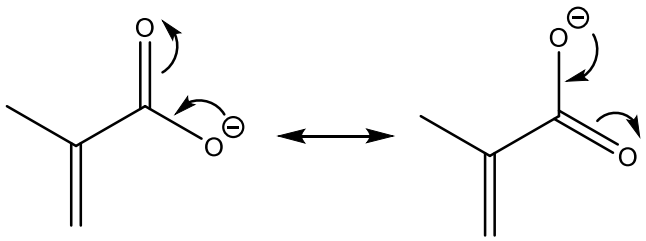


Figure 6.6. *Stabilization by resonance of the fragment $[C_4H_5O_2]^-$.*

Moreover, as the molecular weight increases, the abundance of the fragment $[C_4H_6FO_2]^-$ at *m/z* 105 becomes progressively higher, ultimately becoming the base ion for the 10:2 FTMAc (Figure 6.7). Beyond these diagnostic fragments, the NCI spectra exhibit a homologous series corresponding to successive C_2F_4 iterations. These include the series $[C_nHF_{2n-7}O]^-$, $[C_nH_2F_{2n-6}O]^-$, $[C_nH_2F_{2n-5}O]^-$, $[C_nH_2F_{2n-4}O]^-$, and $[C_nH_4F_{2n-3}O]^-$ (with n = 6, 8, 10, 12).

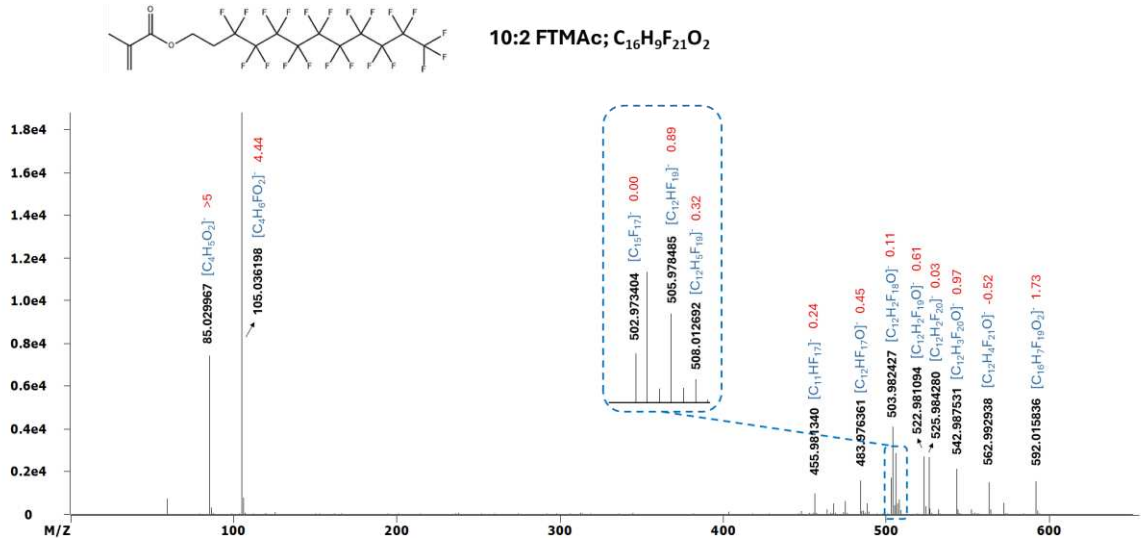


Figure 6.7. *NCI-(HR)MS spectrum of the heneicosafluorododecylethyl methacrylate (10:2 FTMAc, $C_{16}H_9F_{21}O_2$). According to the previous figure, the proposed fragmentation molecular formulas are shown in blue, corresponding to the red mass accuracy values.*

As previously described, NCI fragmentation of iodinated PFAS (FTIs, FDIAs, and FIAs) yields almost exclusively the intense iodide ion $[I]^-$ at *m/z* 127, resulting in non-discriminant spectra that do not allow differentiation among congeners based solely on their mass spectral peaks. The only relevant trend emerges in the FIAs, which exhibit a homologous $[C_nF_{2n}]^-$ series (n = 10-12), arising from the concurrent loss of fluorine and iodine.

A general summary of the base ion characteristics across the three ionization modes is provided in Tables 6.2 (EI), 6.3 (PCI), and 6.4 (NCI).

Table 6.2. *Base ion fragment summary obtained in EI-(HR)MS mode per PFAS chemical class. BP m/z: Base Peak m/z; Acc.: Mass accuracy in ppm.*

Compound	Formula	Base ion	BP *m/z*	Acc. (ppm)
1,2,4-TFB	C_6BrF_5	$[C_6H_3F_3]^+$	132.018039	-0.74
BrPFB	$C_6H_3F_3$	$[C_6BrF_5]^+$	245.909845	0.16
4:2 FTMAc	$C_{10}H_9F_9O_2$	$[C_4H_5O]^+$	69.033469	-0.33
6:2 FTMAc	$C_{12}H_9F_{13}O_2$	$[C_4H_5O]^+$	69.033448	-0.63
8:2 FTMAc	$C_{14}H_9F_{17}O_2$	$[C_4H_5O]^+$	69.033438	-0.77
10:2 FTMAc	$C_{16}H_9F_{21}O_2$	$[C_4H_5O]^+$	69.033434	-0.82
6:2 FTI	$C_8H_4F_{13}I$	$[C_8H_4F_{13}I]^+$	473.914396	-0.15
8:2 FTI	$C_{10}H_4F_{17}I$	$[CF_3]^+$	68.994634	-0.39
10:2 FTI	$C_{12}H_4F_{21}I$	$[CF_3]^+$	68.994633	-0.41
PFBuDiI	$C_4F_8I_2$	$[C_4F_8I]^+$	326.891236	0.26
PFHxDiI	$C_6F_{12}I_2$	$[I]^+$	126.903870	-0.43
PFODiI	$C_8F_{16}I_2$	$[I]^+$	126.903865	-0.47

PFDeI	$C_{10}F_{21}I$	$[CF_3]^+$	68.994634	-0.39
PFDoI	$C_{12}F_{25}I$	$[CF_3]^+$	68.994621	-0.58

Table 6.3. *Base ion fragment summary obtained in PCI-(HR)MS mode per PFAS chemical class. BP m/z: Base Peak m/z; Acc.: Mass accuracy in ppm.*

Compound	Formula	Base ion	BP *m/z*	Acc. (ppm)
1,2,4-TFB	C_6BrF_5	$[M+H]^+$	133.025964	0.02
BrPFB	$C_6H_3F_3$	$[M+H]^+$	246.917756	0.51
4:2 FTMAc	$C_{10}H_9F_9O_2$	$[M+H]^+$	333.053531	1.12
6:2 FTMAc	$C_{12}H_9F_{13}O_2$	$[M+H]^+$	433.048299	3.52
8:2 FTMAc	$C_{14}H_9F_{17}O_2$	$[M+H]^+$	533.041031	1.12
10:2 FTMAc	$C_{16}H_9F_{21}O_2$	$[M+H]^+$	633.034837	1.32
6:2 FTI	$C_8H_4F_{13}I$	$[M-F]^+$	454.916359	0.65
8:2 FTI	$C_{10}H_4F_{17}I$	$[M-F]^+$	554.910645	1.75
10:2 FTI	$C_{12}H_4F_{21}I$	$[M-F]^+$	654.903672	0.58
PFBuDiI	$C_4F_8I_2$	$[M-I]^+$	326.891528	1.15
PFHxDiI	$C_6F_{12}I_2$	$[M-I]^+$	426.884585	0.42
PFODiI	$C_8F_{16}I_2$	$[M-I]^+$	526.879262	1.68
PFDeI	$C_{10}F_{21}I$	$[M+C_2H_5]^+$	674.910057	0.80
PFDoI	$C_{12}F_{25}I$	$[C_{11}F_{21}]^+$	530.965874	-0.09

Table 6.4. *Base ion fragment summary obtained in PCI-(HR)MS mode per PFAS chemical class. BP m/z: Base Peak m/z; Acc.: Mass accuracy in ppm.*

Compound	Formula	Base ion	BP *m/z*	Acc. (ppm)
1,2,4-TFB	C_6BrF_5	$[M-H]^-$	131.011518	0.84
BrPFB	$C_6H_3F_3$	$[M]^-$	245.911144	0.98
4:2 FTMAc	$C_{10}H_9F_9O_2$	$[C_4H_5O_2]^-$	85.029503	5.50
6:2 FTMAc	$C_{12}H_9F_{13}O_2$	$[C_4H_5O_2]^-$	85.029937	5.11
8:2 FTMAc	$C_{14}H_9F_{17}O_2$	$[C_4H_5O_2]^-$	85.029828	3.82
10:2 FTMAc	$C_{16}H_9F_{21}O_2$	$[C_4H_6FO_2]^-$	105.036198	4.44
6:2 FTI	$C_8H_4F_{13}I$	$[I]^-$	126.905136	0.90
8:2 FTI	$C_{10}H_4F_{17}I$	$[I]^-$	126.905132	0.87
10:2 FTI	$C_{12}H_4F_{21}I$	$[I]^-$	126.905201	1.41

PFBuDiI	$C_4F_8I_2$	$[I]^-$	126.905181	1.26
PFHxDiI	$C_6F_{12}I_2$	$[I]^-$	126.905168	1.16
PFODiI	$C_8F_{16}I_2$	[I]-	126.905169	1.20
PFDeI	$C_{10}F_{21}I$	[I]-	126.905151	0.59
PFDoI	$C_{12}F_{25}I$	[I]-	126.905196	1.37

As shown in the electron ionization section, excellent mass accuracy, below 1 ppm, enabled the reliable assignment of the molecular formulas associated with the base peaks. The same applies to PCI and NCI, except for the methacrylates in NCI, where the mass calibration table, necessary for tuning the time-of-flight, begins from fragment C_4F_8 of the calibrant at *m/z* 200 (nominal mass). Considering this, non-optimal mass accuracies (from 3.82 to 5.5ppm) for *m/z* values well below 200 are due to the extension of the mass calibration curve beyond its valid calibration range.

6.3.2. Ion source performance comparison

As previously described, the PFAS mixture solution was injected across multiple ionization modes (EI, PCI, and NCI) to evaluate the analytical response in terms of ionization efficiency. A visualization of this comparison is given in Figure 6.8 using both the base ion and total ion current (TIC) S/N ratios.

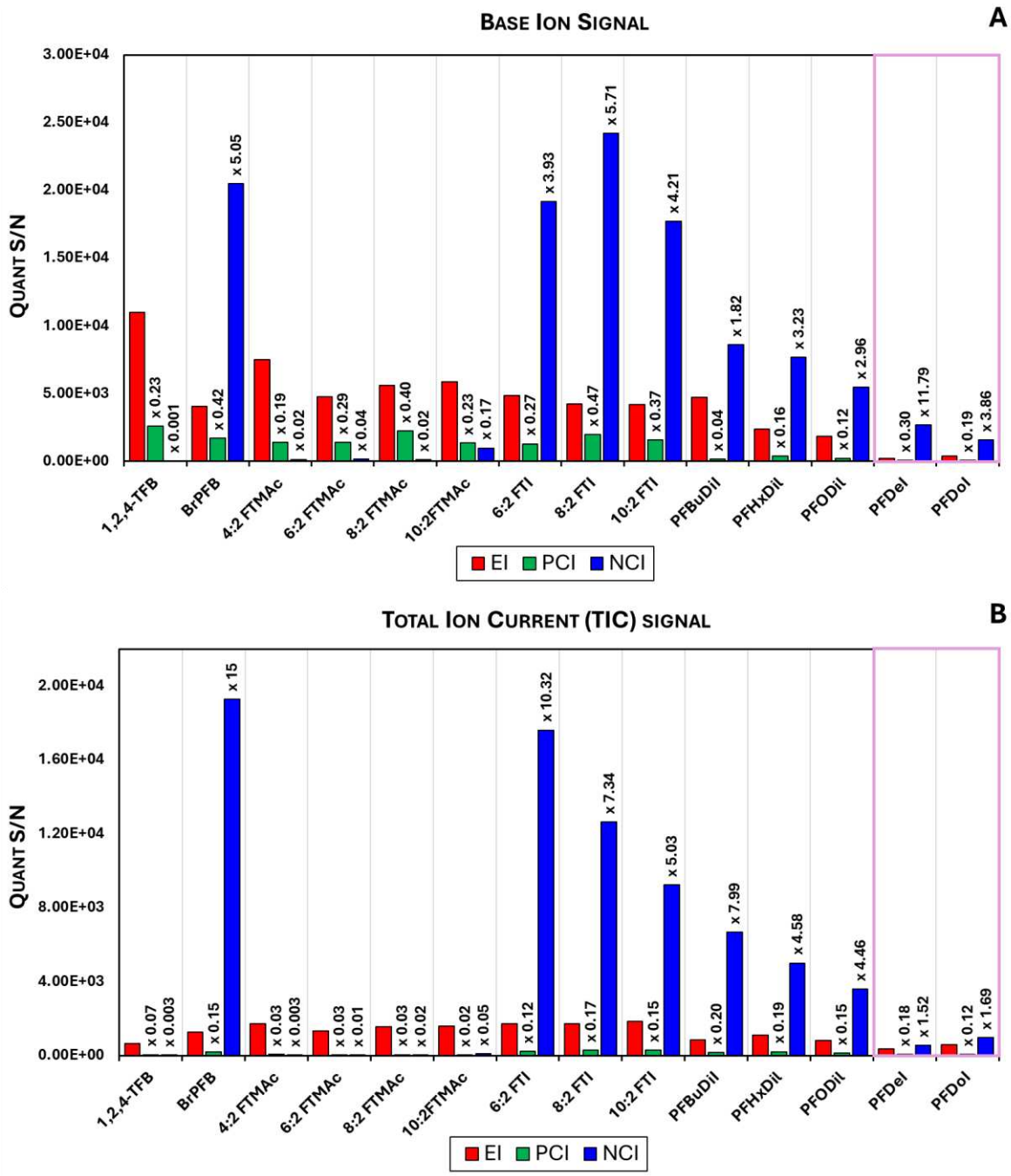


Figure 6.8. *Comparison of the ionization efficiency across multiple ionization modes in terms of signal-to-noise ratio, calculated based on the base ion signal (A) and the total ion current signal (B) in GC-(HR)MS. All compounds are at 50 µg/mL, while fluorinated iodine alkanes (FIAs) are highlighted in pink at 30 µg/mL (due to poor solubility of the stock standards in methanol).*

The comparative evaluation reveals marked differences in analytical response among the various PFAS classes. In particular, NCI delivers the highest S/N ratios for most of the PFAS investigated. This behavior is evident for iodine- and bromine-containing standards (FTIs, FDIAs, FIAs, and BrPFB), where the high electron-capture affinity results in the enhancements of the signal intensities from one order to a twelve-fold magnitude with respect to EI. Although structurally informative, EI produced lower S/N ratios for these iodinated and brominated PFAS. Nevertheless, for the methacrylate series and 1,2,4-trifluorobenzene, EI remains the most effective ionization method. In particular, FTMAcs compounds, lacking highly electronegative moieties, do not benefit from NCI, and don't ionize efficiently under PCI either. Positive chemical ionization (PCI) represents the least

effective of the three ionization modes examined. Essentially, the low proton affinity limits the formation of stable $[M]^+$ and $[M+H]^+$ or adduct ions, resulting in weak signals across all PFAS chemical classes. Moreover, TIC-based evaluation in Figure 6.1B further confirms these trends. Essentially, while base ion integrates a single diagnostic fragment, TIC captures the overall ionization efficiency of each source, providing a general representation of total signal output. Under this signal, NCI still provides the strongest overall response for halogenated compounds, followed by EI, whereas PCI again exhibits lower signal outputs.

The limits of quantification (LOQ) were also determined for each PFAS standard under EI, PCI, and NCI, following the Eurachem guidelines [37], using the blank standard deviation. The resulting LOQs, expressed in pg/µL (ppb) on a logarithmic scale, are presented in Figure 6.9.

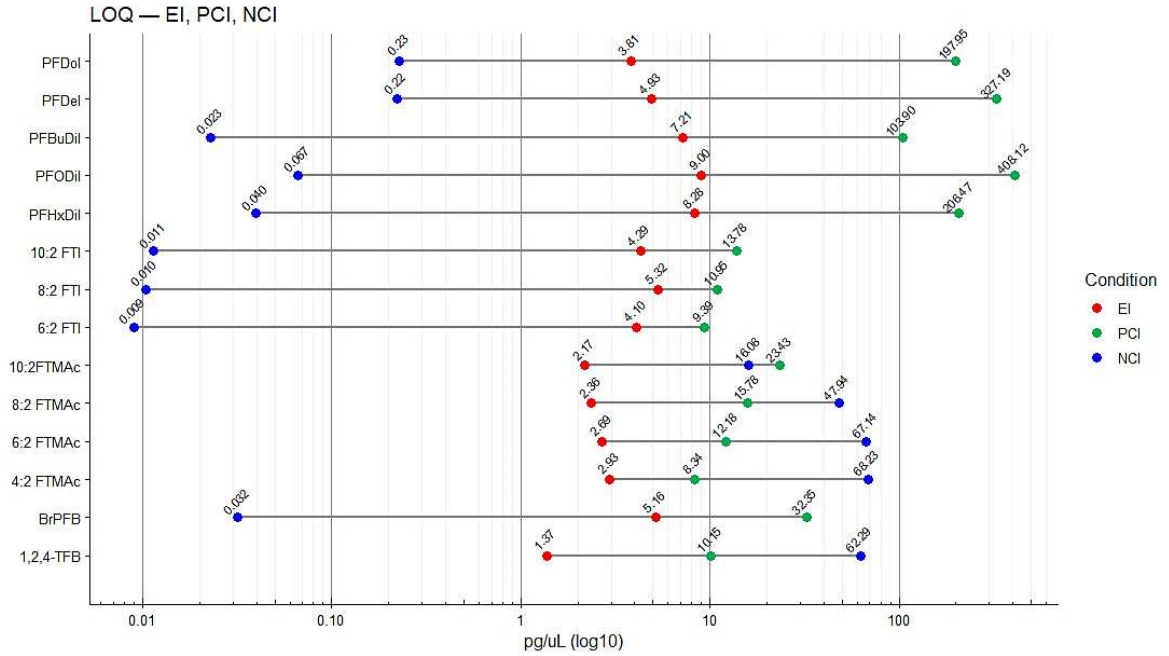


Figure 6.9. *LOQ values across different ionization modes (EI, PCI, and NCI). PFAS standards are listed on the y-axis, ordered by chemical class, while the x-axis presents LOQs in pg/µL according to a logarithmic scale in order to restrict the visualization of the differences.*

As shown in the plot, NCI consistently provides the lowest LOQs for iodinated and brominated PFAS, yielding values between approximately 9 ppt and 0.2 ppb, in line with the higher electron-capture of heavily halogenated species under NCI. EI produces intermediate LOQs, generally ranging between 9 ppb and 1.37 ppb for methacrylate derivatives. Conversely, PCI yields the highest LOQs overall, often in the sub-ppm range for heavily halogenated PFAS, reflecting its poor ionization efficiency toward analytes with low proton affinity. Nevertheless, PCI outperforms NCI for several methacrylates and for 1,2,4-TFB.

Finally, some measurements were also conducted in GC×GC across the different ionization modes. The signal-to-noise ratio obtained in GC and GC×GC modes essentially yields a pair of values that can be displayed as a scatter plot (Figure 6.10).

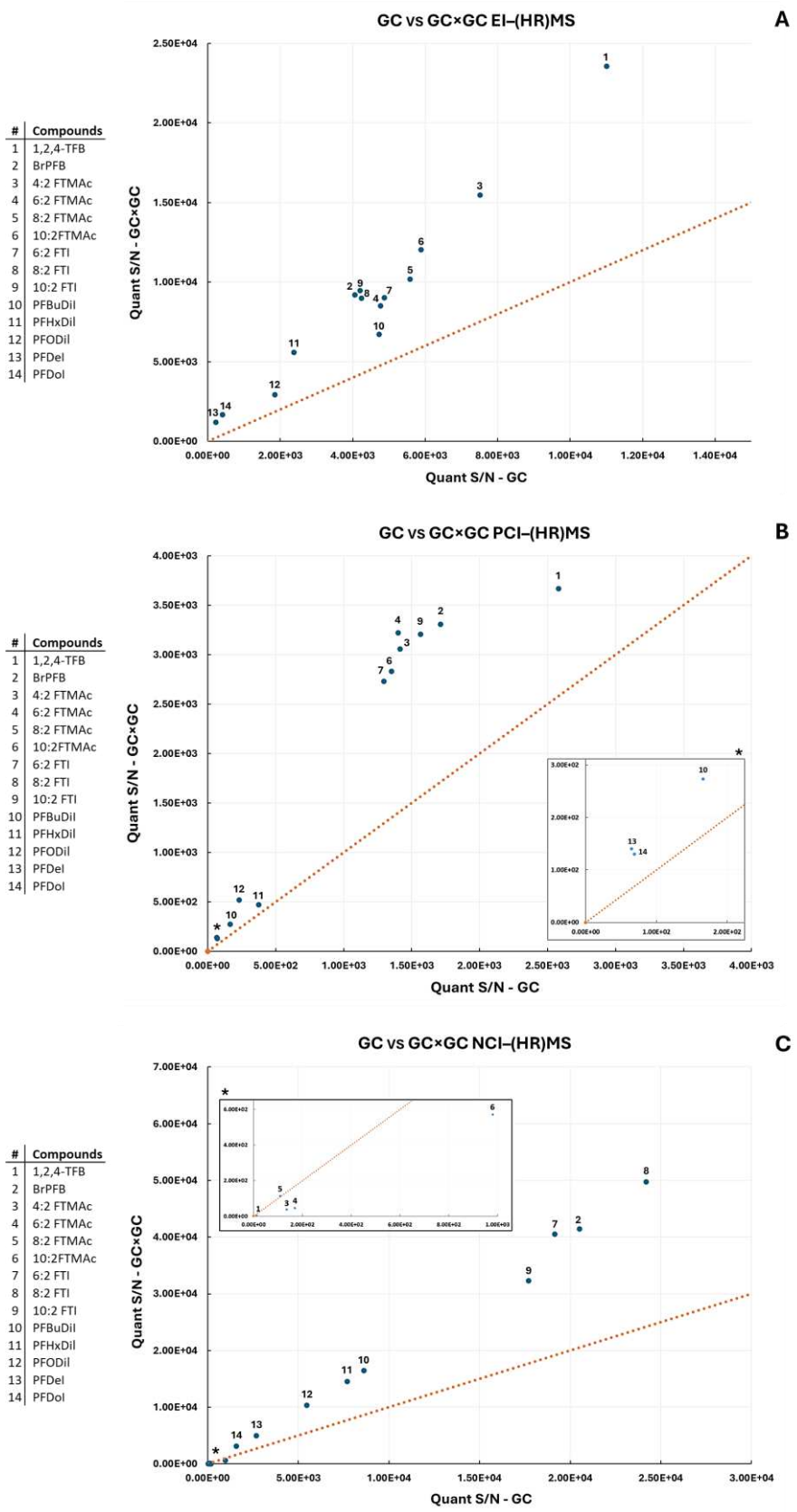


Figure 6.10. *S/N comparison of GC and GC×GC-(HR)MS across electron ionization (A), positive chemical ionization (B), and negative chemical ionization (C).*

Across all ionization modes (EI, PCI, and NCI), GC×GC demonstrates an enhancement in the S/N values over conventional one-dimensional GC. Every analyte of Figure 6.10A (EI) and 6.10B (PCI) lies above the 1:1 reference line, indicating that GC×GC systematically

increases S/N through the compression of the chromatographic band. Also, in NCI (Figure 6.10C), most of the compounds show higher S/N in GC×GC. However, for a small subset of analytes, GC does not provide the expected increase in signal-to-noise ratio. This reflects the case where there is no chemical background within a selective mass window, and the detection limits of GC and GC×GC are essentially the same or lower for GC×GC. However, when chemical noise is present and a baseline exists, the S/N can differ between GC and GC×GC. Overall, even in light of this, the combination of GC×GC and multiple ionization modes underscores its superior detectability, highlighting its potential to improve trace-level PFAS identification and quantification.

Considering the previous discussion, the question "*Which ionization mode is best suited for studying PFAS?"* may appear natural. However, this question depends on a fundamental assumption; specifically, that different ionization techniques are treated as separate compartments, and selecting one necessarily excludes other ionization modes (Figure 11A).

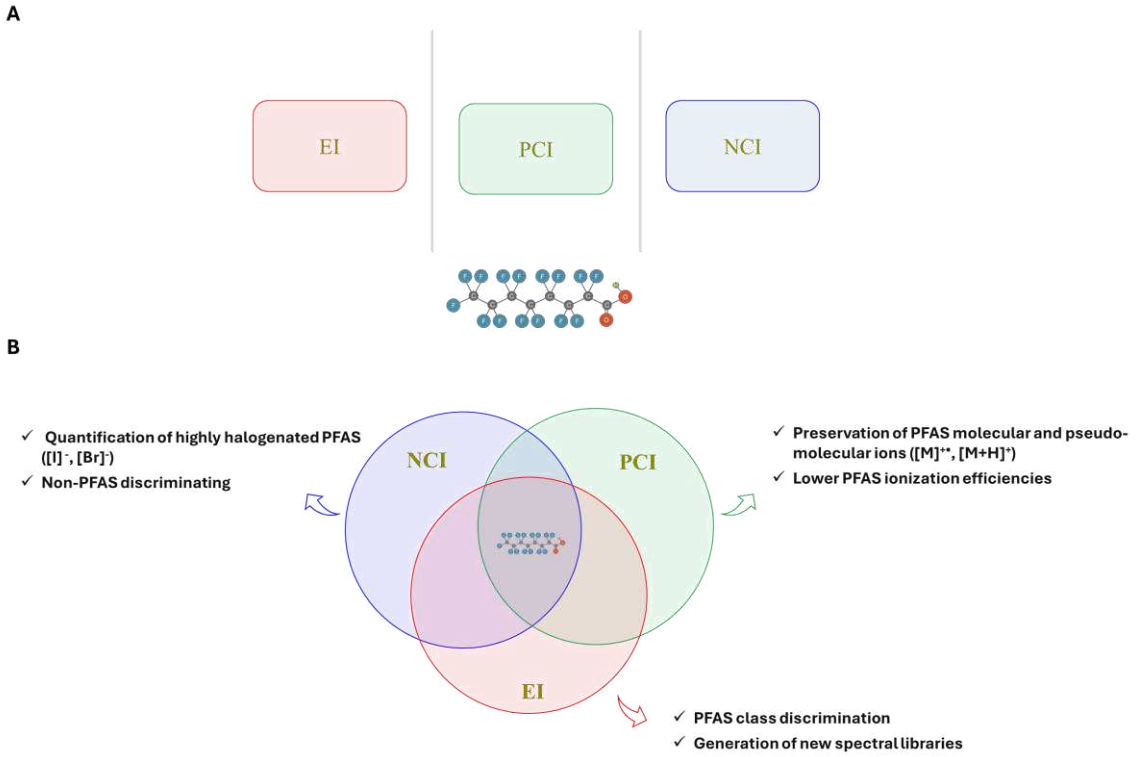


Figure 6.11. *Isolated (A) and complementary (B) views of different ionization modes.*

A more appropriate question is instead: "*What kind of information does each ionization source provide?".* In light of this, EI, PCI, and NCI should be viewed not as competing alternatives but as complementary tools (Figure 11B). Especially in non-targeted analysis, where the expansion of the observable sample chemical space is the goal, a combined use of hard (EI) and soft ionization (PCI and NCI) represents a better approach to maximizing the continuity of the analytical information (AIC) obtainable from PFAS. NCI enables highly sensitive quantification of iodine- and bromine-containing PFAS due to their strong electron-capture affinity, although its spectra are inherently non-discriminatory across PFAS classes and are often dominated solely by halide ions. PCI, by contrast, simplifies spectral interpretation through the preservation of molecular and pseudo-molecular ions ($[M]^{+\bullet}$ and $[M+1]^{+}$), although ionization efficiencies are generally lower for most of the PFAS examined. Finally, EI offers the most structurally informative patterns, producing extensive fragmentation that supports PFAS class discrimination and facilitates the generation of new spectral libraries.

6.4. Conclusions

A comprehensive spectral characterization of the fragmentation patterns of multiple PFAS chemical classes was conducted using both hard (EI) and soft (PCI and NCI) ionization under GC-(HR)MS conditions. For each PFAS class, a variety of diagnostic fragments was observed. In EI, numerous common ions for all probe compounds, were detected, originating primarily from the fragmentation of the fluorinated carbon backbone (*e.g.*, *m/z* 100, 119, 131, 169, 181, 219 and 269, corresponding to $[C_2F_4]^+$, $[C_2F_5]^+$, $[C_3F_5]^+$, $[C_3F_7]^+$, $[C_4F_7]^+$, $[C_4F_9]^+$, and $[C_5F_{11}]^+$, respectively). EI also produced class specific fragments, including iodine ions ($[I]^+$ for FTIs and FIAs, $[I_2]^+$ for FDIAs), as well as diagnostic fragments such as $[C_4H_5O]^+$ for FTMAcs and $[CF_3]^+$ for FTIs and FIAs. In PCI, as expected, spectra were dominated by $[M+H]^+$ ions arising from proton transfer reactions driven by $[CH_5]^+$. Moreover, in PCI, methane ions efficiently generates electrophilic addition products, yielding intense $[M+C_2H_5]^+$ and $[M+C_3H_5]^+$ species. NCI spectra, by contrast, consist almost exclusively of the intense $[I]^-$ ion at *m/z* 126 for all iodinated PFAS. FTMAcs, however, exhibited a moderate degree of fragmentation, with a diagnostic $[C_4H_5O_2]^-$ fragment at *m/z* 85. Overall, the fragments formed across multiple ionization modes reflect the ability (or inability) of the molecule's moieties to stabilize negative or positive charge through electron rearrangement. This study also revealed several iterative mass series associated with the addition of CF_2 units, highlighting the potential to investigate these compounds through the Kendrick Mass Defect.

Analytical response was also assessed using the signal-to-noise ratio of the base ion for each compound. Iodine- and bromine-containing species showed markedly higher ionization efficiency in NCI, achieving quantification limits in the ng/L (ppt) range. For the other PFAS classes, EI remained the most efficient ionization mode, while GC×GC analysis provided an additional detectability boost for PFAS detection. Overall, this study emphasizes that each ionization mode brings intrinsic advantages and limitations, and a complementary strategy of EI, PCI, and NCI enables the broadest expansion of PFAS chemical space from a non-targeted workflow perspective. EI essentially remains the cornerstone for PFAS structural elucidation due to its harmonization across laboratories and its highly reproducible fragmentation. Soft ionization techniques, however, offer valuable complementary information. In this sense, PCI is advantageous thanks to the preservation of molecular and pseudomolecular ions, although its lower ionization efficiency limits its utility for non-targeted analysis. NCI, in contrast, provides higher sensitivity for highly halogenated PFAS, delivering LOQ values at least two orders of magnitude lower than those of EI. Finally, considering that these enhancements can be achieved by simply switching ionization modes without essentially altering the chromatographic configuration, the complementarity of multiple ionization modes further strengthens their applicability in non-targeted PFAS analysis.

References

[1] M. Ateia, A. Maroli, N. Tharayil, and T. Karanfil, "The overlooked short- and ultrashort-chain poly- and perfluorinated substances: A review," Apr. 01, 2019, *Elsevier Ltd*. doi: 10.1016/j.chemosphere.2018.12.186.

[2] I. T. Cousins *et al.*, "Strategies for grouping per-and polyfluoroalkyl substances (PFAS) to protect human and environmental health," Jul. 01, 2020, *Royal Society of Chemistry*. doi: 10.1039/d0em00147c.

[3] S. Kurwadkar *et al.*, "Per- and polyfluoroalkyl substances in water and wastewater: A critical review of their global occurrence and distribution," Feb. 25, 2022, *Elsevier B.V.* doi: 10.1016/j.scitotenv.2021.151003.

[4] M. K. Wijayahena, R. Kachangoon, C. Witmer, J. S. Wallace, J. Vichapong, and D. S. Aga, "PFAS Isomers Matter: Distribution Patterns of Linear and Branched PFOS and PFOA in Consumed Fish Revealed by Cyclic Ion Mobility Separation," *J. Agric. Food Chem.*, vol. 73, no. 41, pp. 25967–25977, Oct. 2025, doi: 10.1021/acs.jafc.5c08608.

[5] J. Wang *et al.*, "Critical Review of Thermal Decomposition of Per- and Polyfluoroalkyl Substances: Mechanisms and Implications for Thermal Treatment Processes," May 03, 2022, *American Chemical Society*. doi: 10.1021/acs.est.2c02251.

[6] E. M. Sunderland, X. C. Hu, C. Dassuncao, A. K. Tokranov, C. C. Wagner, and J. G. Allen, "A review of the pathways of human exposure to poly- and perfluoroalkyl substances (PFASs) and present understanding of health effects," Mar. 01, 2019, *Nature Publishing Group*. doi: 10.1038/s41370-018-0094-1.

[7] D. Dias, J. Bons, A. Kumar, M. H. Kabir, and H. Liang, "Forever Chemicals, Per-and Polyfluoroalkyl Substances (PFAS), in Lubrication," Apr. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/lubricants12040114.

[8] K. Dasu, X. Xia, D. Siriwardena, T. P. Klupinski, and B. Seay, "Concentration profiles of per- and polyfluoroalkyl substances in major sources to the environment," Jan. 01, 2022, *Academic Press*. doi: 10.1016/j.jenvman.2021.113879.

[9] J. Glüge *et al.*, "An overview of the uses of per- And polyfluoroalkyl substances (PFAS)," *Environ. Sci. Process. Impacts*, vol. 22, no. 12, pp. 2345–2373, Dec. 2020, doi: 10.1039/d0em00291g.

[10] J. Glüge *et al.*, "Information Requirements under the Essential-Use Concept: PFAS Case Studies," *Environ. Sci. Technol.*, vol. 56, no. 10, pp. 6232–6242, Oct. 2021, doi: 10.1021/acs.est.1c03732.

[11] C. Bach *et al.*, "Simultaneous determination of perfluoroalkyl iodides, perfluoroalkane sulfonamides, fluorotelomer alcohols, fluorotelomer iodides and fluorotelomer acrylates and methacrylates in water and sediments using solid-phase microextraction-gas chromatography/mass spectrometry," *J. Chromatogr. A*, vol. 1448, pp. 98–106, May 2016, doi: 10.1016/j.chroma.2016.04.025.

[12] S. Mok, S. Lee, Y. Choi, J. Jeon, Y. H. Kim, and H. B. Moon, "Target and non-target analyses of neutral per- and polyfluoroalkyl substances from fluorochemical industries using GC-MS/MS and GC-TOF: Insights on their environmental fate," *Environ. Int.*, vol. 182, Dec. 2023, doi: 10.1016/j.envint.2023.108311.

[13] H. Holmquist, S. Schellenberger, I. van der Veen, G. M. Peters, P. E. G. Leonards, and I. T. Cousins, "Properties, performance and associated hazards of state-of-the-art durable water repellent (DWR) chemistry for textile finishing," May 01, 2016, *Elsevier Ltd*. doi: 10.1016/j.envint.2016.02.035.

[14] A. Podder, A. H. M. A. Sadmani, D. Reinhart, N. Bin Chang, and R. Goel, "Per and poly-fluoroalkyl substances (PFAS) as a contaminant of emerging concern in surface water: A transboundary review of their occurrences and toxicity effects," Oct. 05, 2021, *Elsevier B.V.* doi: 10.1016/j.jhazmat.2021.126361.

[15] E. Panieri, K. Baralic, D. Djukic-Cosic, A. B. Djordjevic, and L. Saso, "PFAS Molecules: A Major Concern for the Human Health and the Environment," Feb. 01, 2022, *MDPI*. doi: 10.3390/toxics10020044.

[16] J. W. Washington, T. M. Jenkins, K. Rankin, and J. E. Naile, "Decades-scale degradation of commercial, side-chain, fluorotelomer-based polymers in soils and water," *Environ. Sci. Technol.*, vol. 49, no. 2, pp. 915–923, Jan. 2015, doi: 10.1021/es504347u.

[17] S. E. Fenton *et al.*, "Per- and Polyfluoroalkyl Substance Toxicity and Human Health Review: Current State of Knowledge and Strategies for Informing Future Research," Mar. 01, 2021, *Wiley Blackwell*. doi: 10.1002/etc.4890.

[18] G. Jha *et al.*, "Per-and polyfluoroalkyl substances (PFAS) in integrated crop–livestock systems: Environmental exposure and human health risks," *Int. J. Environ. Res. Public Health*, vol. 18, no. 23, Dec. 2021, doi: 10.3390/ijerph182312550.

[19] K. Singh, N. Kumar, A. Kumar Yadav, R. Singh, and K. Kumar, "Per-and polyfluoroalkyl substances (PFAS) as a health hazard: Current state of knowledge and strategies in environmental settings across Asia and future perspectives," Nov. 01, 2023, *Elsevier B.V.* doi: 10.1016/j.cej.2023.145064.

[20] B. E. Blake and S. E. Fenton, "Early life exposure to per- and polyfluoroalkyl substances (PFAS) and latent health outcomes: A review including the placenta as a target tissue and possible driver of peri- and postnatal effects," Oct. 01, 2020, *Elsevier Ireland Ltd*. doi: 10.1016/j.tox.2020.152565.

[21] J. T. Szilagyi, V. Avula, and R. C. Fry, "Perfluoroalkyl Substances (PFAS) and Their Effects on the Placenta, Pregnancy, and Child Development: a Potential Mechanistic Role for Placental Peroxisome Proliferator–Activated Receptors (PPARs)," Sep. 01, 2020, *Springer*. doi: 10.1007/s40572-020-00279-0.

[22] A. Androulakakis, N. Alygizakis, E. Bizani, and N. S. Thomaidis, "Current progress in the environmental analysis of poly- and perfluoroalkyl substances (PFAS)," Dec. 01, 2022, *Royal Society of Chemistry*. doi: 10.1039/d2va00147k.

[23] L. Sankhla, P. C. Ramamurthy, S. R. Wickramasinghe, and C. Thamaraiselvan, "A comprehensive review of detection techniques for PFAS compounds in aqueous solution," *Journal of Water Process Engineering*, vol. 77, Sep. 2025, doi: 10.1016/j.jwpe.2025.108569.

[24] S. R. Newton *et al.*, "Filling the Gaps in PFAS Detection: Integrating GC-MS Non-Targeted Analysis for Comprehensive Environmental Monitoring and Exposure Assessment," Feb. 11, 2025, *American Chemical Society*. doi: 10.1021/acs.estlett.4c00930.

[25] D. Zanella, J. Focant, and F. A. Franchina, " 30 th Anniversary of comprehensive two-dimensional gas chromatography: Latest advances ," *Analytical Science Advances*, vol. 2, no. 3–4, pp. 213–224, 2021, doi: 10.1002/ansa.202000142.

[26] J. P. Koelmel *et al.*, "Shedding Light on PFAS Dark Matter Using a Novel GC-HRMS Approach," *Environ. Sci. Technol.*, Dec. 2025, doi: 10.1021/acs.est.5c11712.

[27] G. Black *et al.*, "Exploring chemical space in non-targeted analysis: a proposed ChemSpace tool," *Anal. Bioanal. Chem.*, vol. 415, no. 1, pp. 35–44, Jan. 2023, doi: 10.1007/s00216-022-04434-4.

[28] J. Hollender *et al.*, "NORMAN guidance on suspect and non-target screening in environmental monitoring," Dec. 01, 2023, *Springer*. doi: 10.1186/s12302-023-00779-4.

[29] T. Kind and O. Fiehn, "Advances in structure elucidation of small molecules using mass spectrometry," 2010. doi: 10.1007/s12566-010-0015-9.

[30] L. Hejazi, D. Ebrahimi, D. B. Hibbert, and M. Guilhaus, "Compatibility of electron ionization and soft ionization methods in gas chromatography/orthogonal time-of-flight mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 23, no. 14, pp. 2181–2189, Jul. 2009, doi: 10.1002/rcm.4131.

[31] T. Portolés, E. Pitarch, F. J. López, F. Hernández, and W. M. A. Niessen, "Use of soft and hard ionization techniques for elucidation of unknown compounds by gas chromatography/time-of-flight mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 25, no. 11, pp. 1589–1599, Jun. 2011, doi: 10.1002/rcm.5028.

[32] J. H. Gross, *Mass Spectrometry*, Second Edition. Springer, 2017. doi: 10.1007/978-3-319-54398-7.

[33] O. David. Sparkman, Zelda. Penton, and F. G. . Kitson, *Gas chromatography and mass spectrometry : a practical guide*. Academic Press, 2011.

[34] L. Hejazi, M. Guilhaus, D. B. Hibbert, and D. Ebrahimi, "Gas chromatography with parallel hard and soft ionization mass spectrometry," *Rapid Communications in Mass Spectrometry*, vol. 29, no. 1, pp. 91–99, Dec. 2014, doi: 10.1002/rcm.7091.

[35] H. Hübschmann, *Handbookof GC-MS, Fundamentals and Applications*, Third edition. Wiley-VCH Verlag GmbH & Co.KGaA, 2015. doi: 10.1002/9783527625215.ch5.

[36] J. T. Watson, *Introduction to Mass Spectrometry Instrumentation, Applications and Strategies for Data Interpretation*, Fourth Edition. John Wiley & Sons Ltd, 2007. doi: https://pubs.acs.org/doi/abs/10.1021/ed086p810.1.

[37] F. W. McLafferty and F. Turecek, *Interpretation of Mass Spectra*, 4th ed. University Science Books, 1993.

[38] H. Cantwell, "The Fitness for Purpose of Analytical Methods A Laboratory Guide to Method Validation and Related Topics Third edition Acknowledgements." Eurachem, 2025. [Online]. Available: www.eurachem.org.

This page was intentionally left blank

List of publications

- **De Poli M.**, Chenet T., Felletti S., Spadafora D. N., Cavazzini A., Franchina F.A.; "*Sorbent-Based Sampling With Two-Stage Trapping/Desorption Coupled to Comprehensive Two-Dimensional Gas Chromatography and Mass Spectrometry for Terpenoids Profiling in Cannabis*"; Analytical Science Advances 6 (2025), e202400044.
- Corviseri M. C., Polidoro A., **De Poli M.**, Stevanin C., Chenet T., D'Anna C., Cavazzini A., Pasti L., Franchina F.A.; *"Targeted Determination of Volatile Fluoroalkyl Pollutants and Non-Targeted Screening for Environmental Monitoring"*; Talanta 292 (2025), 127944.
- **De Poli M.**, Polidoro A., Franchina F.A.; *"Chapter 1: Definitions and concepts for targeted and non-targeted chemical analysis"*; From: "*Gas Chromatography-Mass Spectrometry for Omics Applications*"; Wiley, ISBN: 978-1-394-20927-9, (2026).
- Borgatti D., Radicetti E., **De Poli M**., Franchina F.A., Bertoli T., Tedeschi P., Mancinelli R., Ben Hassine M.; "*Optimizing Wood Vinegar Application to Enhance Frum Wheat (Triticum durum Desf.) Germination and Early Growth Under Salinity Stress*"; Journal of Plant Growth Regulation. Submitted Dec. 2025.
- Triachini S., Franchina F.A., Vezzulli F., Bertuzzi T., **De Poli M.**, Di Stefano R., Gabrielli M., Capri E., Lucini L., Becchi P.P.; "*The Impact of Different Energy-Saving Roasting Curves on the Metabolomics and Sensory Profile of Arabica Coffee: a Multi-Omics Data Fusion Approach*"; Food Chemistry. Submitted Oct. 2025.
- Di Stefano R., **De Poli M.**, Moltrasio C., Marzano A. V., Rimondi E., Melloni E., Secchiero P., Lodi G., Manfredini M., Cavazzini A., Marcuzzi A., Crovella S., Franchina F.A; "*Volatilome Profiling of Inflammation in Skin Fibroblasts for Biomarker Discovery: a proof of principle*"; Submitted Biomolecules Sep. 2025.
- Feerick A., Bade J., Baumann S.A., Baker M.C., Begley A.L., Cabrera A., Chen Y., Clifton M., Reis de Carbalho A., **De Poli M.**, Franchina F.A., Franklin E., Froment J., Hayes J.M., Hanozin E., Klee S., Koelmel J., Kralles Z.T., Letzel T., McDonough C., McGlynn D.F., McNeil B., Nieto S., O'Connor C., Rostkowsky P., Saleeby B., Sunyer-Caldù A., Geer Wallace M.A., Williams A., Lindsay D.Y., Young T., Newton S.E.; "*Non-targeted Analysis with Gas Chromatography: State of the Science and a Path Forward*"; In preparation.
- **De Poli M.**, Di Stefano R., Galliani V., Vulpio C., Franchina F.A; "*Evaluation of Solvent-Dependent Metabolomic Profiles in Human Serum via Comprehensive Two-Dimensional Gas Chromatography and Mass Spectrometry*"; In preparation.
- **De Poli M.**, Corviseri M.C., Polidoro A., Franchina F.A; "*Hard and Soft Ionization HRMS for the Structural Elucidation of Volatile and Semivolatile PFAS*"; In preparation.