

UNIVERSITY OF NAPLES FEDERICO II

Department of Pharmacy



Ph.D. programme in

‘Nutraceuticals, Functional Foods and Human Health’

Food plants secondary metabolites active on Metabolic Syndrome

Elena Serino

Coordinator

Prof. Angelo Antonio Izzo

Tutor

Prof. Carmen Formisano

XXXVII cycle

(2021-2024)

ABSTRACT

Metabolic syndrome (MetS) is a group of obesity-related symptoms that often occur together and significantly increase the risk of developing cardiovascular complications such as stroke and type 2 diabetes mellitus. In 2019, cardiovascular disease was responsible for about 9.6 million deaths in men and 8.9 million deaths in women, accounting for about one-third of all deaths worldwide. In the same year, diabetes and diabetes-related kidney disease contributed to an estimated 2 million deaths worldwide. In the early stages, in order to limit the worsening of MetS symptoms and its pathological development, it is recommended to reduce the risk factors, not through drug therapy, but by changing the patient's daily habits. In this phase, the use of nutraceuticals can be helpful and provide significant health benefits.

During my PhD, I focused my research on several food plants that had been reported in the literature for their anti-diabetic potential and whose phytochemical characterisation had not been described. The isolated secondary metabolites were therefore tested against targets involved in MetS to validate the use of the plants and identify the bioactive ingredients.

This manuscript is divided into four parts, each of which is divided into chapters. The first part is an introduction discussing the main characteristics of MetS and the targets on which the isolated secondary metabolites were tested (Chapter 1), while the second part describes the main techniques I have used in my research (Chapter 2). The second part describes the phytochemical characterisation of three different *Centaurea* species, *C. kotschyi*, *C. sicula* and *C. patula*. The isolation and structural characterisation (HR-MS, 1D and 2D NMR) of 58 pure secondary metabolites are reported, including different structural families of sesquiterpene lactones, lignans and polyphenols. In particular, these analyses lead to the isolation of 14 unprecedented compounds. In addition, the antidiabetic potential of the isolated secondary metabolites was evaluated on GLUT transporters and on the α -glucosidase enzyme, thanks to our collaboration with Professor Mariano Stornaiuolo of the University Federico II and Professor Sonia Piacente of the University of Salerno. The NMR spectra of the previously undescribed metabolites are given in Appendices A, B and C respectively. This research was partly funded by Ministry of Education, Universities and Research, PRIN 2022 PNRR: P2022PPH4C.

The third part of this manuscript describes the analyses of two different food plants: *Glycyrrhiza foetida* and Aglianico grape pomace. In the first chapter of Part 3, the phytochemical composition of the aerial parts of *G. foetida* is reported, where a total of 29 pure compounds

were isolated and characterised, including 12 previously unreported metabolites. The isolated compounds were tested against various targets related to MetS; in particular, the amorfrutins were tested against PPARs targets thanks to our collaboration with Dr Rosa Maria Vitale at the Institute of Biomolecular Chemistry (IBC) in Pozzuoli, while both the amorfrutins and the other polyphenols were tested in prof. Mariano Stornaiuolo's laboratory as GLUT promoters and as mitochondrial stimulators. This project was partly funded by the Ministry of Education, Universities and Research, PRIN 2022 PNRR: P2022PPH4C. The second chapter of Part 3 reports on the phytochemical analysis of Aglianico grape pomace and the anti-diabetic potential of oleanolic acid, its main component. This project was carried out in collaboration with Professor Martino Forino of the Department of Agronomy of the University Federico II. The NMR spectra of *G. foetida* are reported in Appendix D. The fourth part of this manuscript deals with the analytical analyses of tropical dietary plants with anti-diabetic potential: a less studied species of the genus *Piper* and *Annona* sp. plants. The first chapter presents the previously unknown composition of *P. borbonense*, which was investigated using a metabolomic approach and different analytical techniques, including Feature-Based Molecular Networking analysis. Furthermore, the analyses led to the isolation of two unprecedented compounds. This project was carried out in collaboration with Professor Giovanni Appendino at the Università degli Studi del Piemonte Orientale and partly at the École Polytechnique during my stay abroad. The second chapter deals entirely with my stay in Gif-sur-Yvette, near Paris, at the Centre National de la Recherche Scientifique (CNRS) under the supervision of prof. Véronique Eparvier and in collaboration with prof. David Touboul at the École Polytechnique. Here, thanks to the collaboration of 'Herbier IRD de Guyane', I studied different species of *Annona* genus from French Guyana and developed an LCMS-based chemotaxonomic classification methodology. Then I have focused on acetogenins, a class of specialised metabolites present in the Annonaceae family, developing and optimizing an analytical LC-MS/MS procedure for the detection of these secondary metabolites by performing post-column infusions of different salts to induce a specific fragmentation pattern. I also had the opportunity to carry out ion mobility analyses. Appendices E and F contain the supplementary data for *P. borbonense* and the MS data for acetogenins, respectively.

PUBLICATIONS

- 1) Al-Hmadi HB, Serino E, Pastore A, Chianese G, Hammami S, Stornaiuolo M, Taglialatela-Scafati O. Metabolites from Aerial Parts of Glycyrrhiza foetida as Modulators of Targets Related to Metabolic Syndrome. *Biomolecules*. 2024 Apr 11;14(4):467.
[DOI: 10.3390/biom14040467](https://doi.org/10.3390/biom14040467)
- 2) Serino E, Rigano D, Bruno M, Pastore A, Stornaiuolo M, Formisano C, Taglialatela-Scafati O. Glucose Uptake Stimulating Metabolites from Aerial Parts of *Centaurea sicula*. *Journal of Natural Products*. 2024 Mar 26;87(4):1179-86.
[DOI: 10.1021/acs.jnatprod.4c00134](https://doi.org/10.1021/acs.jnatprod.4c00134)
- 3) Serino E, Iannotti FA, Al-Hmadi HB, Caprioglio D, Moriello C, Masi F, Hammami S, Appendino G, Vitale RM, Taglialatela-Scafati O. PPAR α / γ -Targeting Amorfrutin Phytocannabinoids from Aerial Parts of *Glycyrrhiza foetida*. *Journal of Natural Products*. 2023 Nov 8;86(11):2435-47.
[DOI: 10.1021/acs.jnatprod.3c00509](https://doi.org/10.1021/acs.jnatprod.3c00509)
- 4) Errichiello F, D'Amato M, Gambuti A, Moio L, Pastore A, Hekmat AH, Stornaiuolo M, Serino E, Taglialatela-Scafati O, Forino M. Oleanolic acid: A promising antidiabetic metabolite detected in Aglianico grape pomace. *Journal of Functional Foods*. 2023 May 1;104:105548.
[DOI: 10.1016/j.jff.2023.105548](https://doi.org/10.1016/j.jff.2023.105548)
- 5) Serino E, Chianese G, Musto G, Zengin G, Rigano D, Stornaiuolo M, Formisano C, Taglialatela-Scafati O. Guaiane-rich phytochemical profile of *Centaurea kotschyi* subsp. *persica* (Boiss.) Wagenitz and identification of hypoglycaemic metabolites. *Phytochemistry*. 2022 Jul 1; 199:113189.
[DOI: 10.1016/j.phytochem.2022.113189](https://doi.org/10.1016/j.phytochem.2022.113189)
- 6) Serino E, Chahardoli A, Badolati N, Sirignano C, Jalilian F, Mojarrab M, Farhangi Z, Rigano D, Stornaiuolo M, Shokoohinia Y, Taglialatela-Scafati O. Salvigenin, a trimethoxylated flavone from *Achillea wilhelmsii* C. Koch, exerts combined lipid-lowering and mitochondrial stimulatory effects. *Antioxidants*. 2021 Jun 29;10(7):1042.
[DOI: 10.3390/antiox10071042](https://doi.org/10.3390/antiox10071042)

In addition, three other papers are about to be submitted for publication.

TABLES OF CONTENTS

ABSTRACT.....	iii
PUBLICATIONS.....	v
PART 1	1
CHAPTER 1: INTRODUCTION	1
1.1 Metabolic syndrome	1
1.2 The role of natural products in Metabolic Syndrome	2
1.3 Pharmacological targets studied in this project	5
1.3.1 GLUT transporters	6
1.3.2 Peroxisome-proliferator-activated receptors	8
1.3.3 Mitochondrial activity	11
1.3.4 Alpha-amylase and alpha-glucosidase	12
References	14
CHAPTER 2: Techniques for deconvolution and structural elucidation applied in this study .	17
2.1 Mass Spectrometry	17
2.2 Molecular Networking	21
2.3 Trapped Ion Mobility Spectrometry (TIMS)	23
2.4 Nuclear Magnetic Resonance	24
2.5 Stereochemical assignment	25
2.5.1 Determination of the relative configuration	25
2.5.2 Determination of the absolute configuration	26
2.6 General experimental procedures	28
References	29
PART 2: <i>Centaurea</i> species secondary metabolites and their activity on metabolic syndrome	31
CHAPTER 1: Genus <i>Centaurea</i> (Asteraceae)	32
1.1 Asteraceae family	32
1.2 Genus <i>Centaurea</i>	33
1.2.1 Secondary metabolites of genus <i>Centaurea</i> : sesquiterpenes	34
1.2.2 Bioactivity of genus <i>Centaurea</i> and of sesquiterpenes	36
References	37
CHAPTER 2: <i>Centaurea kotschyi</i>	41
2.1 Preliminary LC-MS/MS analysis.....	41
2.2 Phytochemical composition of the aerial parts	42

2.3 Biological activity	46
2.4 Conclusions	48
2.5 Experimental section	48
2.5.1 Plant material	48
2.5.2 Extraction and isolation	48
2.5.3 Spectral Data.....	49
2.5.4 Acetylation of kotschyol A and cynarinin A	51
2.5.5 LC-MS/MS analysis	51
2.5.6 2-NBDG glucose assays	51
References	53
APPENDIX A	55
CHAPTER 3: <i>Centaurea sicula</i>	64
3.1 Phytochemical characterization of the aerial parts	64
3.2 Bioactivity of the secondary metabolites	72
3.3 Conclusions	73
3.4 Experimental section	74
3.4.1 Plant material	74
3.4.2 Extraction and isolation	74
3.4.3 Spectral Data.....	75
3.4.4 Methanolysis of Melitensin	77
3.4.5 NBDG Glucose Uptake Assay.....	77
References	79
APPENDIX B	83
CHAPTER 4: Alpha-glucosidase inhibitory profile of sesquiterpene lactones from <i>Centaurea</i> species.....	102
4.1 Alpha-glucosidase activity of sesquiterpene lactones	102
4.2 Molecular docking experiments	103
4.3 Enzyme kinetics.....	104
4.4 Conclusions	106
4.5 Experimental Section	107
4.5.1 Plant Material and isolation of compounds	107
4.5.2 α -Glucosidase Inhibition Assay.....	107
4.5.3 Inhibitory kinetic analysis	107
4.5.4 Molecular modeling	108

References	109
CHAPTER 5: NMR-based profiling of <i>Centaurea patula</i> sesquiterpenoids and their alpha-glucosidase inhibitory activity	111
5.1 Phytochemical composition of <i>C. patula</i>	112
4.2 Alpha-glucosidase inhibitory activity	120
5.3 Conclusions	122
5.4 Experimental section	122
5.4.1 Plant Material	122
5.4.2 Extraction and Isolation.....	122
5.4.3 Spectral Data.....	124
5.4.4 Computational Details of compounds 16 and 17.....	127
5.4.6 α -Glucosidase Inhibition Assay.	127
References	129
APPENDIX C.....	132
PART 3: Dietary molecules from food plants active on Metabolic Syndrome.....	152
CHAPTER 1: <i>Glycyrrhiza foetida</i> secondary metabolites and their modulation on MetS	153
1.1 Genus Glycyrrhiza.....	153
1.2 Amorfrutins	155
1.2.1 Amorfrutins biogenesis	156
1.2.2 Amorfrutins bioactivity	157
1.3 <i>Glycyrrhiza foetida</i> phytochemical composition.....	158
1.3.1 Amorfrutin profiling.....	159
1.3.2 Polyphenols profiling.....	166
1.4 Bioactivity of <i>G. foetida</i> secondary metabolites.....	169
1.4.1 Amorfrutins modulation on PPARs target	169
1.4.2 In Silico Study and Structure-Activity Relationships on PPARs	171
1.4.3 Effect of mitochondrial activity.....	180
1.4.4 Promotion of glucose uptake via glucose transporters (GLUTs).....	182
1.5 Conclusions	183
1.6 Experimental Section	185
1.6.1 Plant material and extraction	185
1.6.2 LC-MS/MS Analysis	185
1.6.3 Purification of secondary metabolites from <i>G. foetida</i>	186
1.6.4 Spectral Data.....	187

1.6.5 Computational Methods.....	191
1.6.6 Cell Culture, Transfection, and Luciferase Assay	192
1.6.7 MTT Assay	192
1.6.8 Mitochondrial Activity.....	193
1.6.9 NBDG Glucose Uptake Assay.....	193
1.6.10 RNA Purification and mRNA Quantification by Real-Time RT-PCR.....	194
1.6.11 GC/MS Analysis	194
References	196
APPENDIX D.....	201
CHAPTER 2: Oleanolic acid, a promising antidiabetic metabolite detected in Aglianico grape pomace	229
2.1 Potential bioactive secondary metabolites of Aglianico	229
2.2 Chemical analysis of grape pomace	232
2.2.1 HPLC-DAD analyses of amino acids	236
2.4 Effect of oleanolic acid enriched fractions on mitochondrial activity of in-vitro cultured C2C12 myoblast	238
2.5 Effect of oleanolic acid enriched fractions on glucose uptake activity of in-vitro cultured C2C12 myoblast	240
2.6 Conclusions	241
2.7 Experimental Section	242
2.7.1 Chemicals and reagents.....	242
2.7.2 Grape pomace collection, extraction and chromatographic separation.....	242
2.7.3 LC-MS/MS analyses.....	243
2.7.4 NMR analysis.....	244
2.7.5 High-Performance liquid chromatography analyses of anthocyanins	245
2.7.6 High-Performance liquid chromatography analyses of amino acids	245
2.7.7 Measurement of mitochondrial activity.....	246
2.7.8 Measurement of glucose uptake	247
2.7.9 Statistical analysis	247
References	248
PART 4: Chemical profiling of some antidiabetic dietary plants	253
CHAPTER 1: Untargeted phytochemical analysis of <i>Piper borbonense</i>	254
1.1 <i>Piper</i> species and their antidiabetic potential	254
1.2 <i>Piper borbonense</i>	255

1.3 Metabolomic profiling of <i>Piper borbonense</i> using LC-MS/MS-FBMN and LC-MSMS guided purification	256
1.4 Isolation and characterization of secondary metabolites	261
1.5 Quantification of sesamin	267
1.6 Analysis of the volatile compounds of <i>P. borbonense</i> fruits	268
1.7 Conclusions	269
1.8 Experimental section	270
1.8.1 Plant Material and Extraction	270
1.8.2 LC-MS/MS Analysis and Molecular Networking	270
1.8.4 Isolation of pure compounds	271
1.8.5 Spectral data	273
1.8.6 Computational Details	273
1.8.7 Head-space gas-chromatography	274
References	275
APPENDIX E	281
CHAPTER 2: Chemotaxonomy of <i>Annona</i> species and analysis of acetogenins	288
2.1 <i>Annona</i> genus.....	288
2.1.1 The antidiabetic potential of <i>Annona</i> sp. plants.....	290
2.1.2 Annonaceous acetogenins	291
2.2 Chemotaxonomy of French Guyana <i>Annona</i> species.....	292
2.3 Cerium and europium induced fragmentation of acetogenins	294
2.3.1 Fragmentation pattern analysis of annonacin with different salts	294
2.3.2 Analyses of complex mixtures.....	298
2.3.3 Ion Mobility analyses of acetogenins.	302
2.4 Experimental Section	304
2.4.1 Chemicals.....	304
2.4.2 Plant material and sample preparation.	304
2.4.3 LC-MSMS and ion mobility analyses	304
2.4.4 PLSDA and Hierarchical Cluster Dendrogram generation.	305
References	307
APPENDIX F	311
General Conclusion.....	319

PART 1

CHAPTER 1: INTRODUCTION

1.1 Metabolic syndrome

The metabolic syndrome (MetS) is a complex cluster of disorders that mainly affect lipid and carbohydrate metabolism. It is a multifactorial condition in which a high-sugar, high-fat diet combined with physical inactivity are the main triggering factors, although there may be a genetic predisposition. The difficulties in understanding its exact pathogenesis contribute to the complexity of the diagnosis and the establishment of universal criteria to describe this pathology. In addition, the criteria vary according to ethnicity and gender (van der Made et al., 2012). Nevertheless, the generally accepted parameters are related to the following five main risk factors (Ambroselli et al., 2023):

1. Visceral obesity: the increase in visceral fat in the abdominal cavity. It can be measured as a high waist circumference, which in Europe is set at 102 cm (or more) for men and 88 cm (or more) for women, but the threshold may vary depending on the country;
2. High blood pressure: when blood pressure is higher than 130/85 mmHg.
3. Low HDL cholesterol: when HDL levels are below 40 mg/dL in men and below 50 mg/dL in women.
4. High triglyceride levels: TG levels between 150 and 199 mg/dL are considered borderline, while levels above 200 mg/dL are considered high.
5. Insulin resistance: it's an impaired biological response of tissues to insulin stimulation. It is caused by cells that are less responsive to the action of insulin, resulting in increased levels of sugar in the bloodstream. A fasting blood glucose level between 100 and 125 mg/dL indicates pre-diabetes.

In addition to these major risk factors, also oxidative stress, along with pro-inflammatory conditions and the production of abnormal adipocytokines, such as tumor necrosis factor α , interleukin-1 (IL-1), IL-6, leptin, and adiponectin, play a crucial role in the pathological

development (Srikanthan et al., 2016). MetS is spreading at an alarming rate in developed Western countries due to dietary habits and sedentary lifestyles, and is reaching epidemic proportions in lower-middle-income countries, where the mortality rate due to diabetes increased 13%, with dramatic implications for public health and the global economy. (Shin et al., 2013; Silveira Rossi et al., 2021). The first line of intervention is to change the dietary and lifestyle habits of those affected in order to reduce the risk factors that can lead to the development of pathological disorders. A healthy diet, increased physical activity, weight loss, stopping smoking and limiting alcohol consumption are the first steps in preventing the condition from worsening. If not adequately controlled, MetS can develop into more serious conditions such as obesity, hypercholesterolemia and hypertriglyceridemia, cardiovascular disease (CVD) and type 2 diabetes mellitus. These conditions can often co-exist as dyslipidemia, where hypercholesterolemia, hypertriglyceridemia and low HDL levels co-occur, or as diabetic dyslipidemia, where diabetes is associated with dyslipidemia. The most serious clinical outcomes of MetS are CVD and type 2 diabetes mellitus, which are two of the greatest health threats worldwide. In fact, CVD complications, such as stroke, are more likely to occur in the MetS affected population, while type 2 diabetes mellitus is estimated to cause over one million deaths per year and is expected to increase in the near future (Chandra et al; 2022). According to the World Health Organization (WHO), in 2019 diabetes was the direct cause of 1.5 million deaths and 48% of all deaths due to diabetes occurred before the age of 70 years, another 460 000 kidney disease deaths were caused by diabetes. The expected incidence rate is shown **Figure 1.1**.

1.2 The role of natural products in Metabolic Syndrome

The synthetic drugs developed for the treatment of MetS are generally expensive and associated with several side effects. Moreover, their use is often discouraged in the first stages of the pathological development. On the other hand, natural products have many qualities that make their application highly attractive and that are driving the current interest in investigating their modulation of MetS dysfunctional pathways. First, natural products are usually less expensive than pharmaceuticals, making them more accessible to the affected population. Furthermore, they have fewer side effects than synthetic drugs, giving patients a sense of security that improves the adherence to the therapy. As they are often mixtures of bioactive

ingredients, they have a multi-target mechanism of action that has proven to be highly effective (Bilal et al., 2018).

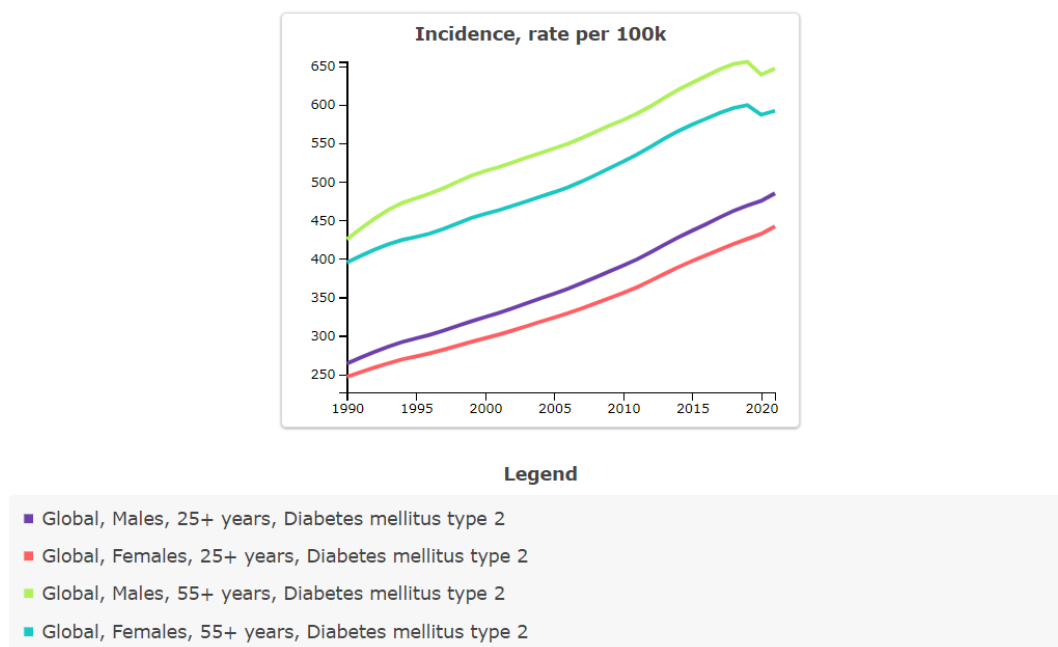


Figure 1.1 Incidence of diabetes from Global Burden of Disease Collaborative Network. Global Burden of Disease Study 2019. Results. Institute for Health Metrics and Evaluation. 2020

In addition, a growing number of scientific studies have demonstrated their potential not only as alternative medicine, but also in combination with synthetic drugs, supporting pharmacological treatments, offering the possibility of reducing the drug dose and consequently reducing side effects. All these qualities have increased their use and improved patient compliance (Jacob et al., 2019). A large number of natural products is known to improve MetS through different mechanisms of action. The most widely used are botanicals, which are plant-based products with traditional uses and claims in health, medicine and human nutrition. Botanicals contain hundreds or thousands of ingredients that can have synergistic effects and are easily marketed as dietary supplements. For example, many nutraceuticals have a cholesterol-lowering effect that can work on three different levels:

1. Reduce intestinal absorption of dietary cholesterol, such as sterols and plant fibres. Phytosterols (β -sitosterol, campesterol and stigmasterol) and stanols, their saturated derivatives (sitostanol and campestanol) (**Figure 1.2**), can displace the

incorporation of cholesterol into mixed micelles in the small intestine, thereby reducing cholesterol absorption and consequently LDL plasma levels (Plat et al., 2000). They can also exert another mechanism by interfering with proteins involved in cholesterol metabolism in both enterocytes and hepatocytes, such as ABCA1 or ABCG5/8 (Calpe-Berdiel et al., 2009). On the other hand, dietary fibres, such as β -glucan, pectin and psyllium, can prolong gastric emptying time, increase satiety and promote faecal excretion of cholesterol and bile salts. These products are commonly used to promote weight loss. Synthetic drugs that have a similar effect may act on Niemann-Pick C1-Like 1 (NPC1L1), a sterol carrier on enterocytes, which is the pharmacological target of ezetimibe, used to treat hypercholesterolaemia and reduce the risk of cardiovascular disease.

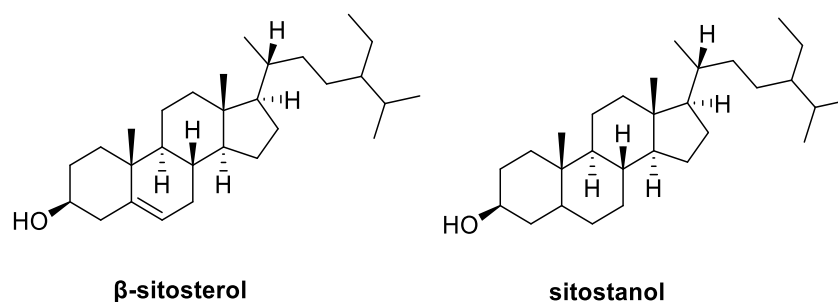


Figure 1.2 Sterols and stanols chemical structures

2. Promoting the elimination of LDL cholesterol, as berberine (**Figure 1.2**). This isoquinoline alkaloid, isolated from the genus *Berberis*, can improve the lipid profile by acting on several targets. It promotes fatty acid oxidation, hepatic LDL clearance and can also interfere with dietary cholesterol absorption. It can also stimulate glycolysis, insulin secretion and inhibit gluconeogenesis and adipogenesis in the liver. However, berberine has poor bioavailability, which makes its therapeutic use difficult (Tabeshpour et al., 2017; Och et al., 2022).
3. Reducing hepatic cholesterol synthesis, such as fermented red rice. This product can inhibit 3-hydroxy-3-methylglutaryl coenzyme A (HMG-CoA) reductase, a hepatic enzyme that plays a key role in the biosynthesis of cholesterol. In particular, monacolin K (**Figure 1.3**), the main bioactive component of fermented red rice

(*Oryza sativa*), was the precursor of a class of synthetic drugs called statins (Patel, 2016).

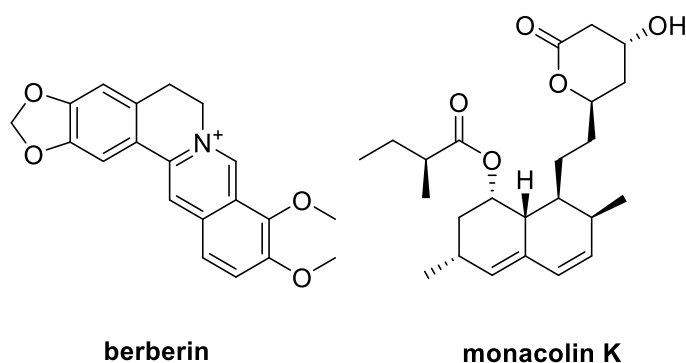


Figure 1.3 Chemical structures of berberin and monacolin K.

Other natural products with health benefits in MetS include omega-3 fatty acids, particularly eicosapentaenoic acid (EPA) and docosahexaenoic acid (DHA), whose role is not fully understood but appears to be related to an improvement in insulin sensitivity, a reduction in VLDL and triglycerides, and anti-inflammatory properties (van der Made et al., 2012). In addition, curcuma extracts (from *Curcuma longa*) have strong antioxidant and anti-inflammatory activity, reducing the basal level of inflammation typical of obese patients, together with an improvement in the leptin-adiponectin ratio, an indicator of insulin resistance and atherogenic risk (Vafaeipour et al., 2022). Finally, among novel foods, spirulina, a type of blue-green algae, has also been shown to have several benefits against MetS, including antioxidant and anti-inflammatory properties, along with an improvement in glycaemic and lipid profiles (Yousefi et al., 2019).

Considering the major impact of MetS on the world population and the great resource of natural substances, my PhD project aimed at the phytochemical characterisation of certain plants with anti-diabetic properties, but whose bioactive compounds had not yet been described.

1.3 Pharmacological targets studied in this project

The secondary metabolites isolated during my PhD project have been tested against different targets depending on the data reported in the literature. Here is a description of the main targets

considered. Sesquiterpene lactones and lignans isolated from *Centaurea* species (described in the second part of this manuscript) have been tested on GLUT transporters and α -glucosidase in order to identify the bioactive compounds responsible for the antidiabetic potential of the organic extract of *Centaurea* species (Fattaheian-Dehkordi et al., 2021). PPAR γ / α modulatory effect was evaluated for amorfrutins isolated from *Glycyrrhiza foetida* (as described in the third part of this manuscript) given the high potential of these compounds as PPAR agonists (Weidner et al., 2012). In addition, amorfrutins and the other polyphenols isolated from *G. foetida* were tested on GLUT carriers and as mitochondrial stimulators.

1.3.1 GLUT transporters

Glucose is a vital element for cells as it is the primary source of energy and its cellular uptake is a fundamental process for all metabolic pathways. Its transport across the membrane is the primary event and is enabled by GLUT transporters, which are multi-pass transmembrane proteins that allow its passage according to the glucose concentration gradient. When glucose binds to its specific site on the transporters, these proteins undergo a conformational change that allows glucose to be transported across the cell membrane, released into the cytoplasm and finally returned to its original conformation. GLUT transporters are characterised by tissue-specific expression patterns, different activation mechanisms and substrate affinities, as summarised in **Table 1.1**. The regulation of GLUT transporters plays a central role in the handling of glycemia, glucose homeostasis and metabolism, with important implications for pathologies such as diabetes and metabolic disorders characterised by insulin resistance and hyperglycemia. There are four major isoforms of GLUT transporters (Mueckler et al., 2013):

1. GLUT1 is ubiquitous, but is most abundantly expressed in the blood-brain barrier, erythrocytes and placenta. It enables the basal transport of glucose in cells.
2. GLUT2 is found in the liver, pancreas and small intestine. It is involved in the glucose sensing process as it has a low affinity for glucose compared to the other isoforms, allowing glucose uptake at high glycaemic levels, such as after meals. It is involved in the glucose sensing process in pancreatic β -cells as it allows these cells to detect changes in blood glucose levels. When glucose enters through GLUT2 transporters, at high glycaemic levels, the increase in intracellular glucose concentration triggers insulin secretion. GLUT2 can be also regulated by insulin and glucagon. For example, GLUT2

translocation to the cell membrane of hepatocytes is promoted by insulin and leads to the accumulation of glucose as glycogen, whereas glucagon stimulates GLUT2 to release glucose into the bloodstream.

3. GLUT3 is expressed in neurons and is therefore essential for energy production in the brain.
4. GLUT4 is mainly found in muscle cells and fat tissue. It is also known as the "insulin-responsive glucose transporter" because its activation is completely regulated by insulin. In fact, GLUT4 is normally present in cytoplasmic vesicles and translocates to the plasma membrane following insulin signalling. Impaired GLUT4 activity is associated with type II diabetes mellitus and obesity.

Table 1.1 GLUT transporters

GLUT 1	GLUT 2	GLUT 3	GLUT 4
Blood, blood brain barrier, placenta	Liver, β pancreatic cells, small intestine	Brain, Neurons	Skeletal muscle, adipose tissue, cardiac muscle
Insulin independent	Insulin independent	Insulin independent	Insulin dependent
Fasting	High K_m Low affinity	Low K_m High affinity	

Another class of transporters that allow glucose to pass through cell membranes are the sodium-glucose linked transporters (SGLTs). The two most studied isoforms are SGLT1 and SGLT2, both of which operate in a 1:1 stoichiometry, meaning that one molecule of glucose is transported along with one sodium ion. The mechanism is summarized in **Figure 1.4**.

1. SGLT1 is mainly expressed in the epithelial cells of the small intestine and in the proximal tubules of the kidneys. It is responsible for the absorption of glucose from the intestinal lumen into the bloodstream and the reabsorption of glucose from the renal tubular fluid. SGLT1 mainly recognises glucose and galactose.

2. SGLT2 is predominantly located in the proximal tubules of the kidney, where it plays a primary role in the reabsorption of glucose, which is the main substrate, from the renal tubular fluid back into the bloodstream.

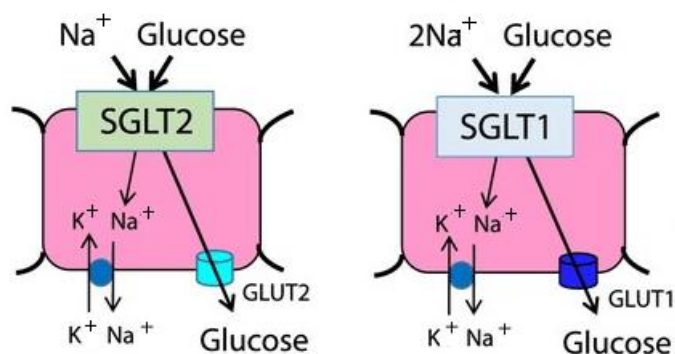


Figure 1.4 SGLT transporters

Drugs that are able to inhibit SGLT2 transporters are currently used to treat type 2 diabetes mellitus by increasing the renal clearance of glucose, which is not reabsorbed but excreted, lowering blood glucose levels (Deng et al., 2016). This class of synthetic drugs has been developed from phlorizin (**Figure 1.5**), a dihydrochalcone isolated from fruit trees (Ehrenkranz et al., 2005).

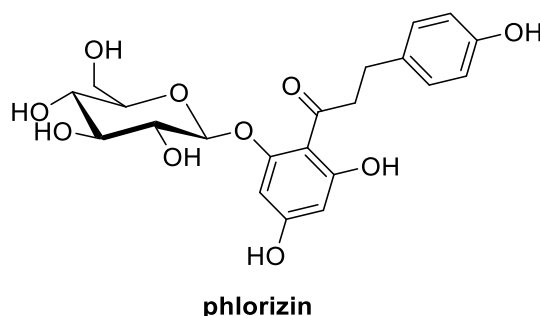


Figure 1.5 Chemical structure oh phlorizin

1.3.2 Peroxisome-proliferator-activated receptors

The peroxisome proliferator-activated receptors (PPARs) are a family of nuclear receptor proteins that play a key role in the modulation of several metabolic pathways, energy homeostasis, inflammation and cell differentiation. They can be involved in several pathological disorders such as obesity, type 2 diabetes mellitus, dyslipidaemia, and their modulation are often a turning point in the treatment of the pathologies in which they are implicated. There are three major PPAR isoforms: PPAR α , PPAR γ , which are the most studied,

and PPAR β/δ . Each has a specific distribution, functions and ligand specificity, but in general the endogenous ligands are unsaturated fatty acids. PPARs are ligand-activated transcription factors located in the cytoplasm of cells. When a specific agonist binds to the ligand binding domain (LBD), PPARs translocate to the nucleus where they heterodimerise with the retinoid X receptor (RXR), inducing interaction between the PPAR/RXR heterodimer and specific DNA sequences contained within regions that promote the expression of target genes, known as PPAR responsive elements (PPRE). This interaction triggers the recruitment of coactivator or corepressor proteins, leading to gene transcription, as show in **Figure 1.6** (Grygiel-Górniak et al., 2014).



Figure 1.6 Mechanism of gene transcription regulated by PPAR

PPAR β/δ , is ubiquitously distributed but expressed mainly in the gastrointestinal tract, kidney and skeletal muscle. It is involved in the regulation of cholesterol and lipid metabolism and its ligands could be potential drugs for the treatment of obesity (Luquet et al., 2005). However, this target has a controversial safety profile due to its activation in cancer progression (Wagner et al., 2020). PPAR α is mainly expressed in tissues with high metabolic activity and energy requirements, such as heart, skeletal muscle and liver. It plays a crucial role in lipid metabolism, particularly fatty acid β -oxidation. Synthetic PPAR α agonists are fibrates, a class of drugs used to treat hypertriglyceridemia. In addition to lowering triglycerides, they also improve HDL cholesterol levels, insulin sensitivity and the glycaemic profile (Bougarne et al., 2018). PPAR γ is predominantly expressed in white and brown adipose tissue, colon and spleen. It plays a crucial role in adipogenesis and is a potent insulin sensitiser. For these reasons, targeting PPAR γ is an attractive prospect for the management of type 2 diabetes mellitus and may offer some key advantages. In fact, the available synthetic agonists are thiazolidinediones (TZDs), which have a peculiar mechanism of action in that they improve insulin sensitivity and increase glucose uptake in peripheral tissues, thereby controlling glycaemia through a non-hypoglycaemic mechanism. This means that there is no direct stimulation of β -pancreatic cells

to produce insulin, which preserves β -pancreatic cells and slows down pathological development. On the other hand, TZDs have several severe side effects. They can increase the risk of myocardial infarction, which is already high in people with MetS, as well as fluid retention, which leads to edema in the lower limbs and can contribute to heart failure. In addition, these drugs cause weight gain, a critical issue for obese patients, most of whom have type 2 diabetes mellitus. Other problems include an increased risk of osteoporosis, bladder cancer and hypoglycaemia. Since TZDs are the only available drugs that selectively act on PPAR γ , it is a challenging goal to find new selective PPAR γ modulators that are not associated with such a critical side effect (Bermudez et al., 2010; Lindberg et al., 2007).

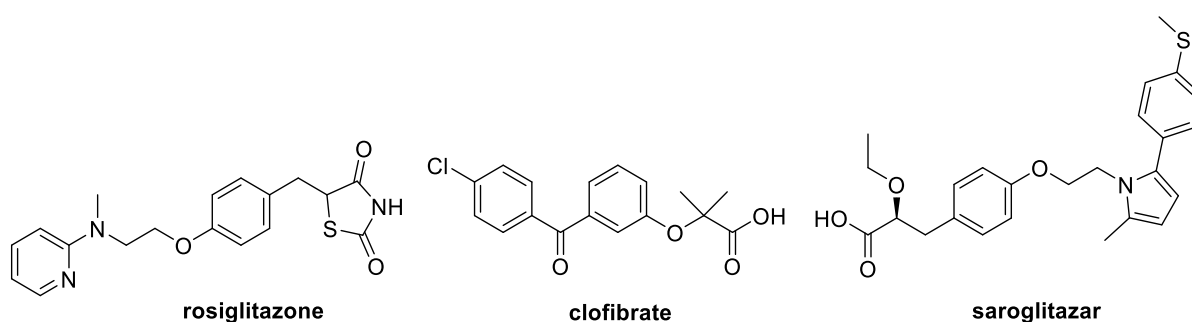
Table 1.2 Main PPAR isoforms, distribution and functions

PPAR ISOFORMS	TISSUE DISTRIBUTION	FUNCTIONS
PPAR γ	Brown and white adipose tissue, large intestine, spleen	Adipogenesis, lipid biosynthesis, glucose control
PPAR α	Liver, heart, brown fat, skeletal muscle	Fatty acids metabolism, lipid lowering levels

In summary, PPAR α promotes energy burning, whereas PPAR γ contributes to energy storage by promoting adipogenesis, as briefly reported in **Table 1.2**. As mentioned above, in many cases both glucose and lipid metabolism are dysfunctional, such as in diabetes dyslipidaemia. Recent studies have shown that the simultaneous activation of more than one PPAR isoform improves the therapeutic efficacy of pharmacological treatment. Dual PPAR γ/α agonists have a multi-target mechanism that is perfectly suited to this condition. They also appear to have beneficial effects on endothelial function, reduce blood pressure and have anti-inflammatory and anticoagulant properties. However, the drugs currently present on the market are still associated with the side effects of TZDs and therefore has a controversial safety profile (Chandra et al., 2022).

Table 1. 3 Synthetic agonists of PPAR, clinical applications and side effects

SYNTHETIC LIGANDS	CLINICAL USE	SIDE EFFECTS
Glitazones (PPAR γ agonist)	Type 2 diabetes mellitus	Hepatotoxicity, weight gain, fluid retention, cardiovascular complications
Fibrates (PPAR α agonist)	Hypertriglyceridemia	Hepatotoxicity, gallstones

**Figure 1.7** Chemical structures of a PPAR γ agonist (rosiglitazone), PPAR α agonist (clofibrate) and a dual agonist (saroglitazar)

1.3.3 Mitochondrial activity

Mitochondria are master regulators of cellular function. They are the powerhouses of cells, providing ATP through oxidative phosphorylation and are essential for many metabolic processes such as the citric acid cycle, fatty acid β -oxidation, calcium ion signalling and the biosynthesis of some metabolites. They also regulate apoptosis and are involved in developmental and ageing processes. Although the relationship between mitochondrial dysfunction and metabolic inflammation is still not clearly defined, many scientific publications have reported that dysfunctional mitochondrial respiratory capacity is associated with insulin resistance and the development of type 2 diabetes mellitus (Holmström et al., 2012). In fact, increasing mitochondrial activity, supporting mitochondrial biogenesis, reducing oxidative stress and improving mitochondrial health may have significant benefits in the treatment of a variety of diseases, including diabetes and obesity. Indeed, accelerated catabolism can promote lipolysis, glucose uptake and inhibition of lipid and cholesterol biosynthesis in liver cells and adipose tissue. Moreover, mitochondria regulates ROS balance, and an imbalanced ROS production promotes the progression of metabolic inflammation. On

the other hand, overnutrition can also affect mitochondria. In fact, the metabolic overload caused by overnutrition initially increases the number of mitochondria as a compensatory mechanism to meet the new cellular needs, leading to oxidative stress and chronic inflammation (Yuan et al., 2022). Among natural products, polyphenols have been shown to stimulate mitochondrial functionality, accelerating intracellular catabolic reactions with beneficial effects on metabolic syndrome (Serino et al., 2021).

1.3.4 Alpha-amylase and alpha-glucosidase

Alpha-amylase and alpha-glucosidase are two enzymes that modulate the digestion of complex sugars into monosaccharides, thus they can regulate their absorption rate. A schematic representation of α -amylase and α -glucosidase is reported in **Figure 1. 8**. Alpha-amylase is a hydrolytic enzyme with a key role in carbohydrates digestion. This enzyme recognizes and breaks down α -(1,4)-glycosidic linkages present in starch and glycogen, its substrates, turning complex polysaccharides in smaller ones. The process takes place in the mouth and in the small intestine, where the salivary and pancreatic amylases are present respectively.

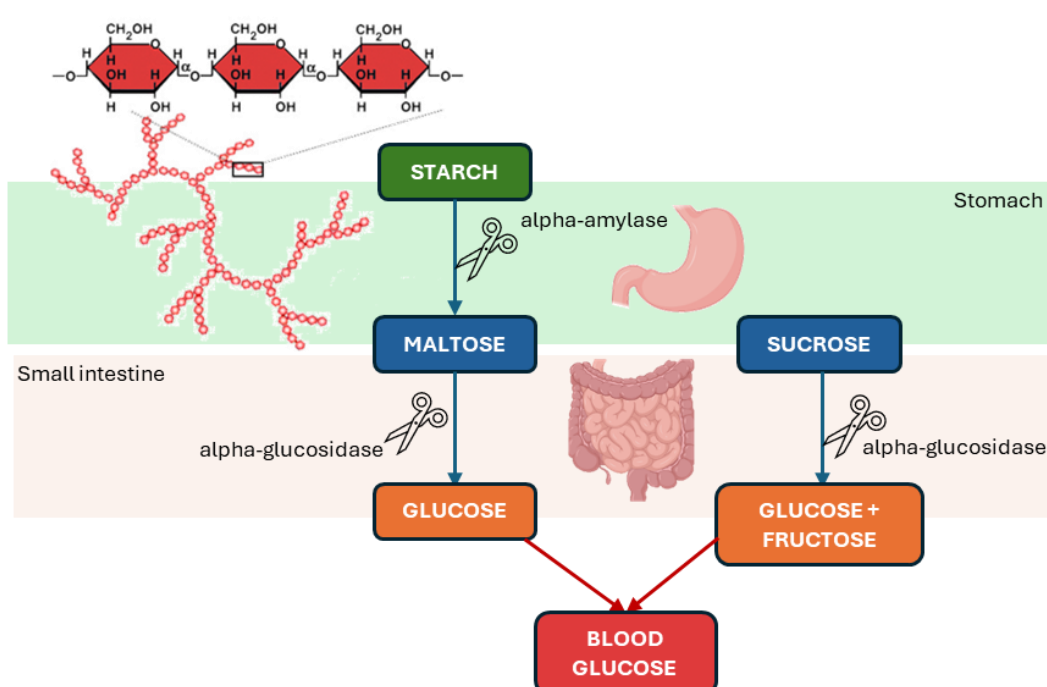


Figure 1.8 Hydrolysis of starch to glucose catalyzed by alpha-amylase and alpha-glucosidase

Alpha-glucosidase is an intestinal enzyme located in the brush border of the cells with a carbohydrate-hydrolase function. It acts on the O-glycosidic α -(1,4) linkage, breaking down oligosaccharides into single α -glucose monosaccharides. Its inhibition reduces the digestion rate of carbohydrates, lowering glucose levels in the blood. (Bashkin et al., 2021). Therefore, inhibitors of these enzymes represent a class of drugs used to control glycemia in non-insulin-dependent diabetes mellitus (type 2), particularly postprandial hyperglycemia, as shown in **Figure 1.9** (Bashkin et al., 2021; Fattaheian-Dehkordi et al., 2021).

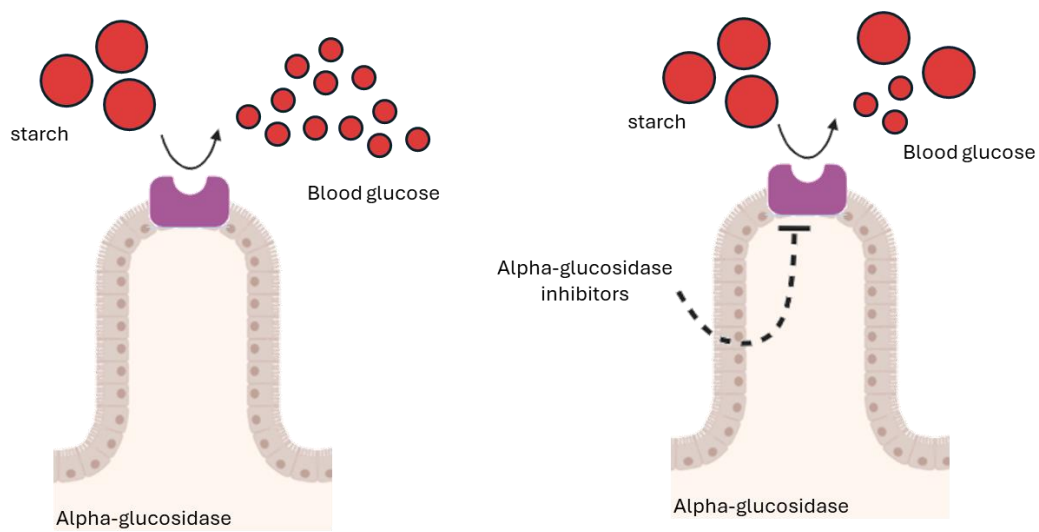


Figure 1.9 Schematic representation of alpha-glucosidase on an intestinal villus and its inhibition effect

References

- Ambroselli, D., Masciulli, F., Romano, E., Catanzaro, G., Besharat, Z.M., Massari, M.C., Ferretti, E., Migliaccio, S., Izzo, L., Ritieni, A. and Grosso, M., **2023**. New advances in metabolic syndrome, from prevention to treatment: the role of diet and food. *Nutrients*, *15*(3), p.640.
- Bashkin, A., Ghanim, M., Abu-Farich, B., Rayan, M., Miari, R., Srouji, S., Rayan, A. and Falah, M., **2021**. Forty-one plant extracts screened for dual antidiabetic and antioxidant functions: Evaluating the types of correlation between α -amylase inhibition and free radical scavenging. *Molecules*, *26*(2), p.317.
- Bermudez, V., Finol, F., Parra, N., Parra, M., Pérez, A., Penaranda, L., Vílchez, D., Rojas, J., Arráiz, N. and Velasco, M., **2010**. PPAR- γ agonists and their role in type 2 diabetes mellitus management. *American journal of therapeutics*, *17*(3), pp.274-283.
- Bilal, M., Iqbal, M.S., Shah, S.B., Rasheed, T. and Iqbal, H., **2018**. Diabetic complications and insight into antidiabetic potentialities of ethno-medicinal plants: A review. *Recent patents on inflammation & allergy drug discovery*, *12*(1), pp.7-23.
- Bougarne, N., Weyers, B., Desmet, S.J., Deckers, J., Ray, D.W., Staels, B. and De Bosscher, K., **2018**. Molecular actions of PPAR α in lipid metabolism and inflammation. *Endocrine reviews*, *39*(5), pp.760-802.
- Calpe-Berdiel, L., Escolà-Gil, J.C. and Blanco-Vaca, F., **2009**. New insights into the molecular actions of plant sterols and stanols in cholesterol metabolism. *Atherosclerosis*, *203*(1), pp.18-31.
- Chandra, A., Kaur, P., Sahu, S.K. and Mittal, A., **2022**. A new insight into the treatment of diabetes by means of pan PPAR agonists. *Chemical Biology & Drug Design*, *100*(6), pp.947-967.
- Deng, D. and Yan, N., **2016**. GLUT, SGLT, and SWEET: Structural and mechanistic investigations of the glucose transporters. *Protein Science*, *25*(3), pp.546-558.
- Ehrenkranz, J.R., Lewis, N.G., Ronald Kahn, C. and Roth, J., **2005**. Phlorizin: a review. *Diabetes/metabolism research and reviews*, *21*(1), pp.31-38.
- Fattaheian-Dehkordi, S., Hojjatifard, R., Saeedi, M. and Khanavi, M., **2021**. A review on antidiabetic activity of *Centaurea* spp.: A new approach for developing herbal

- remedies. *Evidence-Based Complementary and Alternative Medicine*, 2021(1), p.5587938.
- Grygiel-Górniak, B., **2014**. Peroxisome proliferator-activated receptors and their ligands: nutritional and clinical implications-a review. *Nutrition journal*, 13, pp.1-10.
- Holmström, M.H., Iglesias-Gutierrez, E., Zierath, J.R. and Garcia-Roves, P.M., **2012**. Tissue-specific control of mitochondrial respiration in obesity-related insulin resistance and diabetes. *American Journal of Physiology-Endocrinology and Metabolism*, 302(6), pp.E731-E739.
- Jacob, B. and Narendhirakannan, R.T., **2019**. Role of medicinal plants in the management of diabetes mellitus: a review. *3 Biotech*, 9, pp.1-17.
- Lindberg, M. and Astrup, A., **2007**. The role of glitazones in management of type 2 diabetes. A dream or a nightmare?. *Obesity Reviews*, 8(5), pp.381-384.
- Luquet, S., Gaudel, C., Holst, D., Lopez-Soriano, J., Jehl-Pietri, C., Fredenrich, A. and Grimaldi, P.A., **2005**. Roles of PPAR delta in lipid absorption and metabolism: a new target for the treatment of type 2 diabetes. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease*, 1740(2), pp.313-317.
- Mueckler, M. and Thorens, B., **2013**. The SLC2 (GLUT) family of membrane transporters. *Molecular aspects of medicine*, 34(2-3), pp.121-138.
- Och, A., Och, M., Nowak, R., Podgórska, D. and Podgórski, R., **2022**. Berberine, a herbal metabolite in the metabolic syndrome: The risk factors, course, and consequences of the disease. *Molecules*, 27(4), p.1351.
- Patel, S., **2016**. Functional food red yeast rice (RYR) for metabolic syndrome amelioration: a review on pros and cons. *World journal of microbiology and biotechnology*, 32, pp.1-12.
- Plat, J., Kerckhoffs, D.A. and Mensink, R.P., **2000**. Therapeutic potential of plant sterols and stanols. *Current opinion in Lipidology*, 11(6), pp.571-576.
- Serino, E., Chahardoli, A., Badolati, N., Sirignano, C., Jalilian, F., Mojarab, M., Farhangi, Z., Rigano, D., Stornaiuolo, M., Shokoohinia, Y. and Taglialatela-Scafati, O., **2021**. Salvigenin, a trimethoxylated flavone from *Achillea wilhelmsii* C. Koch, exerts combined lipid-lowering and mitochondrial stimulatory effects. *Antioxidants*, 10(7), p.1042.
- Shin, J.A., Lee, J.H., Lim, S.Y., Ha, H.S., Kwon, H.S., Park, Y.M., Lee, W.C., Kang, M.I., Yim, H.W., Yoon, K.H. and Son, H.Y., **2013**. Metabolic syndrome as a predictor of type 2 diabetes,

and its clinical interpretations and usefulness. *Journal of diabetes investigation*, 4(4), pp.334-343.

- Silveira Rossi, J.L., Barbalho, S.M., Reverete de Araujo, R., Bechara, M.D., Sloan, K.P. and Sloan, L.A., **2021**. Metabolic syndrome and cardiovascular diseases: Going beyond traditional risk factors. *Diabetes/metabolism research and reviews*, 38(3), pp.e3502-e3502.
- Srikanthan, K., Feyh, A., Visweshwar, H., Shapiro, J.I. and Sodhi, K., **2016**. Systematic review of metabolic syndrome biomarkers: a panel for early detection, management, and risk stratification in the West Virginian population. *International journal of medical sciences*, 13(1), p.25.
- Tabeshpour, J., Imenshahidi, M. and Hosseinzadeh, H., **2017**. A review of the effects of *Berberis vulgaris* and its major component, berberine, in metabolic syndrome. *Iranian journal of basic medical sciences*, 20(5), p.557.
- Vafaeipour, Z., Razavi, B.M. and Hosseinzadeh, H., **2022**. Effects of turmeric (*Curcuma longa*) and its constituent (curcumin) on the metabolic syndrome: An updated review. *Journal of integrative medicine*, 20(3), pp.193-203.
- van der Made, M. S., Baumgartner, S., González-Muniesa, P., Angeles Zulet, M. and Alfredo Martinez, J., **2012**. The use of functional foods in the metabolic syndrome. *Current Nutrition & Food Science*, 8(1), pp.25-44.
- Wagner, N. and Wagner, K.D., **2020**. PPAR beta/delta and the hallmarks of cancer. *Cells*, 9(5), p.1133.
- Weidner, C., de Groot, J.C., Prasad, A., Freiwald, A., Quedenau, C., Kliem, M., Witzke, A., Kodelja, V., Han, C.T., Giegold, S. and Baumann, M., **2012**. Amorphutins are potent antidiabetic dietary natural products. *Proceedings of the National Academy of Sciences*, 109(19), pp.7257-7262.
- Yousefi, R., Saidpour, A. and Mottaghi, A., **2019**. The effects of Spirulina supplementation on metabolic syndrome components, its liver manifestation and related inflammatory markers: A systematic review. *Complementary therapies in medicine*, 42, pp.137-144.
- Yuan, Q., Zeng, Z.L., Yang, S., Li, A., Zu, X. and Liu, J., **2022**. Mitochondrial stress in metabolic inflammation: modest benefits and full losses. *Oxidative Medicine and Cellular Longevity*, 2022(1), p.8803404.

CHAPTER 2: Techniques for deconvolution and structural elucidation applied in this study

Natural product research involves a number of different steps. After the extraction of the biomass, the extract can be subjected to a preliminary LC-MS/MS analysis to obtain a preliminary phytochemical fingerprint to guide the subsequent purification steps. Chromatographic methods (MPLC and HPLC) are then used to isolate the metabolites in order to obtain purified compounds for structural characterisation and pharmacological testing. During my PhD, I used spectrometric and spectroscopic techniques to identify the known compounds and to accurately determine the structural features of the newly discovered products. The relative and absolute configurations were also elucidated using NMR, chemical conversions or computational approaches.

2.1 Mass Spectrometry

Mass spectrometry (MS) is an analytical technique commonly used in organic chemistry to analyse ions in the gas phase by measuring their mass-to-charge ratio (m/z). It is a destructive but highly sensitive technique used to determine molecular weights, quantify known compounds, and determine the elemental formula and structural features of compounds present in a sample. A mass spectrometer consists of three main components: an ion source, a mass analyser and an ion detection system.

- 1) The ion source produces ions in the gas phase which are directed to the mass analyser by applying an electric-magnetic field. Different types of ion sources can be used to analyse samples with specific chemical-physical properties and analytical requirements.
- 2) The mass analyser separates ions based on their m/z , and plays a key role in determining the accuracy and resolution of the analysis. There are several types of analysers, each with its own operating principles and advantages.

- 3) Finally, the detector, usually working in conjunction with the analyser, converts the ion signals into measurable electronic signals which are then processed to produce the final mass spectrum.

The compounds described in the following chapters were mainly analysed using an electrospray ionisation (ESI) source through an Orbitrap system. ESI mass spectrometry is a widely used ionisation technique that is best suited to polar molecules, a schematic representation is shown in **Figure 2.1**. ESI allows the ionisation of compounds directly from a liquid solution and is particularly advantageous because it can be easily coupled to liquid chromatography (LC) systems, allowing the analysis of non-volatile samples. The sample solution emerges from a capillary outlet to which a high voltage is applied, resulting in the formation of an aerosol. The finely charged droplets undergo a further process of evaporation, which increases the surface charge and ultimately allows the ionised analytes to pass into the gas phase. A neutral gas, such as nitrogen, is used as a 'vehicle' to assist the atomisation of the liquid and the evaporation of the neutral solvent within the droplet. The ions produced can be protonated species $[M+H]^+$ or adducts with another cation, such as sodium $[M+Na]^+$. In addition, multiply charged ions, $[M+nH]^{n+}$, can also be observed. On the other hand, the expulsion of a hydrogen can produce an $[M-H]^-$ anion, which can be observed in the negative ion mode (Banerjee et al., 2012).

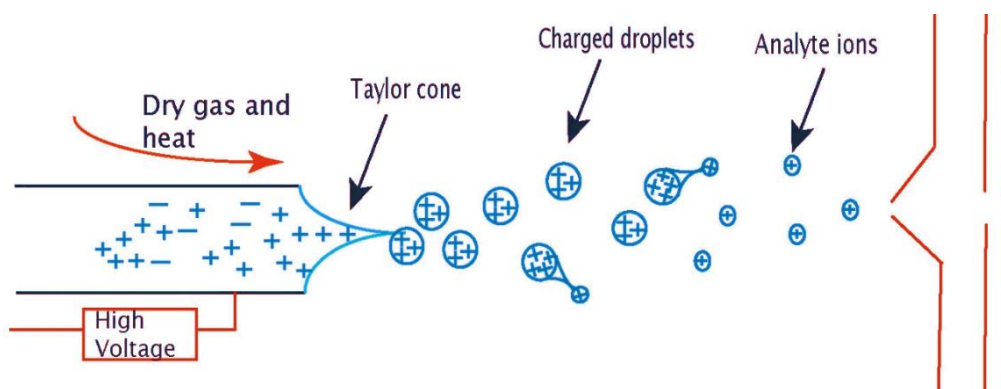


Figure 2. 1 Schematic representation of the electrospray ionization process

Orbitrap is a mass analyser based on the principle of trapping ions in an electrostatic field and measuring the frequency of their movement (**Figure 2.2**). The ions produced are trapped around an electrode by an electrostatic field and oscillate back and forth along the central axis in

response to an alternating current applied to the outer electrode. The frequency at which the ions oscillate depends on their mass-to-charge ratio. The oscillations of the ions induce a current in the electrodes, which is detected and converted by Fourier transformation into individual signals (Savaryn et al., 2016). The mass spectrum represents the abundance of ions at different m/z values. Orbitrap has several advantages:

- High resolving power (up to 200,000); high sensitivity, enabling the detection of low-abundance ions in complex mixtures; mass accuracy (1-2 ppm);
- Versatility, as it can be coupled with various ionisation techniques;
- High Resolution Mass Spectrometry (HR-MS), which combines Ion Trap and Fourier Transform technologies. HR-MS offers exceptional mass resolution and accuracy in determining the mass-to-charge ratio (m/z) of ions.

In addition, sequential fragmentation (MS^n spectra) can be performed, providing valuable information on the fragmentation patterns of ions and aiding in the identification of the molecules. In tandem mass spectrometry (MS/MS), the precursor ions selected on the basis of their m/z values are subjected to fragmentation and the resulting fragment ions are analysed to provide structural information. The high resolution and accuracy of the Orbitrap allows precise measurement of both precursor and fragment ions, increasing the reliability of MS/MS data (Alvarez-Rivera, 2019).

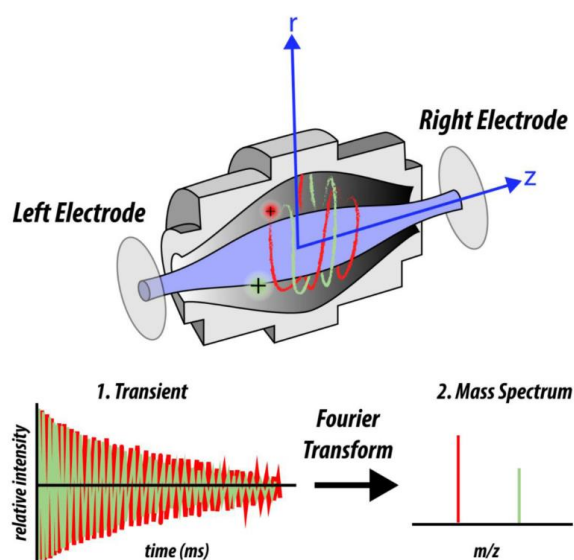


Figure 2. 2 The Orbitrap FT mass analyzer

Another commonly used mass spectrometer is the time-of-flight (ToF), which I had the opportunity to use during my visiting period at the École Polytechnique (Palaiseau, France). In this spectrometer, ions are separated based on the time it takes them to travel through a flight tube of known length to reach the detector. Ions with lower m/z will travel faster and reach the detector first, while ions with higher m/z will travel slower and reach the detector last (**Figure 2.3**). Thus, in the ToF spectrum, the recorded m/z peak corresponds to the sum of signals corresponding to multiple and independent ions arriving at the mass detector. The flight tube is often equipped with a reflectron, which corrects the poor resolution obtained with a linear ToF. The trajectory of the ions depends on their momentum and kinetic energy due to an applied pulsed acceleration voltage and the m/z ratios of the ions (Greaves et al., 2014).

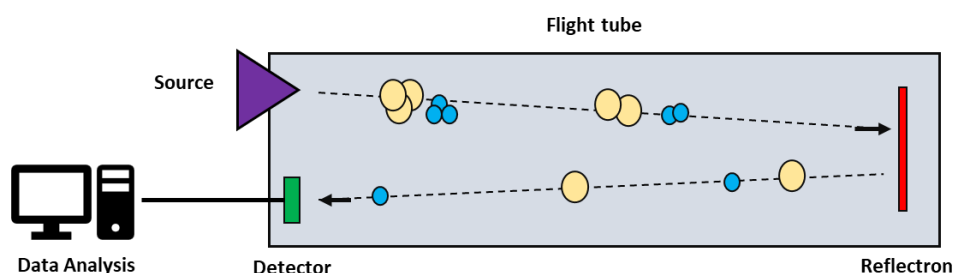


Figure 2. 3 Schematic representation of a ToF mass analyzer

Besides liquid-chromatography, also gas chromatography can be easily coupled with mass spectrometry. Gas chromatography mass-spectrometry (GC-MS) is an analytical technique widely used to identify and quantify small volatile lipophilic compounds in a sample (Sparkman et al., 2011). To be analyzed through this technique, the compounds must be thermostable due to the high temperatures involved. In this technique, the sample is vaporised and the components are separated by gas chromatography, then channeled into the mass spectrometer where they are ionised and fragmented. Typically, an electron impact (EI) ionisation source is used for this technique. In this source, a heated filament emits a beam of high-energy electrons that interact with the sample molecules, causing ionisation and the removal of one or more electrons, resulting in the formation of cations. The high-energy impact also causes fragmentation of the molecules, providing the structural information needed for the annotation, usually by comparing the experimental spectra with those reported in the libraries (NIST and Wiley Mass Spectral Libraries). A similar technique is Headspace

Chromatography (HS-GC), a particular type of gas chromatography that I had the opportunity to use during my visiting period at École polytechnique. This technique does not require any sample preparation. In fact, in this method, the sample is placed directly into a sealed vial and heated to allow volatile compounds to equilibrate between the sample and the gas phase (headspace) above it before injection (Wang et al., 2008).

2.2 Molecular Networking

Molecular Networking (MN) is an important computational dereplication tool that is valuable in metabolomics, natural product research and drug discovery. It is based on tandem mass spectrometry and allows faster identification of compounds in complex mixtures. For example, in an LC-MS/MS analysis, MN provides an overview of the acquired MS/MS data, organising it according to the similarities that occur between the spectra using a cosine spectral similarity algorithm (Quinn et al., 2017). Thus, MN allows a schematic visualisation of the structural relationships along the compounds present in a sample as a function of their MS/MS spectra. Indeed, if molecules share a similar scaffold, differently decorated or substituted, they will share a similar fragmentation pattern and thus generate similar MS/MS spectra. As MN algorithms allow each MS/MS spectrum within a data set to be compared with every other spectrum, if there is a similarity between the spectra, they will be clustered together to create a network of MS/MS spectral relationships from which the molecular network is originated. In general, after MS/MS data acquisition, the workflow consists of converting instrument-specific raw data files to an open format (mzXML, .mzML or .mgf), pre-processing the raw data using MZmine software (Schmid et al., 2023) and uploading the processed file to the Global Natural Products Social Molecular Networking platform (GNPS, <https://gnps.ucsd.edu/>) or on MetGem software (Olivon et al., 2018). Finally, Cystoscape open source software can be used to better visualise and finalise the network format (Shannon et al., 2003) (see **Figure 2.4**). In addition, metadata files can be added to enrich the MN with information (as in feature-based molecular networking) (Nothias et al., 2020). GNPS aligns each MS/MS spectrum in a dataset with each of the others, and then assigns a cosine score to each combination to describe their similarity. Depending on the degree of similarity, spectra are collapsed into a single node (consensus cluster) if they are from identical masses and have a high degree of similarity, or represented as

distinct nodes connected by edges if they are structurally related. A network is thus formed from nodes, which are the individual spectra, connected by edges whose thickness represents the degree of similarity between them. All consensus spectra (nodes) are then compared with spectral library databases (**Figure 2.5**). On the other hand, the MetGem is able to generate molecular networks using a different type of algorithm that results in a tSNE-based projection. In this case, the nodes are not connected by edges, but it's their spatial proximity that indicates the structural similarities between the nodes.

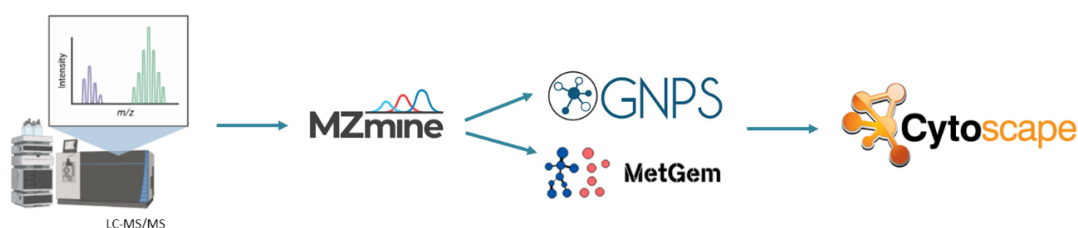


Figure 2. 4 General workflow to generate a Molecular Network

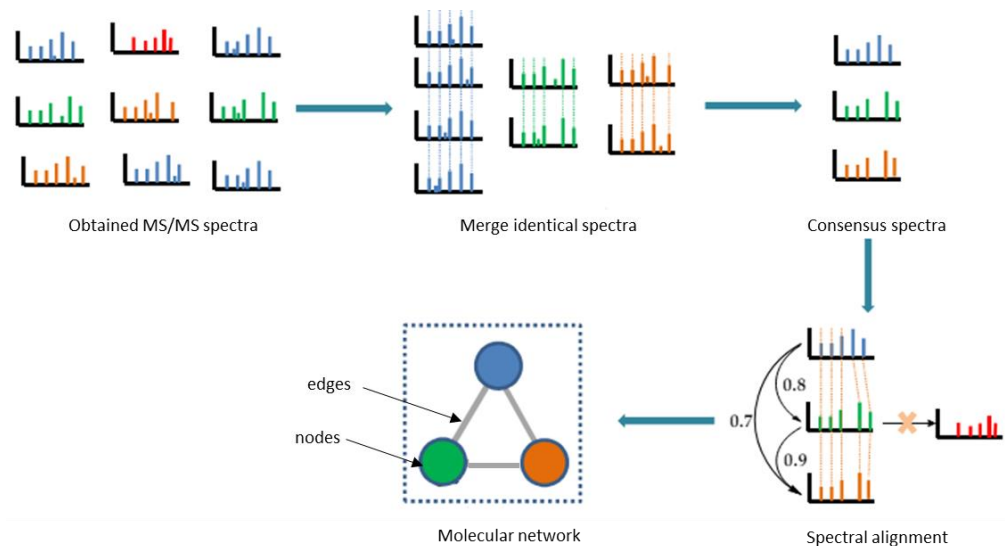


Figure 2. 5 Schematic representation of the steps to build a molecular network

2.3 Trapped Ion Mobility Spectrometry (TIMS)

During my stay abroad, I had the opportunity to acquire some basic knowledge of ion mobility measurements using a Bruker timsToF instrument.

Ion Mobility Mass Spectrometry (IMS-MS) is a sophisticated analytical technique that couples ion mobility to LC-MS, enhancing the analytical capabilities of mass spectrometry as it adds an extra dimension of separation to LC-MS, helping to separate and identify ions. In fact, mass spectrometry cannot provide information about the stereochemical properties of the compounds analysed, whereas ion mobility spectrometry can separate ions according to their three-dimensional properties, thus providing complementary information (Gabelica, 2021). Ion mobility refers to the ability of ions to move through a medium (a neutral gas) under the influence of an electric field. The ion mobility (K) is defined as the proportionality constant between the drift velocity (v_d) of an ion in a gas and the strength of the applied electric field (E) and is generally expressed in square metres per volt-second ($\text{m}^2/\text{V}\cdot\text{s}$).

$$K = v_d / E$$

Thus, ions are separated in an electric field in the presence of a collision gas depending on their m/z value and their spatial arrangement. In fact, for the same m/z , the mobility value will be lower if the ion has a more extended structure, corresponding to a lower drift velocity. Vice versa, the mobility value will be higher if the ion has a more compact structure, resulting in a higher drift velocity. IMS-MS can be particularly useful for the separation of isomeric compounds based on their unique mobilities in complex mixtures (Gabelica et al., 2019). The mobility value is usually given as the collision cross section (CCS), which represents the effective area of an ion interacting with neutral gas molecules and depends on the molecular formula, mobility value, collision gas and temperature. The CCS is a key parameter for the characterisation of molecular structures, as the structural interpretation can be based on the agreement between experimental and calculated CCS values (Gabelica et al., 2019). Trapped Ion Mobility Spectrometry (TIMS) is one of the most modern IMS methods recently launched by Bruker Daltonics and is highly efficient in terms of ion separation compared to other IMS (Dodds et al., 2019). The TIMS cell consists of three main sections: entrance funnel, analyser section and exit funnel (**Figure 2.6**) (Hernandez et al., 2014). Compared to other ion mobility drift tubes that use a stationary gas, in the TIMS drift tube ions are analysed using an electric field that holds the ions stationary against a moving gas, so the ions are initially trapped and

then eluted from the TIMS analyser over time according to their mobility (Michelmann et al., 2014).

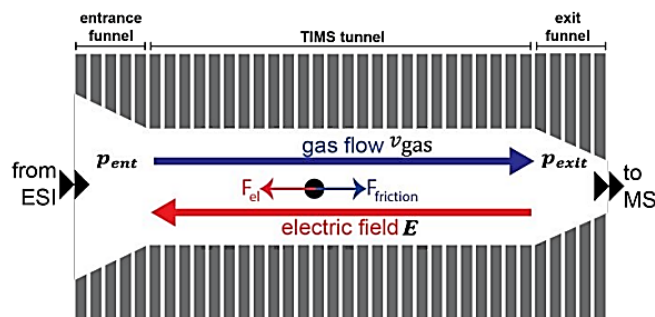


Figure 2. 6 Trapped Ion Mobility cell operation (Liu et al., 2022)

2.4 Nuclear Magnetic Resonance

Nuclear magnetic resonance (NMR) spectroscopy is a powerful analytical tool used to elucidate the structure of molecules in a non-destructive way. It exploits the magnetic properties of certain isotopic nuclei, which are characterised by a specific spin number (I). Spin is an intrinsic property that refers to the angular momentum possessed by atomic nuclei. When nuclei with $I \neq 0$ are placed in an external magnetic field (B_0), they adopt different orientations corresponding to specific energy levels. The number of different orientations is equal to $2I+1$. For example, hydrogen (^1H) and carbon (^{13}C) have $I=1/2$, which means they can be analysed using this technique, while the nuclei with $I=0$ are recognised as "NMR inactive". When an external magnetic field is applied to nuclei with $I = 1/2$, they can adopt two possible orientations, α parallel ($I = 1/2$), corresponding to the lower energy level, or β antiparallel ($I = -1/2$), corresponding to the higher level. The energy difference (ΔE) is equal to:

$$\Delta E = \gamma B_0$$

where γ is the gyromagnetic ratio specific to certain nuclei. The protons are distributed between the two levels (α and β), with an excess of protons in the α state, with its lower energy. When an energy is applied (a radiofrequency pulse), the nuclei undergo transitions from lower to higher energy levels; when the radiofrequency is turned off, it is possible to observe the Free Induction Decay (FID), which is transformed in the frequency domain by applying the Fourier Transform (FT). To carry out my research, one-dimensional (1D) and two-dimensional (2D) NMR

experiments have been extensively used for the structural elucidation of the metabolites (Soukup et al., 2014). In particular, the 2D NMR experiments used are listed below.

- COSY (COrrrelation SpectroscopY) is a homonuclear 2D NMR experiment used to identify the spin systems of a compound by the transfer polarisation between directly J-coupled spins. It provides information about the connectivity of protons within a molecule.
- HSQC (Heteronuclear Single Quantum Coherence) is a 2D heteronuclear NMR experiment used to correlate the proton signals with the chemical shift of the directly bonded carbon. The correlation peaks represent the one-bond proton-carbon couplings ($^1J_{CH}$).
- HMBC (Heteronuclear Multiple Bond Correlation) is a heteronuclear 2D NMR experiment that allows the correlation of proton signals with bound heteronuclei separated by multiple chemical bonds. It provides valuable information about the long-range connectivity between ^1H and neighbouring ^{13}C through $^{2,3}J_{CH}$ couplings or four bond correlations in the case of conjugated systems, while suppressing $^1J_{CH}$ coupling. This experiment allows the connection of different fragments of a molecule.

2.5 Stereochemical assignment

Natural compounds can be characterised by complex structures with one or more chiral centers. Once the planar structure of a newly described compound has been determined, its stereochemical features must also be elucidated. In fact, since the spatial arrangement of the various atoms in a molecule is a crucial aspect that influences its biological activities, the relative and absolute configuration of new metabolites must be thoroughly outlined.

2.5.1 Determination of the relative configuration

The relative configuration of a molecule refers to the spatial arrangement of two atoms or groups of atoms in relation to a reference molecule. It can be elucidated by detailed NMR analysis through the evaluation of chemical shifts (δ), coupling constants (J) and the NOE effect. For example, the chemical shift of a proton within a molecule is influenced by its chemical environment, so protons belonging to different diastereoisomers will have two

different chemical shifts. The influence of the relative proton arrangements in a molecule on the values of the coupling constants (J_{H-H}) can be described by the Karplus equation:

$$^3J = A\cos\theta^2 + B\cos\theta + C$$

This means that the coupling constant depends on the dihedral angle (θ) between the atoms considered, while the values A, B and C are empirically determined parameters that depend on the substituents involved. According to this equation, the coupling constant is small when the torsion angle is close to 90° , whereas it is large when the angles are 0° or 180° (Figure 2. 7). This provides a first indication of the relative spatial distribution of atoms within a molecule.

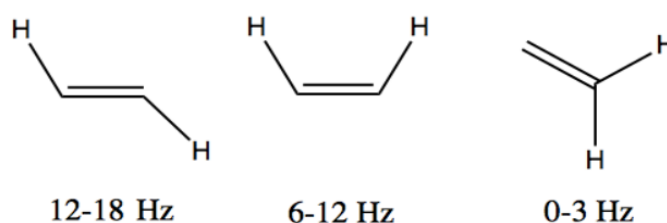


Figure 2. 7 Different J-J coupling constants between protons in geometric isomers.

On the other hand, a powerful tool for assigning the relative configuration is the NOESY 2D NMR experiment. NOESY (Nuclear Overhauser Enhancement Spectroscopy) is a homonuclear experiment that can show the spatial relationship of two protons in a molecule that do not belong to the same spin system. It is based on the NOE effect, a phenomenon in which the signal intensity is enhanced when two protons are spatially close to each other, affecting their nuclear dipole-dipole interactions. This effect can be observed for all protons within $2.5 \text{ \AA}$ of an irradiated proton (Claridge, 2016).

2.5.2 Determination of the absolute configuration

The absolute configuration of a molecule refers to the precise three-dimensional arrangement of substituents around each chiral center. It is often described using the R/S system for molecules with multiple chiral centers. ECD (electronic circular dichroism) is a spectroscopic technique used to determine the absolute configuration of chiral molecules by exploiting the interaction between an optically active molecule and plane polarised UV/Vis light (170-800 nm). Circular dichroism (CD) consists of the differential absorption of the two components of circularly polarised light by the chiral molecule. The two components, one rotating counter-

clockwise (L) and the other clockwise (R), simultaneously illuminate the sample to the same extent (Figure 2. 8).

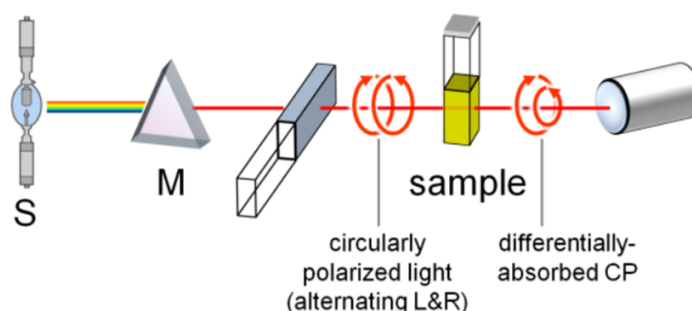


Figure 2. 8 Schematic representation of ECD

An asymmetric sample, containing one or more chiral centres, absorbs E_L and E_R in different ways, so the resulting electric field appears elliptically polarised (CD) and rotated (Optical Rotatory Dispersion) as a function of wavelength. On the other hand, a symmetric or racemic sample absorbs E_L and E_R equally, giving an $E_L + E_R$ vector that is plane polarised and not rotated. An electronic transition can be associated with a positive or negative ECD band, which is called the Cotton effect. The data are interpreted by comparing the experimentally obtained ECD spectra with a theoretical one (Figure 2. 9).

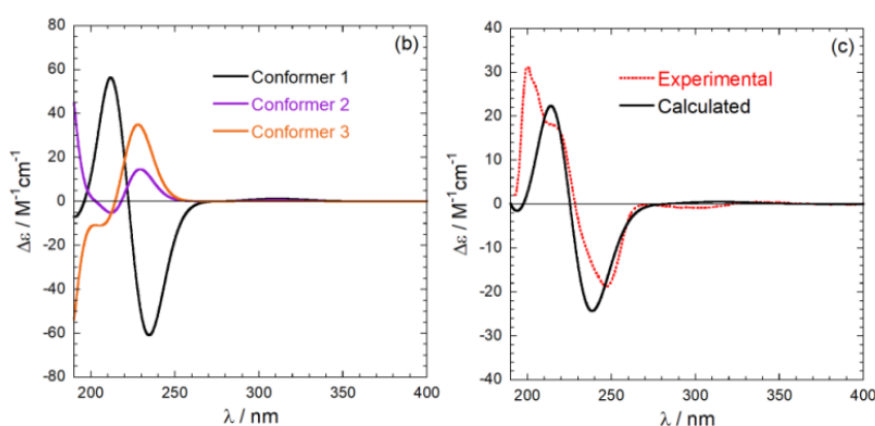


Figure 2. 9 Example of ECD spectra comparison

The CD spectrum depends directly on the electronic and geometric properties of a molecule, so a complex series of steps must be taken to obtain a calculated spectrum. The first

minimisation step is carried out using Molecular Mechanics (MM2, MMFF94 and Extended Hückel), and it is based on a preliminary conformational analysis of the molecule to evaluate the minimum energy of the possible conformers. The conformers characterised by the lowest energy ($\Delta\Delta G < 2$ kcal/mol) are then taken into account in the UV/ECD TDDFT (time-dependent density functional theory) calculations (Nugroho et al., 2014).

2.6 General experimental procedures

Thin-layer chromatography (TLC) was performed on plates coated with silica gel 60 F254 Merck, 0.25 mm. Chemicals and solvents were from Merck Life Science S.r.l. and were used without any further purification unless stated otherwise. Chromatographic purifications were carried out using MPLC (Medium Pressure Liquid Chromatography) and HPLC (High Pressure Liquid Chromatography). MPLC-DAD separations were performed on an Interchim® instrument, puriFlash XS 520 Plus (Sepachrom s.r.l., Milan, Italy). HPLC-UV-vis separations were performed on an Agilent instrument, using 1260 Quat Pump VL system, equipped with a 1260 VWD VL UV-vis detector and Knauer 1800 apparatus equipped with a refractive index detector and a Rheodyne injector. ^1H (400 and 600 MHz) and ^{13}C (100 and 150 MHz) NMR spectra were measured on a Bruker spectrometer. Chemical shifts are referenced to the residual solvent signal (CDCl_3 : δ_{H} 7.26, δ_{C} 77.0; CD_3OD : δ_{H} 3.31, δ_{C} 49.3). Homonuclear ^1H connectivities were determined by COSY (CORrelation SpectroscopY) experiments. Throughspace ^1H connectivities were evidenced using a NOESY (Nuclear Overhauser Enhancement SpectroscopY) experiment with a mixing time of 300 ms. One-bond heteronuclear ^1H – ^{13}C connectivities was determined by the HSQC (Heteronuclear Single Quantum Correlation) experiment; two- and three-bond ^1H – ^{13}C connectivities by gradient-HMBC (Heteronuclear Multiple Bond Correlation) experiments optimized for a ^2J of 8 Hz. HR-ESIMS experiments were performed on LTQ-Orbitrap mass spectrometer equipped with an ESI interface and Excalibur data system. Optical rotations (CHCl_3 and MeOH) were measured at 589 nm on a JASCO P2000 polarimeter. ECD spectra were obtained on a Jasco J-815 CD spectropolarimeter (Jasco, Mary's Court, Easton, MD, USA) with a 0.1 mm cuvette and four accumulations. LC-MS/MS analysis were performed on an LTQ-XL Ion Trap mass spectrometer (Thermo Fisher, Waltham, MA, USA) equipped with an Ultimate 3000 HPLC or on a LTQ Orbitrap XL FT-MS mass spectrometer (Thermo Fisher Scientific Spa) equipped with an Ultimate 3000 HPLC (Agilent Technology).

Chromatographic separation was achieved on a Kinetex 2.6 μm polar C18 100 Å (100 × 3 mm) column (Phenomenex).

References

- Alvarez-Rivera, G., Ballesteros-Vivas, D., Parada-Alfonso, F., Ibañez, E. and Cifuentes, A., **2019**. Recent applications of high resolution mass spectrometry for the characterization of plant natural products. *TrAC Trends in Analytical Chemistry*, 112, pp.87-101.
- Banerjee, S. and Mazumdar, S., **2012**. Electrospray ionization mass spectrometry: a technique to access the information beyond the molecular weight of the analyte. *International journal of analytical chemistry*, 2012(1), p.282574.
- Claridge, T.D., **2016**. High-resolution NMR techniques in organic chemistry (Vol. 27). Elsevier.
- Dodds, J.N. and Baker, E.S., **2019**. Ion mobility spectrometry: fundamental concepts, instrumentation, applications, and the road ahead. *Journal of the American Society for Mass Spectrometry*, 30(11), pp.2185-2195.
- Gabelica, V., Shvartsburg, A.A., Afonso, C., Barran, P., Benesch, J.L., Bleiholder, C., Bowers, M.T., Bilbao, A., Bush, M.F., Campbell, J.L. and Campuzano, I.D., **2019**. Recommendations for reporting ion mobility Mass Spectrometry measurements. *Mass spectrometry reviews*, 38(3), pp.291-320.
- Gabelica, V., **2021**. Ion mobility–mass spectrometry: an overview.
- Greaves, J. and Roboz, J., **2014**. *Mass spectrometry for the novice* (pp. 1-107). Boca Raton: CRC Press.
- Hernandez, D.R., DeBord, J.D., Ridgeway, M.E., Kaplan, D.A., Park, M.A. and Fernandez-Lima, F., **2014**. Ion dynamics in a trapped ion mobility spectrometer. *Analyst*, 139(8), pp.1913-1921.
- Liu, F.C., Ridgeway, M.E., Park, M.A. and Bleiholder, C., **2022**. Tandem-trapped ion mobility spectrometry/mass spectrometry (t TIMS/MS): a promising analytical method for investigating heterogenous samples. *Analyst*, 147(11), pp.2317-2337.
- Michelmann, K., Silveira, J.A., Ridgeway, M.E. and Park, M.A., **2014**. Fundamentals of trapped ion mobility spectrometry. *Journal of the American Society for Mass Spectrometry*, 26(1), pp.14-24.

- Nothias, L.F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., Protsyuk, I., Ernst, M., Tsugawa, H., Fleischauer, M. and Aicheler, F., **2020**. Feature-based molecular networking in the GNPS analysis environment. *Nature methods*, 17(9), pp.905-908.
- Nugroho, A.E. and Morita, H., **2014**. Circular dichroism calculation for natural products. *Journal of natural medicines*, 68, pp.1-10.
- Olivon, F., Elie, N., Grelier, G., Roussi, F., Litaudon, M. and Touboul, D., **2018**. MetGem software for the generation of molecular networks based on the t-SNE algorithm. *Analytical chemistry*, 90(23), pp.13900-13908.
- Quinn, R.A., Nothias, L.F., Vining, O., Meehan, M., Esquenazi, E. and Dorrestein, P.C., **2017**. Molecular networking as a drug discovery, drug metabolism, and precision medicine strategy. *Trends in pharmacological sciences*, 38(2), pp.143-154.
- Savaryn, J.P., Toby, T.K. and Kelleher, N.L., **2016**. A researcher's guide to mass spectrometry-based proteomics. *Proteomics*, 16(18), pp.2435-2443.
- Schmid, R., Heuckeroth, S., Korf, A., Smirnov, A., Myers, O., Dyrlund, T.S., Bushuiev, R., Murray, K.J., Hoffmann, N., Lu, M. and Sarvepalli, A., **2023**. Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nature biotechnology*, 41(4), pp.447-449.
- Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T., **2003**. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome research*, 13(11), pp.2498-2504.
- Soukup, R.W. and Soukup, K., **2014**. The series "Progress in the chemistry of organic natural products": 75 years of service in the development of natural product chemistry. *Progress in the Chemistry of Organic Natural Products* 100, pp.453-588.
- Sparkman, O.D., Penton, Z. and Kitson, F.G., **2011**. *Gas chromatography and mass spectrometry: a practical guide*. Academic press.
- Wang, Y., McCaffrey, J. and Norwood, D.L., **2008**. Recent advances in headspace gas chromatography. *Journal of Liquid Chromatography & Related Technologies*®, 31(11-12), pp.1823-1851.

PART 2: *Centaurea* species secondary metabolites and their activity on metabolic syndrome

This part of the manuscript describes the results obtained from the phytochemical characterisation of the apolar extracts of different *Centaurea* species aerial parts. These plants have been used for culinary purposes, for their edible flowers or as herbal teas, as well as for their medicinal properties, as described in the next chapter. Following my previous research activity on the phytochemical characterisation of the Iranian *Achillea wilhelmsii* (Asteraceae) and the activity of its secondary metabolites on MetS dysfunctional pathways (Serino et al., 2021), and in the frame of my research group on the investigation of less known species of the genus *Centaurea* (Formisano et al., 2017; De Cicco et al., 2021), I focused my attention on less studied species of this genus in order to identify their bioactive constituents. I have performed a comprehensive phytochemical profile of the organic extract of the aerial parts of *C. kotschyi* subsp. *persica*, *Centaurea sicula* and *Centaurea patula*. The extracts of *C. kotschyi* and *C. patula* were provided by our collaboration with Professor Gokhan Zengin of Selcuk University in Turkey, where *Centaurea* plants are most endemic. *C. sicula* extract was provided by Professor Maurizio Bruno of the University of Palermo in Sicily, where this species is native. After the purification and the spectroscopic identification, all the isolated pure metabolites, mainly sesquiterpene lactones, were tested in order to identify the bioactive compounds responsible for the hypoglycaemic activity of the extracts thanks to the collaboration of Professor Mariano Stornaiuolo of Department of Pharmacy at the University Federico II and Professor Sonia Piacente from Salerno University.

CHAPTER 1: Genus *Centaurea* (Asteraceae)

1.1 Asteraceae family

Asteraceae is one of the largest plant families, with more than 1600 genera and 25000 species. It includes a wide variety of plants, from small herbs to shrubs and even some trees, and is widespread throughout the world, especially in temperate and subtropical regions. These plants produce a wealth of secondary metabolites such as terpenes, flavonoids and alkaloids, which contribute to the ecological interactions of the plants as well as their medicinal properties. In fact, Asteraceae plants are economically important as they are used as herbal remedies for their beneficial effects on human health (Rolnik et al., 2021). Here are some common medicinal properties associated with Asteraceae plants:

- The antioxidant activity is closely related to the presence of flavonoids and phenolic compounds, which have the ability to scavenge free radicals and reduce oxidative stress. Aqueous, ethanolic and methanolic extracts of several Asteraceae plants have demonstrated this property in in vitro and in vivo experiments (Koc et al., 2015; Piątkowska et al., 2022).
- The anti-inflammatory properties have been described for a wide range of Asteraceae, including extracts of *Arnica montana* and *Calendula officinalis*. These plants are useful in the treatment of arthritis, rheumatism and skin irritations as they modulate the levels of pro-inflammatory mediators such as TNF- α , NF-kB, leading to the suppression of COX-2, along with various interleukins (Macedo et al., 2004; Preethi et al., 2009).
- These plants have traditionally been used to treat infections, wounds and respiratory problems. More recently, antimicrobial activity against bacteria, fungi and viruses has been evaluated (Gou et al., 2023; Lawson et al., 2020).
- Many plants in this family, including *Cynara scolymus* (artichoke), have been traditionally used to support digestive health and relieve gastrointestinal discomfort, as well as having antispasmodic, anti-ulcer and anti-diabetic properties (Jalili et al., 2020).
- The anti-diabetic effect demonstrated by *Centaurea* and *Achillea* species, whose extracts were able to modulate blood glucose levels (Fattaheian-Dehkordi et al., 2021; Serino et al., 2021; Hanalp et al 2023).

- The diuretic effect shown by extracts of *Taraxacum officinale* (dandelion). This plant stimulates urine production and supports kidney function, so it can be used to treat conditions such as oedema and urinary tract infections (Jalili et al., 2020).
- The immunomodulatory activity of *Echinacea* sp. enhances immune function and supports the body's defence mechanisms against infection and disease (Burlou-Nagy et al., 2022).

Although Asteraceae plants have a long tradition of use as medicinal remedies, it is important to validate their efficacy and safety through scientific research, as well as to discover other potential uses of these plants (Achika et al., 2014; Awwad et al., 2020).

1.2 Genus *Centaurea*

The genus *Centaurea* is one of the largest genera in the family Asteraceae, native to the Mediterranean and western Asia (**Figure 1.1**). It contains about 700 species and is one of the most numerous and taxonomically complex in the tribe Cardueae. The genus takes its name from Chiron, the centaur of Greek mythology who, according to legend, discovered the medicinal properties of these plants. *Centaurea* are flowering plants with a wide range of growth habits and flower forms. They can be annual, biennial or perennial, often with spiny leaves and stems. The flowers are typically dense clusters of small florets surrounded by bracts, giving them a distinctive appearance, whose colours vary from species to species and can include shades of purple, pink, blue, yellow and white (**Figure 1.2**). *Centaurea* plants are known for their ability to colonise and dominate disturbed ecosystems, sometimes becoming invasive in regions where they are introduced (Özel et al., 2023). Given this botanical richness, it is not surprising that several *Centaurea* species have long been used in folk medicine for a variety of ailments. Compounds belonging to virtually dozens of classes of secondary metabolites have been characterised in *Centaurea* species, but sesquiterpene lactones (Bruno et al., 2013; Sokovic et al 2017), flavonoids (Formisano et al., 2012) and alkaloids (Xia et al., 2018) are undoubtedly the hallmarks of this genus.

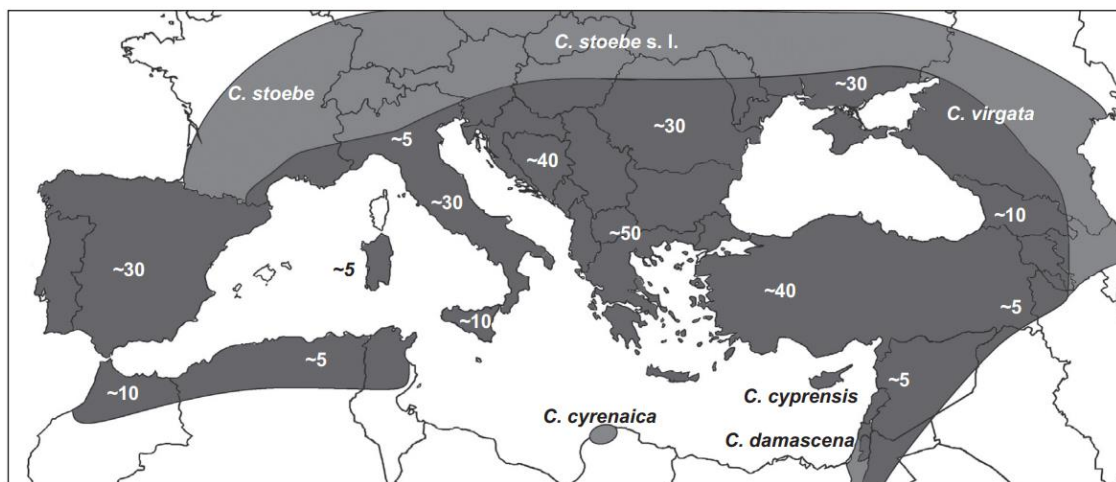


Figure 1.1 *Centaurea* species distribution in the Mediterranean basin. In dark grey: areas with more than one species occurring. In light grey: only one species occurring. Numbers show the approximate species number in the area (Hilpold et al., 2014).



Figure 1.2 Flowers of different *Centaurea* species: (A) *C. drabifolia*; (B) *C. kotschyi* var. *persica*; (C) *C. sicula*; (D) *C. patula*.

1.2.1 Secondary metabolites of genus *Centaurea*: sesquiterpenes

During my PhD, I focused on the apolar organic extracts of the aerial part of three different *Centaurea* species, whose most representative class of secondary metabolites are sesquiterpenoids. This class of chemical compounds is characterised by a variable rearranged C-15 skeleton. In general, a main feature is the presence of a γ -lactone ring closed at C-6 or C-8, with an exomethylene group conjugated to the carbonyl group, or a $-\text{CH}_2\text{OH}$ unit due to a reduction reaction. In most cases, the stereochemistry of lactonisation is trans, with the α -orientation of H-7, but cis analogues have also been reported, since the lactone ring can be fused to the main scaffold in either orientation (Patricia et al., 2018). There are four main scaffolds, elemanolides, eudesmanolides (6-6 bicyclic compounds), germacranolides (10-membered ring) and guaianolides (5-7 bicyclic compounds), as reported in **Figure 1.3**. All

terpenes are derived from the mevalonate biosynthetic pathway (MVA) using acetyl-CoA as the starting molecule and NADPH as the energy source. The acetyl-CoA building blocks lead first to the formation of mevalonic acid, which then leads to the formation of reactive five-carbon units, isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP). These give rise to farnesyl pyrophosphate (FPP) and finally the germacrene backbone, the precursor of all sesquiterpenes (Matejić et al., 2010).

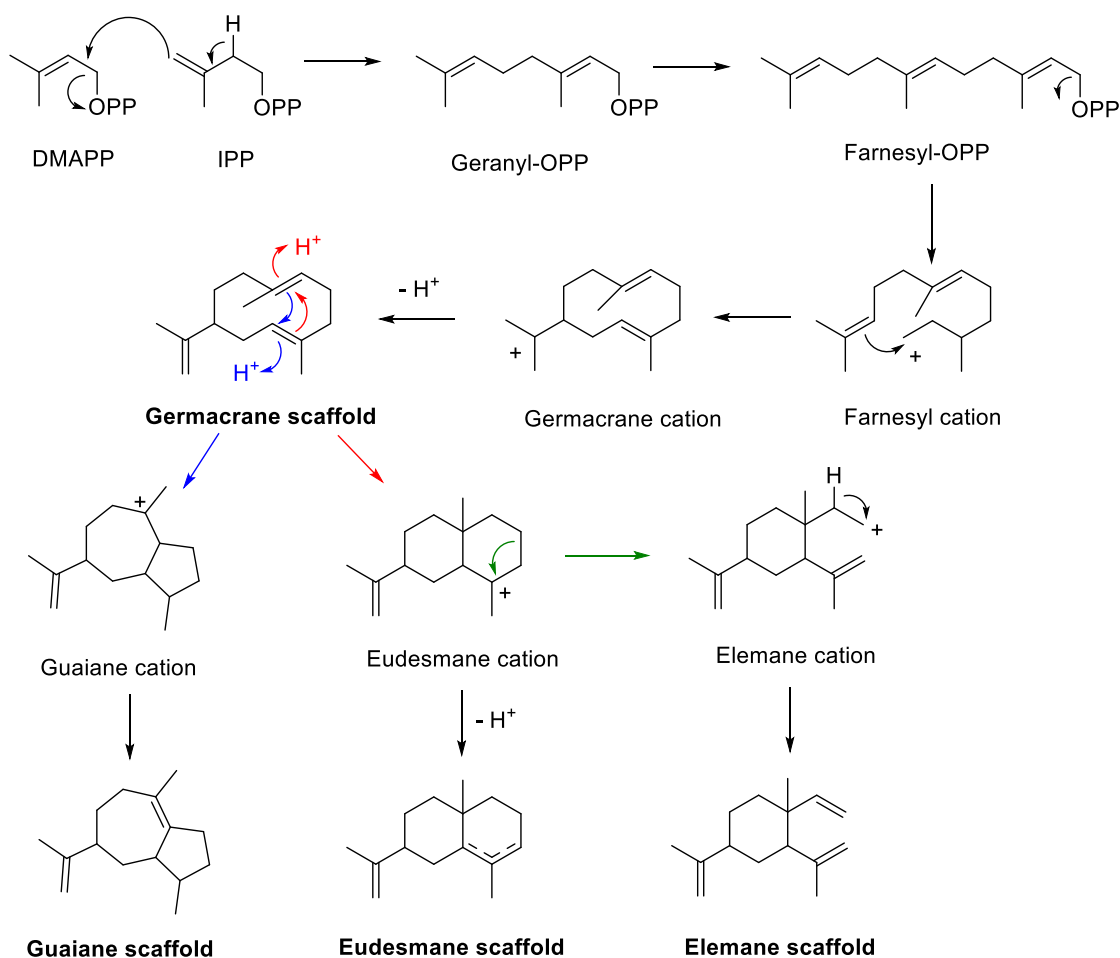


Figure 1. 3 Biogenesis of sesquiterpenoids skeletons

In general, these compounds are variously hydroxylated and decorated with different types of lateral side chains, creating a wide range of analogues with high chemical diversity that can be related to the plant taxonomic classification (Bruno et al., 2013).

Besides sesquiterpenoids, lignans and polyphenols are a well-represented class of specialized metabolites present in the organic extracts of *Centaurea* species, as shown in the following

chapters (Serino et al., 2022; Serino et al., 2024), along with indole alkaloids (Hodaj et al., 2017; Shoeb et al., 2005; Sarker et al., 1997).

1.2.2 Bioactivity of genus *Centaurea* and of sesquiterpenes

Combined with this exuberant chemical diversity, a plethora of pharmacological activities have been reported for *Centaurea* extracts. *Centaurea* species have been used in folk medicine to treat diarrhea, hypertension, headache, muscle pain, rheumatism and malaria (Gou et al., 2023). Other important properties are antimicrobial, antiviral and anti-inflammatory, which may be related to the presence of sesquiterpenes and polyphenols, together with antidiabetic activity (Sokovic et al., 2017). In particular, a recent investigation showed that several organic extracts of *Centaurea* aerial parts have a high inhibitory activity on α -amylase and α -glucosidase enzymes, thus controlling blood glucose levels (Fattaheian-Dehkordi et al., 2021). In these extracts, *Centaurea* species are generally rich in sesquiterpenoids and lignans, and their role in MetS modulation and related molecular targets is not yet fully understood. Sesquiterpenes have been studied mainly for their cytotoxic activity (Araki et al., 2024; De Cicco et al., 2021), which may be particularly relevant for chlorinated sesquiterpenoids (Estévez-Sarmiento et al., 2020), however some sesquiterpenoids have been reported for their potential role in the management of MetS. In general, these compounds have demonstrated to improve blood glucose and lipid levels, as well as improving insulin sensitivity. The molecular targets that may be involved include PPARs agonism, as described for the germacranolides isolated from *Tithonia diversifolia* (Lin et al., 2012); AMPK agonism, as postulated for the dimeric guaianolides byrsonines A and B isolated from *Byrsonima crassifolia* (Gutierrez and Ramirez, 2016); regeneration and revitalisation of residual β -cells, as suggested by the guaianolide eremanthin isolated from *Costus speciosus* (Salazar-Gomez et al., 2020) and α -glucosidase inhibition. In addition, sesquiterpene lactones have anti-inflammatory effects by suppressing the degradation of I- κ B (inhibitor of NF- κ B), reducing the activation of glucose-stimulated NF- κ B and ameliorating the pro-inflammatory state caused by hyperglycaemia (Awwad et al., 2020).

Thus, the following chapters of this part of the manuscript describe the phytochemical characterisation of *C. kotschy*, *C. sicula* and *C. patula* with insights into the biological activity of their secondary metabolites on targets related to MetS.

References

- Achika, J.I., Arthur, D.E., Gerald, I. and Adedayo, A., **2014**. A review on the phytoconstituents and related medicinal properties of plants in the Asteraceae family. *IOSR J Appl Chem*, 7(8), pp.1-8.
- Araki, K., Hara, M., Hamada, S., Matsumoto, T. and Nakamura, S., **2024**. Antiproliferative Activities of Cynaropicrin and Related Compounds against Cancer Stem Cells. *Chemical and Pharmaceutical Bulletin*, 72(2), pp.200-208.
- Awwad, A., Poucheret, P., Idres, A.Y., Bidel, L. and Tusch, D., **2020**. The bitter Asteraceae: An interesting approach to delay the metabolic syndrome progression. *NFS journal*, 18, pp.29-38.
- Bruno, M., Bancheva, S., Rosselli, S. and Maggio, A., **2013**. Sesquiterpenoids in subtribe Centaureinae (Cass.) Dumort (tribe Cardueae, Asteraceae): Distribution, ¹³C NMR spectral data and biological properties. *Phytochemistry*, 95, pp.19-93.
- Burlou-Nagy, C., Bănică, F., Jurca, T., Vicaș, L.G., Marian, E., Muresan, M.E., Bacskey, I., Kiss, R., Feher, P. and Pallag, A., **2022**. Echinacea purpurea (L.) Moench: Biological and pharmacological properties. A review. *Plants*, 11(9), p.1244.
- De Cicco, P., Busà, R., Ercolano, G., Formisano, C., Allegra, M., Taglialatela-Scafati, O. and Iannaro, A., **2021**. Inhibitory effects of cynaropicrin on human melanoma progression by targeting MAPK, NF-κB, and Nrf-2 signaling pathways in vitro. *Phytotherapy Research*, 35(3), pp.1432-1442.
- Estévez-Sarmiento, F., Saavedra, E., Ruiz-Estévez, M., León, F., Quintana, J., Brouard, I. and Estévez, F., **2020**. Chlorinated guaiane-type sesquiterpene lactones as cytotoxic agents against human tumor cells. *International Journal of Molecular Sciences*, 21(24), p.9767.
- Fattaheian-Dehkordi, S., Hojjatifard, R., Saeedi, M. and Khanavi, M., **2021**. A review on antidiabetic activity of Centaurea spp.: A new approach for developing herbal remedies. *Evidence-Based Complementary and Alternative Medicine*, 2021(1), p.5587938.
- Formisano, C., Rigano, D., Senatore, F., Bancheva, S., Maggio, A., Rosselli, S. and Bruno, M., **2012**. Flavonoids in Subtribe Centaureinae (Cass.) Dumort.(Tribe Cardueae, Asteraceae): Distribution and ¹³C-NMR Spectral Data. *Chemistry and Biodiversity*, 9(10), p.2096.

- Formisano, C., Sirignano, C., Rigano, D., Chianese, G., Zengin, G., Seo, E.J., Efferth, T. and Taglialatela-Scafati, O., **2017**. Antiproliferative activity against leukemia cells of sesquiterpene lactones from the Turkish endemic plant *Centaurea drabifolia* subsp. *detonsa*. *Fitoterapia*, 120, pp.98-102.
- Gou, J., Lu, Y., Xie, M., Tang, X., Chen, L., Zhao, J., Li, G. and Wang, H., **2023**. Antimicrobial activity in Asteraceae: the selected genera characterization and against multidrug resistance bacteria. *Heliyon*, 9(4).
- Gutiérrez, R.M.P. and Ramirez, A.M., **2016**. Hypoglycemic effects of sesquiterpene lactones from *Byrsonima crassifolia*. *Food science and biotechnology*, 25, pp.1135-1145.
- Hanalp, H.C., Dogan, A., Saygi, T.K., Donmez, F. and Battal, A., **2023**. Exploring phytochemical constituents of *Achillea arabica* Kotschy. ethanolic flower extract by LC-MS/MS and its possible antioxidant and antidiabetic effects in diabetic rats. *Zeitschrift Für Naturforschung C*, 78(5-6), pp.189-199.
- Hilpold, A., Garcia-Jacas, N., Vilatersana, R. and Susanna de la Serna, A., **2014**. Taxonomical and nomenclatural notes on *Centaurea*: A proposal of classification, a description of new sections and subsections, and a species list of the redefined section *Centaurea*.
- Hodaj, E., Tsiftoglou, O., Abazi, S., Hadjipavlou-Litina, D. and Lazari, D., **2017**. Lignans and indole alkaloids from the seeds of *Centaurea vlachorum* Hartvig (Asteraceae), growing wild in Albania and their biological activity. *Natural product research*, 31(10), pp.1195-1200.
- Jalili, C., Moradi, S., Babaei, A., Boozari, B., Asbaghi, O., Lazaridi, A.V., Kermani, M.A.H. and Miraghajani, M., **2020**. Effects of *Cynara scolymus* L. on glycemic indices: A systematic review and meta-analysis of randomized clinical trials. *Complementary therapies in medicine*, 52, p.102496.
- Jalili, C., Taghadosi, M., Pazhouhi, M., Bahrehmand, F., Miraghaee, S.S., Pourmand, D. and Rashidi, I., **2020**. An overview of therapeutic potentials of *Taraxacum officinale* (dandelion): A traditionally valuable herb with a reach historical background. *WCRJ World Cancer Res. J*, 7, p.e1679.
- Koc, S., Isgor, B.S., Isgor, Y.G., Shomali Moghaddam, N. and Yildirim, O., **2015**. The potential medicinal value of plants from Asteraceae family with antioxidant defense enzymes as biological targets. *Pharmaceutical biology*, 53(5), pp.746-751.

- Lawson, S.K., Sharp, L.G., Powers, C.N., McFeeters, R.L., Satyal, P. and Setzer, W.N., **2020**. Volatile compositions and antifungal activities of native american medicinal plants: Focus on the asteraceae. *Plants*, 9(1), p.126.
- Lin, H.R., **2012**. Sesquiterpene lactones from *Tithonia diversifolia* act as peroxisome proliferator-activated receptor agonists. *Bioorganic & medicinal chemistry letters*, 22(8), pp.2954-2958.
- Macedo, S.B., Ferreira, L.R., Perazzo, F.F. and Carvalho, J.T., **2004**. Anti-inflammatory activity of *Arnica montana* 6cH: preclinical study in animals. *Homeopathy*, 93(02), pp.84-87.
- Matejić, J., Šarac, Z. and Randelović, V., **2010**. Pharmacological activity of sesquiterpene lactones. *Biotechnology & Biotechnological Equipment*, 24(sup1), pp.95-100.
- Özel, Ç.A. and Maesaroh, S., **2023**. An Outlook of the Cultivation, Medicinal Properties, and Tissue Culture Techniques of *Centaurea* in Türkiye: A Comprehensive Review. *MAS Journal of Applied Sciences*, 8(2), pp.362-373.
- Patricia, V., Virginia, S. and Martino, S., **2018**. Sesquiterpene lactones advances in their chemistry and biological Aspects. *Switzerland: Springer*.
- Perez RM, Perez C, Zavala MA, Perez S, Hernandez H, Lagunes F. Hypoglycemic effects of lactucin-8-O-methylacrylate of *Parmentiera edulis* fruit. *Journal of Ethnopharmacology*. **2000** Aug 1;71(3):391-4.
- Piątkowska, E., Biel, W., Witkiewicz, R. and Kępińska-Pacelik, J., **2022**. Chemical composition and antioxidant activity of Asteraceae family plants. *Applied Sciences*, 12(23), p.12293.
- Preethi, K.C., Kuttan, G. and Kuttan, R., **2009**. Anti-inflammatory activity of flower extract of *Calendula officinalis* Linn. and its possible mechanism of action.
- Rolnik, A. and Olas, B., **2021**. The plants of the Asteraceae family as agents in the protection of human health. *International Journal of molecular sciences*, 22(6), p.3009.
- Salazar-Gómez, A., Ontiveros-Rodríguez, J.C., Pablo-Pérez, S.S., Vargas-Díaz, M.E. and Garduno-Siciliano, L., **2020**. The potential role of sesquiterpene lactones isolated from medicinal plants in the treatment of the metabolic syndrome—A review. *South African Journal of Botany*, 135, pp.240-251.
- Sarker, S.D., Savchenko, T., Whiting, P., Šik, V. and Dinan, L.N., **1997**. Moschamine, cis-moschamine, moschamindole and moschamindolol: four novel indole alkaloids from *Centaurea moschata*. *Natural Product Letters*, 9(3), pp.189-199.

- Serino, E., Chahardoli, A., Badolati, N., Sirignano, C., Jalilian, F., Mojarreb, M., Farhangi, Z., Rigano, D., Stornaiuolo, M., Shokoohinia, Y. and Taglialatela-Scafati, O., **2021**. Salvigenin, a trimethoxylated flavone from *Achillea wilhelmsii* C. Koch, exerts combined lipid-lowering and mitochondrial stimulatory effects. *Antioxidants*, 10(7), p.1042.
- Serino, E., Chianese, G., Musto, G., Zengin, G., Rigano, D., Stornaiuolo, M., Formisano, C. and Taglialatela-Scafati, O., **2022**. Guaiane-rich phytochemical profile of *Centaurea kotschyi* subsp. *persica* (Boiss.) Wagenitz and identification of hypoglycaemic metabolites. *Phytochemistry*, 199, p.113189.
- Serino, E., Rigano, D., Bruno, M., Pastore, A., Stornaiuolo, M., Formisano, C. and Taglialatela-Scafati, O., **2024**. Glucose Uptake-Stimulating Metabolites from Aerial Parts of *Centaurea sicula*. *Journal of Natural Products*, 87(4), pp.1179-1186.
- Shoeb, M., Celik, S., Jaspars, M., Kumarasamy, Y., MacManus, S.M., Nahar, L., Thoo-Lin, P.K. and Sarker, S.D., **2005**. Isolation, structure elucidation and bioactivity of schischkiniin, a unique indole alkaloid from the seeds of *Centaurea schischkinii*. *Tetrahedron*, 61(38), pp.9001-9006.
- Sokovic, M., Ciric, A., Glamoclija, J. and Skaltsa, H., **2017**. Biological activities of sesquiterpene lactones isolated from the genus *Centaurea* L.(Asteraceae). *Current pharmaceutical design*, 23(19), pp.2767-2786.
- Xia, C. and Tong, X., **2018**. Moschamine-related indole alkaloids. *The Alkaloids: Chemistry and Biology*, 79, pp.139-189.

CHAPTER 2: *Centaurea kotschy*

Genus *Centaurea* is endemic in Turkey, where more than 100 species are endemic and 200 were identified (Wagenitz, 1975). *C. kotschy* (Boiss.) Hayek is a Turkish species that includes at least two subspecies, *C. kotschy* subsp. *persica* (Boiss.), which is the most common one, and *C. kotschy* subsp. *floccosa* (Boiss.). *C. kotschy* subsp. *persica* has been used in folk medicine for dermatological and gynecological dysfunction, along with diabetes. The phytochemical investigations performed on this plant were basically related to the essential oil composition (Ertugrul et al., 2003) and the fatty acids profile (Zengin et al., 2011) and the characterization of the major sesquiterpenoids (Oksuz et al., 1982). Thus, we have performed a comprehensive phytochemical characterization of aerial parts organic extracts of this plant and tested the isolated compounds on GLUT carriers to identify the bioactive compounds responsible for the hypoglycemic activity of the extract.

2.1 Preliminary LC-MS/MS analysis

A Turkish sample of *Centaurea kotschy* subsp. *persica* aerial parts was exhaustively extracted with CH₂Cl₂ after being defatted with n-hexane. A preliminary metabolite profiling was performed through an LC-HRESIMS/MS analysis (**Figure 2.1**). The peaks included in the first part of the base peak chromatogram could be tentatively assigned to sesquiterpenoids and lignans (Rt 10-24 min), while the others were mostly fatty acids. Taxonomically, this species belongs to the same monophyletic core group (Jacea) and the same Cheirolepis clade of *C. drabifolia*, previously investigated in our lab (Formisano et al., 2017), thus we wondered if these two species share a similar phytochemical profile. *C. drabifolia* biosynthesize exclusively guaiane-type sesquiterpene lactones, variously decorated with acyl moieties and different side chains. However, the comparison of the two LC-MS/MS profiles highlighted that, although *C. kotschy* subsp. *persica* seemed to produce guaianolides according to the fragmentation pattern analysis, it appeared to be a significant difference in the substituents bind on guaiane scaffold.

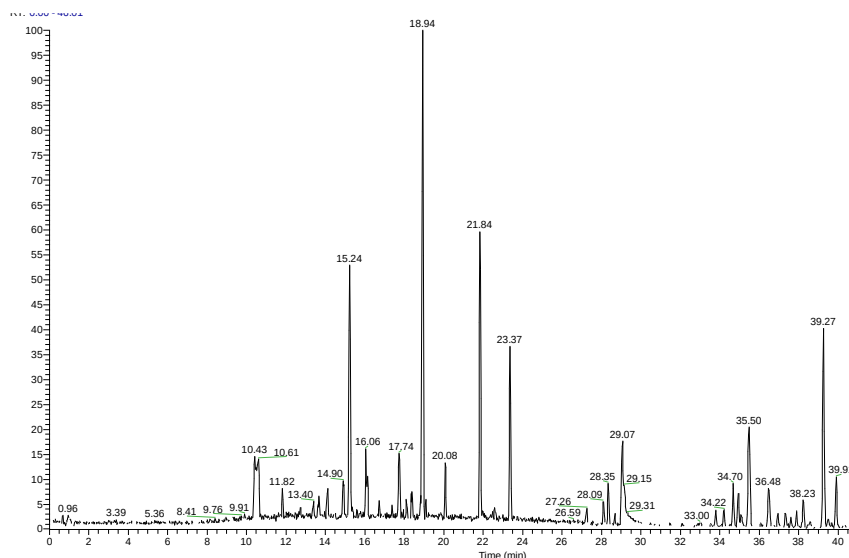


Figure 2. 1 Base peak chromatogram of *C. kotschy* subsp. *persica* aerial parts organic extract

2.2 Phytochemical composition of the aerial parts

In order to conduct a detailed characterization and to obtain the isolated pure compounds to investigate the responsible ones for the bioactivity of the extracts, the CH_2Cl_2 extract was firstly purified on MPLC, and successively on RP-HPLC. The chromatographic purifications led to the isolation of 15 compounds (**Figure 2.2**), including the guaianolides **1–9**, a monoterpene lactone (**10**), four known lignans (**11–14**), and an undescribed neolignan (**15**). Moreover, the guaiane sesquiterpenes **7** and **8** were unprecedented in literature, and here described for the first time. The compounds already described in literature were identified by comparison of their spectral data with those reported as 3 β ,8 α ,11 α ,13-tetrahydroxy-10(14)-guaien-1 α ,4 β ,5 α ,6 β H-6 α ,12-olide (**1**) (Rong et al., 2009), rhizantholide A (**2**) (Shakeri et al., 2019), desacylcynaropicrin (**3**) (Choi et al., 2005), salograviolide A (**4**) (Daniewski et al., 1992), kandavanolide (**5**) (Choi et al., 2005), cynarinin A (**6**) (Li et al., 2005), 4 β ,15-dihydro-3-dehydrosolstitialin A monoacetate (**9**) (Formisano et al., 2017), daphnauranin E (**10**) (Huang et al., 2018), matairesinol (**11**) (Yang et al., 2013), matairesinoside (**12**) (Ban et al., 2018), arctiin (**13**) (Yang et al., 2013) and arctigenin (**14**) (Yang et al., 2013). Compound **1** had been previously isolated only from *Cynara scolymus*, thus it is the first report of this compound in a *Centaurea* plant. Similarly, the monoterpene daphnauranin E (**10**) had been reported only from plants of the genus *Daphne* (family *Thymelaeaceae*).

PART 2

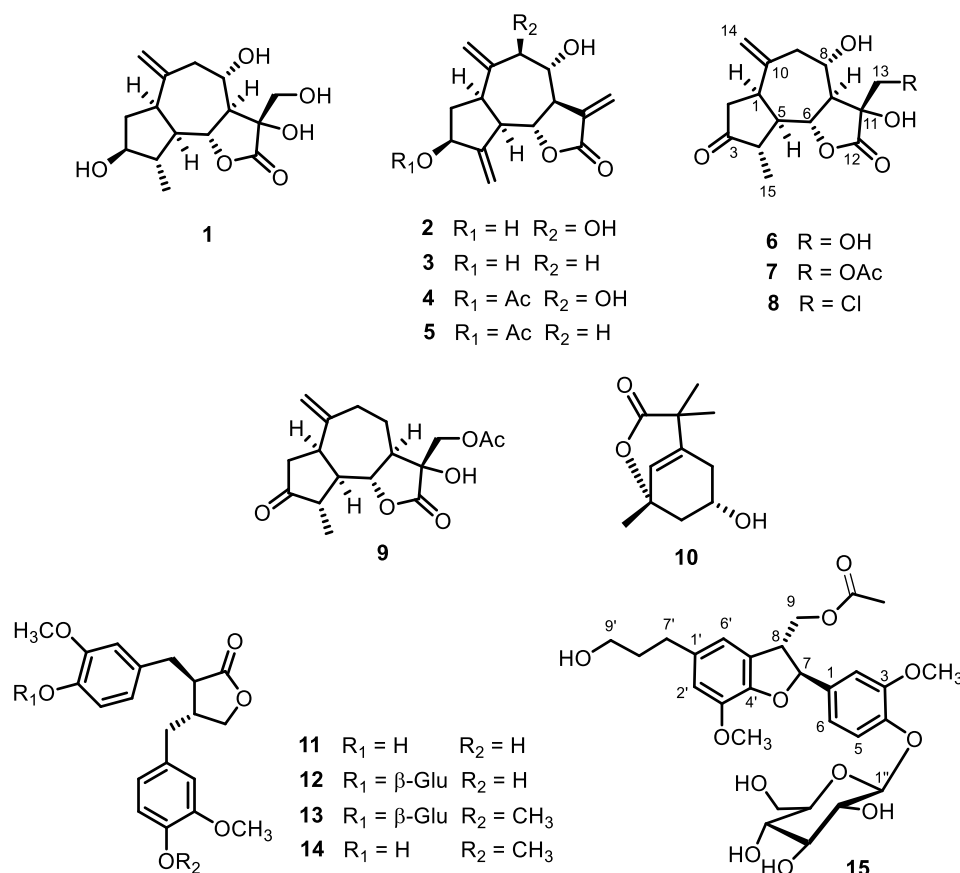


Figure 2. 2 Secondary metabolites isolated from *C. kotschyi* subsp. *persica* aerial parts

Kotschyol A (**7**) appeared as a pale yellow amorphous solid. From HRESIMS analysis a peak at m/z 361.1266 $[M+Na]^+$ was detected suggesting $C_{17}H_{22}O_7$ as molecular formula (calculated for $C_{17}H_{22}O_7Na$ 361.1263). The ^1H NMR analysis disclosed strong similarities between **7** and the known cynarinin A (**6**), including one exomethylene group at δ_H 5.09 and 4.83 and two oxymethines at δ_H 4.08 and 3.91, but an acetyl signal at δ_H 2.10 was also present in the spectrum of **7**. As expected, the ^{13}C data of **7** resemble those of **6**, including the two carbonyl groups, a ketone at δ_C 218.0 and a lactone at δ_C 174.8, four oxygenated sp^3 carbons (δ_C 81.2, 70.0, 76.5, 64.1) and two additional sp^2 carbons. To elucidate the planar structure of **7**, 2D NMR experiments were performed, including HSQC, through which each proton was assigned to its carbon atom, COSY, which evidenced the single extend spin system in line with the guaianane skeleton, and finally HMBC which confirmed the structure hypothesis and the position of the substituents (**Figure 2.3**). In fact, the HMBC crosspeaks between the from both the allylic H-1 (δ_H 3.13) and H₃-15 to the ketone carbonyl at C-3, from the hexomethylene protons H₂-14 to C-1 and C-9, from H-6 to C-12 and from the isolated oxymethylene signals at δ_H 4.62 and 4.41 to

the oxygenated C-11 (δ_c 76.5), the lactone carbonyl C-12, and the acetyl carbonyl at δ_c 170.0, this latter correlation unambiguously placing the acetyl group at C-13. The NMR data of compound **7** are reported in the Experimental Section.

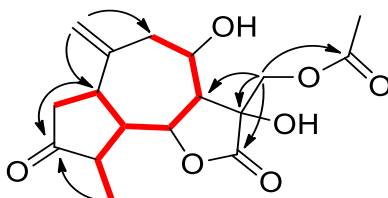


Figure 2.3 COSY (red bold) and diagnostic HMBC correlations (arrows) detected for kotschyol A (**7**)

The stereochemical features of kotschyol A were elucidated through chemical conversions taking advantage of the strong similarities with **6**. Both kotschyol A and cynarinin A were submitted to a peracetylation reaction with acetic anhydride in dry pyridine. As both sesquiterpenes gave the same triacetylated compound (**7a**), as evidenced by ^1H NMR and $[\alpha]_D$, we could affirm that they share the unprecedented kotschyol A shares the same absolute configuration of the known cynarinin A (**Figure 2.4**).

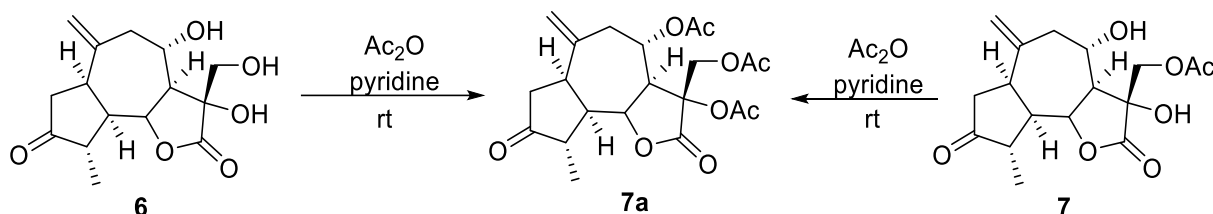


Figure 2.4 Peracetylation reaction for the determination of kotschyol A (**7**) absolute configuration.

Although somewhat unexpected and never described for sesquiterpenoid derivatives, the mild acetylation of tertiary α -hydroxy- γ -lactones is not unprecedented in the literature (Rath et al., 2005). Incidentally, the structure of kotschyol A has been mentioned by Macias et al. (2006) as a possible synthetic target but, to our knowledge, it had never been isolated or synthesized before the present study.

Kotschyol B (**8**) was obtained as a colorless amorphous solid with molecular formula $\text{C}_{15}\text{H}_{19}\text{ClO}_5$ as determined by positive HR-ESIMS peak at m/z 337.0822 $[\text{M}+\text{Na}]^+$ (calculated for $\text{C}_{15}\text{H}_{19}^{35}\text{ClO}_5$ 337.0819). The presence of the chlorine atom was clearly visible by the isotopic pattern showing a $\text{M}+2$ peak in 3:1 ratio. Both the molecular formula and 1D NMR data of **8**

showed close similarity with kotschyol A, but from the ^1H NMR spectrum was observed the absence of the acetoxyl singlet, which was present in **7**, suggesting its replacement with the chlorine atom. This hypothesis was supported with the chemical shift detected for the H₂-13 (δ_{H} 3.76 and 4.15 in **7** in place of δ_{H} 4.41 and 4.52 in **8**) and the ^{13}C NMR resonances that showed the C-13 resonance at δ_{C} 43.5, in full agreement with chlorine substitution. The structural assignment was confirmed with 2D NMR experiment that showed an almost complete overlapping of the remaining ^{13}C NMR resonances, along with the proton-proton coupling constants, suggesting that the two kotschyol also share the same relative configuration. The NOESY correlations between H₂-13/H-8, H-5/H-7, H-1/H-5 and H₃-15/H-5 fully supported this hypothesis. The NMR data of compound **8** are reported in the Experimental Section.

Compound **15** was isolated as a colorless amorphous solid with $\text{C}_{28}\text{H}_{36}\text{O}_{12}$ as molecular formula as determined through positive HRESI-MS analysis that showed a peak at m/z 587.2100 $[\text{M}+\text{Na}]^+$ (calculated for $\text{C}_{28}\text{H}_{36}\text{O}_{12}\text{Na}$ 587.2104). The ^1H NMR spectrum of **15** showed the presence of five signals attributable to substituted aromatic rings (two doublets and three singlets located between δ_{H} 7.16 and δ_{H} 6.71), two methoxy groups (δ_{H} 3.83 and 3.86), one acetyl methyl (δ_{H} 2.03), and a series of partly overlapped multiplets located in the midfield region. Interpretation of the 2D COSY spectrum evidenced the presence of four moieties (**Figure 2.5**), including a sugar unit. All the proton signals were associated to the directly linked carbon atoms through the 2D HSQC spectrum. The doublet at δ_{H} 4.90 showed a clear HSQC cross-peak with the signal at δ_{C} 101.3, thus confirming its anomeric nature. Then, the glycosylated dihydrobenzofuran neolignan skeleton of compound **15** could be assembled taking advantage of the cross-peaks present in the 2D HMBC spectrum. The dihydrobenzofuran core was indicated by H-7/C-4' (δ_{C} 145.7), H-8/C-5' and H-8/C-6' correlations, while attachment of the 3- hydroxypropyl at C-1' and of the methoxy group at C-3' was indicated by H-2'/C-7' and $-\text{OCH}_3/\text{C}-3'$ correlations, respectively. Similarly, attachment of the disubstituted aromatic ring at C-7 was indicated by correlations between the deshielded oxymethine H-7 (δ_{H} 5.50) and C-1 (δ_{C} 136.0), C-2, and C-6, while the methoxy group was placed at C-3 on the basis of the HMBC correlation $\text{OCH}_3/\text{C}-3$. The anomeric signal H-1'' showed HMBC correlation with C-4, thus allowing the placement of the sugar unit. The structure of compound **15** was completed by attachment of the acetyl group at C-9 (HMBC cross-peak of H₂-9 with the ester carbonyl carbon at δ_{C} 171.0). The available proton-proton coupling constants and the ^{13}C NMR values, deduced by the HSQC spectrum, allowed a confident assignment of the sugar as a β -

glucopyranosyl unit. The *trans* relative arrangement of the substituents on the dihydrobenzofuran ring was evidenced by the NOESY correlation between H-7 and H₂-9. Compound **15** is the 4-*O*-glucosylated analogue of crataegifin A, a neolignan recently reported from the Chinese fruit-bearing *Crataegus pinnatifida* (family Rosaceae) (Guo et al., 2019). Since both enantiomers of crataegifin A were isolated and characterized from this plant, we had the opportunity to assign the absolute configuration of **15** by comparing its physical data with those of the aglycone. A comparison of the $[\alpha]_D$ values was not feasible due to the presence of the sugar moiety in **15**, but we reasoned that the ECD spectrum should not be significantly affected by the presence of the sugar. Thus, the close similarity of the ECD profile of **15** (λ_{max} 233, $\Delta\epsilon$ = -12.50) with the ECD spectrum reported for the (7*S*,8*R*)-crataegifin A (λ_{max} 241, $\Delta\epsilon$ = -13.61) allowed a confident assignment of the absolute configuration of the dibenzofuran moiety of **15**. The NMR data of **15** are reported in the Experimental Section.

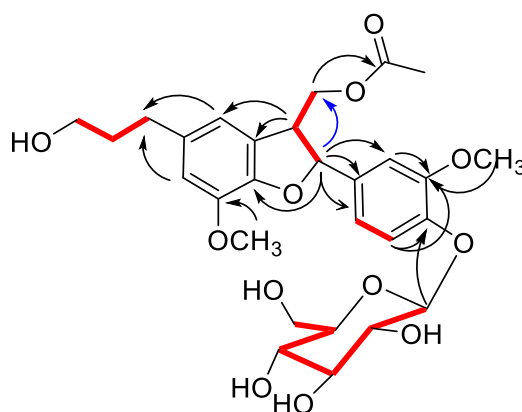


Figure 2.5 COSY (red bold), HMBC correlations (black arrow) and NOESY key correlation (in blue) of **15**.

2.3 Biological activity

C. kotschy subsp. *persica* is one of the several species belonging to genus *Centuraea* that have been reported for their antidiabetic activity (Sokovic et al., 2017), but the bioactive compounds have not been identified. In order to investigate the hypoglycemic properties of this plant, we have tested all the isolated metabolites on C2C12 cells, mouse skeletal muscle cells well responsive to insulin, to monitor the mobilization of GLUT transporters over the cell surface. GLUTs mobilization was assessed by using 2-NBDG, a fluorescent analogue of glucose, covalently linked to the fluorescent molecule nitro blue tetrazolium (NBT). When added to the

culture medium, 2-NBDG binds to GLUTs localized on the cell surface, allowing their identification by fluorescence. Among all the compounds isolated in this study, **1–3** and **6** exerted an insulin-like effect on GLUT mobilization, with EC₅₀ in the range 1.2–1.5 μ M, with rhizantholide A (**2**) showing the highest potency (3.3), while daphnauranin E, lignans **11–14** and neolignan **15** resulted inactive (**Table 2.1**). Considering the guaianolides analogues, that are the compounds responsible for the hypoglycemic activity of the extract, the bioactivity variations can be related to the structural modifications of the compounds. Indeed, in the analogues series **2–5**, the acetylation at C-3 in compound **4** and **5** results in a marked decrease of activity. Moreover, considering the analogues series **6–9**, the replacement of the hydroxylic group at C-13 with other substituents, as for the two new kotschyols, also causes a decrement in the activity. On the other hand, the Michael acceptor behavior seems to be not related to this activity since the non-electrophilic **1** and **6** are much more active than the exomethylene- γ -lactones **4** and **5**.

Table 2.1 GLUT mobilization activity of compounds 1-15 tested on C2C12 cells*

Compound	EC ₅₀ (μ M $\pm$ S.D.)	Potency
1	1.3 $\pm$ 0.3***	1.7 $\pm$ 0.2*
2	1.2 $\pm$ 0.3***	3.3 $\pm$ 0.3***
3	1.4 $\pm$ 0.4***	1.2 $\pm$ 0.1*
4	> 30	-
5	10.3 $\pm$ 0.4**	1.5 $\pm$ 0.2**
6	1.5 $\pm$ 0.4***	1.3 $\pm$ 0.2*
7	> 30	-
8	> 30	-
9	> 30	-
10	> 30	-
11	> 30	-
12	> 30	-
13	> 30	-
14	> 30	-
15	> 30	-
Insulin	0.2 $\pm$ 0.3***	7.2 $\pm$ 0.5***
Vehicle (DMSO)	30.0 $\pm$ 1.8	1.0 $\pm$ 0.1

*Increase in GLUTs localization on the plasma membrane of the cells expressed as EC₅₀ and potency of the compounds (see Experimental for details). n = 3 independent experiments. Results are expressed as mean $\pm$ s.d.; Comparison vs vehicle (for EC₅₀) or vs untreated cells (for potency); p values = * < 0.01, ** < 0.05, *** < 0.001; Potency were calculated only for compound showing EC₅₀ at concentration lower than vehicle.

2.4 Conclusions

In conclusion, the metabolomic analysis of aerial parts of the Turkish endemic species *C. kotschy* subsp. *persica* resulted in the isolation of 15 pure compounds, including two undescribed guaianolides (**7** and **8**) and an undescribed neolignan glucoside (**15**). *C. kotschy* subsp. *persica* sesquiterpenoid metabolism appears to be exclusively devoted to the production of guaianolides, similarly to *C. drabifolia*, previously analyzed by us (Formisano et al., 2017) and belonging to the same clade. The availability of the major metabolites of the *C. kotschy* subsp. *persica* extract allowed us to verify the antidiabetic activity associated to the plant extract. Given the limited amounts of compounds isolated, we chose a GLUT mobilization assay, based on the use of 2-NBDG, a fluorescent analogue of glucose. At least four guaianolides showed a significant bioactivity and these compounds can be viewed as the responsible of the activity of extract. Only a few guaianolides have been reported in the literature for their hypoglycemic activity, namely lactucin-8-O-methylacrilate (Perez et al., 2000) and byrsoninas A-B (Gutierrez and Ramirez, 2016), indicating the presence of strict structural requirements, worthy of further investigation.

2.5 Experimental section

2.5.1 Plant material

The aerial parts of *Centaurea kotschy* subsp. *persica* (Boiss.) Greuter (Asteraceae) were collected at the flowering stage in July 2010 near Beysehir, Konya, Turkey (37.688395, 31.718073). The plant was identified by Evren Yildiztugay from Section of Botany, Department of Biology, Faculty of Science, Selcuk University. A voucher specimen is deposited at KNYA herbarium (Herbarium of the Biology Department of Selcuk University, Voucher No: KNYA 8919). The plant was immediately extracted and the extracts stored at – 18°C until further investigation.

2.5.2 Extraction and isolation

C. kotschy subsp. *persica* aerial parts (500 g) were dried in an oven at 40°C and defatted at room temperature using *n*-hexane (1.3 L for 3 days, three times) and extracted with CH₂Cl₂ (1.3 L for 3 days, three times). After filtration and in vacuo solvent evaporation, 1.9 g of dry crude

CH₂Cl₂ extract were obtained. The CH₂Cl₂ extract was preliminary purified on MPLCDAD on a Cartridge C18 60 Å 50 µm - size 40 (55 g) – column volume (CV) 50 mL, using a mixture of A: water with a 0.1% formic acid, B: acetonitrile as mobile phase. The following gradient method was applied: 90%A – 10%B for CV 0-2; 80%A – 20%B for CV 4-7; 70%A – 30%B for CV 9-12; 60%A – 40%B for CV 14-17; 50%A – 50%B for CV 19- 22; 40%A – 60%B for CV 24-27; 30%A – 70%B for CV 29-32; 20%A - 80%B for CV 34-37; 10%A – 90%B for CV 39-42, with 20.0 mL/ min as flow rate. The UV detection wavelength was set at 227 nm. 17 fractions (labeled A-S) were obtained and salograviolide A (**4**, 265 mg, fraction G) was directly obtained in pure form from the MPLC. All the other fractions were further purified on RP-HPLC by semipreparative or analytical columns, yielding in the isolation of compounds **1–15**. Fraction C (94.0 mg) was purified by HPLC using a mixture of A, B and C (70:25:5) as mobile phase to afford 3β,8α,11α,13-tetrahydroxy-10(14)-guaien- 1α,4β,5α,6βH-6α,12-olide (**1**) (Rt = 9.5 min 1.5 mg), rhizantholide A (**2**, Rt = 13.4 min, 4.2 mg) and cynarinin A (**6**, Rt = 18.0 min, 25.0 mg). Fraction E (47.3 mg) was separated by HPLC-UV on a C-18 column using the following isocratic elution: 75%A – 20%B – 5%C, flow rate 3.0 mL/min, to afford desacylcynaropicrin (**3**, Rt = 12 min, 3.5 mg) and kotschyol A (**7**, Rt = 16 min, 3.5 mg). Fraction F (90.2 mg) was separated on HPLC (A:C 60:40, at 1.0 mL/min flow rate) affording daphnauranin E (**10**, Rt = 11 min, 1.5 mg) matairesinoside (**12**, Rt = 14 min, 6.9 mg), arctiin (**13**, Rt = 21.5 min, 2.5 mg), kotschyol B (**8**, Rt = 20.5 min, 6.8 mg) and the 4-O-glucosylcrataegifin A (**15**, Rt = 26.5 min, 1.5 mg). Fraction H (130.5 mg), was separated by RP-18 HPLC-UV using the following elution gradient: from 60%A to 35%B - 5%C to 30%A – 65%B - 5%C in 32 min, flow rate 3 mL/min to yield **9** (Rt = 10 min, 1.5 mg, matairesinol (**11**, Rt = 11 min 7.5 mg), kandavanolide (**5**, Rt = 12 min, 11.5 mg) and arctigenin (**14**, Rt = 14.9 min 6.5 mg).

2.5.3 Spectral Data

Kotschyol A (**7**) : Pale yellow amorphous solid; $[\alpha]_D^{25} + 109.3$ (c 0.01, MeOH); ¹H NMR (600 MHz, CDCl₃) δ: 5.09 (1H, s, H-14b), 4.83 (1H, s, H-14b), 4.62 (1H, d, *J* 11.4 Hz, H-13b), 4.41 (1H, d, *J* 11.4 Hz, H-14a), 4.08 (1H, t, *J* 10.4 Hz, H-6), 3.91 (1H, overlapped), 3.13 (1H, m, H-1), 2.88 (1H, dd, *J* 12.7, 5.6 Hz, H-9b), 2.72 (1H, t, *J* 10.4 Hz, H-7), 2.55 (1H, dd, *J* 17.5, 3.9 Hz, H-2b), 2.51 (1H, dd, *J* 17.5, 8.4 Hz, H-2a), 2.33 (1H, overlapped, H-4), 2.32 (1H, overlapped, H-5), 2.24 (1H, overlapped, H-9a), 2.10 (3H, s, OAc), 1.25 (3H, d, *J* 5.6 Hz, H-15). ; ¹³C NMR data (125 MHz, CDCl₃): 218.0 (C-3), 174.8 (C-12), 170.0 (OAc), 143.0 (C-10), 115.6 (C-14), 81.2 (C-6), 76.5 (C-

11), 70.0 (C-8), 64.1 (C-13), 56.1 (C-7), 51.9 (C-5), 48.3 (C-9), 47.4 (C-4), 43.3 (C-2), 39.8 (C-1), 20.5 (OAc), 14.7 (C-15). HR-ESIMS: m/z 361.1266 $[M+Na]^+$, ($C_{17}H_{22}O_7Na$ requires m/z 361.1263).

Kotschyol B (**8**) : Colorless amorphous solid; $[\alpha]_D^{25} + 133.5$ (c 0.01, MeOH); ^1H NMR and ^{13}C NMR (600 MHz, $CDCl_3$) δ : 5.11 (1H, s, H-14b), 4.83 (1H, s, H-14a), 4.23 (1H, overlapped, H-6), 4.20 (1H, overlapped, H-8), 4.15 (1H, d, J 11.2 Hz, H-13b), 3.76 (1H, d, J 11.2 Hz, H-13a), 3.13 (1H, m, H-1), 2.91 (1H, dd, J 12.0, 5.6 Hz, H-9b), 2.76 (1H, t, J 10.5 Hz, H-7), 2.56 (1H, dd, J 17.5, 3.9 Hz, H-2b), 2.34 (1H, overlapped, H-5), 2.33 (1H, overlapped, H-4), 2.50 (1H, dd, J 17.5, 8.4 Hz, H-2a), 2.29 (1H, t, J 12.0 Hz, H-9a), 1.25 (3H, d, J 5.6 Hz, H-15). ^{13}C NMR data (125 MHz, $CDCl_3$): 218.2 (C-3), 175.3 (C-12), 142.9 (C-10), 115.7 (C-14), 82.2 (C-6), 77.6 (C-11), 70.0 (C-8), 56.1 (C-7), 51.8 (C-5), 48.6 (C-9), 47.2 (C-4), 43.5 (C-13), 43.4 (C-2), 39.6 (C-1), 14.5 (C-15). ^1H NMR (600 MHz, CD_3OD) δ : 5.05 (1H, s, H-14b); 4.76 (1H, s, H-14a); 4.30 (1H, d, J 10.8 Hz, H-13b), 3.73 (2H, d, J 10.8 Hz, H-13a); 4.20 (1H, t, J 8.8 Hz, H-6); 4.13 (1H, m, H-8); 3.20 (1H, m, H-1), 2.86 (1H, dd, J 12.5, 5.9 Hz, H-9b); 2.23 (1H, t, J 10.9 Hz, H-9a); 2.72 (1H, t, J 10.5 Hz, H-7); 2.58 (1H, dd, J 19.0, 9.3 Hz, H-2b); 2.45 (3H, dd, J 19.0, 9.3 Hz, H-2a); 2.32 (2H, m, H-4 and H-5 overlapped signals), 1.20 (3H, d, J 6.0 Hz, H-15); HR-ESIMS: m/z 337.0822 $[M+Na]^+$, ($C_{15}H_{19}^{35}ClO_5$ requires m/z 337.0819).

4-O-glucosylcrataegifin A (**15**) : Colorless amorphous solid; $[\alpha]_D^{25} - 6.8$ (c 0.01, MeOH); ^1H NMR (600 MHz, CD_3OD): 7.16 (1H, d, J 8.2 Hz, H-5), 7.03 (1H, s, H-2), 6.93 (1H, d, J 8.2 Hz, H-6), 6.76 (1H, s, H-2'), 6.71 (1H, s, H-6'), 5.50 (1H, d, J 6.5 Hz, H-7), 4.90 (1H, d, J 7.4 Hz, H-1''), 4.43 (1H, dd, J 10.9, 5.4, H-9b), 4.33 (1H, dd, J 10.9, 8.0 Hz, H-9a), 3.88 (1H, overlapped, H-6''b), 3.86 (3H, s, OMe-3'), 3.83 (3H, s, OMe-3), 3.69 (1H, overlapped, H-6''a), 3.68 (1H, m, H-8), 3.58 (2H, overlapped, H-9'), 3.48 (1H, t, J 7.4 Hz, H-2''), 3.46 (1H, overlapped, H-3''), 3.39 (2H, overlapped, H-4'' and H-5''), 2.64 (2H, t, J 7.6 Hz, H-7'), 2.03 (3H, s, OAc) 1.81 (2H, m, H-8'). ^{13}C NMR (125MHz, $CDCl_3$): 171.0 (OAc), 149.0 (C-3), 146.0 (C-4), 145.7 (C-4'), 143.3 (C-3'), 136.0 (C-1), 135.1 (C-1'), 127.0 (C-5'), 118.3 (C-6), 116.6 (C-5), 116.3 (C-6'), 112.7 (C-2'), 109.8 (C-2), 101.3 (C-1''), 87.5 (C-7), 76.6 (C-3''), 76.5 (C-5''), 73.6 (C-2''), 70.0 (C-4''), 65.3 (C-9), 61.0 (C-6''), 60.7 (C-9'), 55.3 (OMe-3 and OMe-3'), 50.6 (C-8), 34.6 (C-8'), 31.5 (C-7'), 19.4 (OAc). ECD (MeOH): λ_{max} 233, $\Delta\epsilon = -12.50$. HR-ESIMS: m/z 587.2100 $[M+Na]^+$ ($C_{28}H_{36}O_{12}Na$ requires m/z 587.2104).

Compound **7a** : $[\alpha]_D^{25} + 47.3$ (c 0.01, MeOH); ^1H NMR (600 MHz, CDCl_3) δ_{H} 5.14 (1H, s, H-14b); 4.96 (1H, m, H-8); 4.87 (1H, s, H-14a); 4.35 (1H, d, J 11.2 Hz, H-13b), 4.26 (1H, d, J 11.2 Hz, H-13a), 4.08 (1H, t, J 9.5 Hz, H-6), 3.53 (1H, t, J 10.4 Hz, H-7), 3.12 (1H, m, H-1), 2.95 (1H, dd, J 12.8, 5.6 Hz, H-9b), 2.55 (1H, dd, J 19.0, 8.8 Hz, H-2b), 2.48 (2H, overlapped signals, H-2a and H-5), 2.34 (1H, m, H-4), 2.15 (1H, overlapped, H-9a), 2.15 (3H, s, Ac), 2.11 (3H, s, Ac), 2.08 (3H, s, Ac), 1.29 (3H, d, J 7.2 Hz, H-15). ESIMS: m/z 445.14 $[\text{M}+\text{Na}]^+$.

2.5.4 Acetylation of kotschyol A and cynarinin A

2.0 mg of kotschyol A (**7**) were solubilized in dry pyridine (600 μL) and treated with acetic anhydride (50 μL) overnight at rt. After addition of MeOH and evaporation, the obtained compound was purified by HPLC (n-hexane/EtOAc 7:3) to afford compound **7a** (1.9 mg). The same procedure was applied to cynarinin A (**6**, 1.9 mg) obtaining 1.6 mg of **7a**.

2.5.5 LC-MS/MS analysis

The CH_2Cl_2 extract of *C. kotschyi* subsp. *perisca* aerial parts was subjected to a preliminary LC-MS/MS analysis to perform a qualitative profile of the extract using an LTQ-XL Ion Trap mass spectrometer equipped with an Ultimate 3000 HPLC. Chromatographic separation was achieved using a Kinetex Polar C18 column (100 $\times$ 3.0 mm, 100 $\text{\AA}$, 2.6 μm). The injection volume was 0.5 μL /min and a mobile phase consisting of a combination of A (0.1% formic acid in water, v/v) and B (0.1% formic acid in CH_3CN , v/v); applying the following linear gradient method: from 5% to 60% B in 25 min, from 60% to 95% B in 10 min, and held at 95% B for 5 min. The MS and MS ^{n} spectra, in positive mode, were recorded in Data Dependent Acquisition mode inducing fragmentation of the most intense 5 peaks for each scan. Source conditions were: spray voltage: 3.5 kV (positive mode); capillary voltage: 25 V; source temperature: 320°C; normalized collision energy: 25. The acquisition range was m/z 150–1500.

2.5.6 2-NBDG glucose assays

C2C12 muscle myoblasts were grown in DMEM 10% FBS supplemented with glutamine, penicillin and streptomycin and sub-cultivated every two days to avoid over-confluency. For compounds testing, cells were plated (5 $\times$ 10³/well) in 96-well microtiter plates (PerkinElmer, Waltham, USA) in a final volume of 100 μL /well of culture medium. 24 h after plating, the culture medium was supplemented with crude fractions (30 mg/L) or with the indicated isolated

compounds (range of concentration tested from 0.1 up to 30 μ M) for 6 h. GLUT transporters localized on the surface were identified upon cell staining with the fluorescent analogue of glucose, 2-(N-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)-2-deoxyglucose (2-NBDG, 400 μ M working solution, Invitrogen) diluted in DMEM containing 10% FBS. Plates were incubated with the fluorescent probe for 30 min to be then washed twice in PBS and fixed in 3.7% PFA. Cells presenting a plasma membrane localized 2- NBDG signals were identified and counted under a fluorescent microscope and their total fluorescence intensity measured. EC_{50} of the compounds indicates the concentration of the compounds at which the number of cells presenting a diffuse plasma membrane localized 2- NBDG staining reaches 50% of the maximum. EC_{50} were considered relevant only when statistically different from the vehicle (DMSO). Potency of the compounds indicates the fold increase of total cell fluorescence measured upon incubation of the compounds diluted at their EC_{50} , compared to untreated cells. Comparisons and differences were analyzed for statistical significance by one-way ANOVA and Bonferroni multiple comparisons test with a single pooled variance. Statistical analysis was performed using GraphPad Prism (GraphPad Software 7.03, San Diego, USA). Values were considered statistically different when p value were * <0.01, ** <0.05, *** <0.001.

References

- Ban, N.K., Thoa, N.T.K., Linh, T.M., Giang, V.H., Trang, D.T., Nhiem, N.X., Tai, B.H., Quang, T.H., Yen, P.H., Van Minh, C. and Van Kiem, P., **2018**. Chemical constituents of *Vitex trifolia* leaves. *Natural Product Communications*, 13(2), p.1934578X1801300205.
- Choi, S.Z., Choi, S.U. and Lee, K.R., **2005**. Cytotoxic sesquiterpene lactones from *Saussurea calciccola*. *Archives of pharmacal research*, 28, pp.1142-1146.
- Daniewski, W.M., Nowak, G., Routsis, E., Rychlewska, U., Szczepańska, B. and Skibicki, P., **1992**. Salograviolide A, a sesquiterpene from *Centaurea salonitana*. *Phytochemistry*, 31(8), pp.2891-2893.
- Ertugrul, K., Dura, H., Tugay, O., Flamini, G., Cioni, P.L. and Morelli, I., **2003**. Essential oils from flowers of *Centaurea kotschyi* var. *kotschyi* and *C. kotschyi* var. *decumbens* from Turkey. *Flavour and fragrance journal*, 18(2), pp.95-97.
- Fattaheian-Dehkordi, S., Hojjatifard, R., Saeedi, M. and Khanavi, M., **2021**. A review on antidiabetic activity of *Centaurea* spp.: A new approach for developing herbal remedies. *Evidence-Based Complementary and Alternative Medicine*, 2021(1), p.5587938.
- Formisano, C., Sirignano, C., Rigano, D., Chianese, G., Zengin, G., Seo, E.J., Efferth, T. and Taglialatela-Scafati, O., **2017**. Antiproliferative activity against leukemia cells of sesquiterpene lactones from the Turkish endemic plant *Centaurea drabifolia* subsp. *detonsa*. *Fitoterapia*, 120, pp.98-102.
- Guo, R., Lv, T.M., Shang, X.Y., Yao, G.D., Lin, B., Wang, X.B., Huang, X.X. and Song, S.J., **2019**. Racemic neolignans from *Crataegus pinnatifida*: Chiral resolution, configurational assignment, and cytotoxic activities against human hepatoma cells. *Fitoterapia*, 137, p.104287.
- Gutiérrez, R.M.P. and Ramirez, A.M., **2016**. Hypoglycemic effects of sesquiterpene lactones from *Byrsonima crassifolia*. *Food science and biotechnology*, 25, pp.1135-1145.
- Huang, S.Z., Ma, Q.Y., Wang, Q., Dai, H.F., Liu, Y.Q., Zhou, J. and Zhao, Y.X., **2018**. Daphnauranins C–E, three new antifeedants from *Daphne aurantiaca* roots. *Molecules*, 23(10), p.2429.
- Li, X.L., Qian, P.L., Liu, Z.Y., Zhao, Y., Xu, G., Tao, D.D., Zhao, Q.S. and Sun, H.D., **2005**. Sesquiterpenoids from *Cynara scolymus*. *Heterocycles*, 65(2), pp. 287 – 291.

- Macías, F.A., Viñolo, V.M., Fronczek, F.R., Massanet, G.M. and Molinillo, J.M., **2006**. 11, 16 Oxetane lactones. Spectroscopic evidences and conformational analysis. *Tetrahedron*, 62(33), pp.7747-7755.
- Oksuz, S., Ulubelen, A., Aynechi, Y. and Wagner, H., **1982**. A guaianolide from *Centaurea behen*. *Phytochemistry*, 21(11), pp.2747-2749.
- Perez, R.M., Perez, C., Zavala, M.A., Perez, S., Hernandez, H. and Lagunes, F., **2000**. Hypoglycemic effects of lactucin-8-O-methylacrylate of *Parmentiera edulis* fruit. *Journal of Ethnopharmacology*, 71(3), pp.391-394.
- Rath, J.P., Kinast, S. and Maier, M.E., **2005**. Synthesis of the fully functionalized core structure of the antibiotic abyssomicin C. *Organic Letters*, 7(14), pp.3089-3092.
- Rong, L.I.U., Kun-Lung, H. and Ji-Kai, L.I.U., **2009**. A new sesquiterpene lactone from the leaves of *Cynara scolymus* (Compositae). *Plant Diversity*, 31(04), p.383.
- Shakeri, A., Masullo, M., Bottone, A., Asili, J., Emami, S.A., Piacente, S. and Iranshahi, M., **2019**. Sesquiterpene lactones from *Centaurea rhizantha* CA Meyer. *Natural product research*, 33(14), pp.2016-2023.
- Sokovic, M., Ciric, A., Glamoclija, J. and Skaltsa, H., **2017**. Biological activities of sesquiterpene lactones isolated from the genus *Centaurea* L.(Asteraceae). *Current pharmaceutical design*, 23(19), pp.2767-2786.
- Wagenitz, G., **1975**. *Centaurea* L. In: Davis, P.H. (Ed.), *Flora of Turkey and the East Aegean Islands*, 5. Edinburgh University Press, Edinburgh, pp. 465–585.
- Yang, M., Xu, X., Xie, C., Xie, Z., Huang, J. and Yang, D., **2013**. Separation and purification of arctiin, arctigenin, matairesinol, and lappaol F from *Fructus Arctii* by high-speed counter-current chromatography. *Separation Science and Technology*, 48(11), pp.1738-1744.
- Zengin, G., Guler, G.O., Cakmak, Y.S. and Aktumsek, A., **2011**. Antioxidant capacity and fatty acid profile of *Centaurea kotschyi* (Boiss. & Heldr.) Hayek var. *persica* (Boiss.) Wagenitz from Turkey. *Grasas y aceites*, 62(1), pp.90-95.

APPENDIX A

Here are reported the spectra of the previously undescribed compounds isolated from *Centaurea kotschyi* (described in PART 2 – Chapter 1 of this manuscript).

APPENDIX A	55
Figure S1-S4: NMR spectra of compound 7 (CDCl ₃ , 600MHz)	56-57
Figure S5-S8: NMR spectra of compound 8 (CDCl ₃ , 600MHz)	58-59
Figure S9-13: NMR spectra of compound 15 (MeOD, 600MHz)	60-62
Figure S14-15: NMR spectra of compound 7a (CDCl ₃ , 600MHz)	62-63

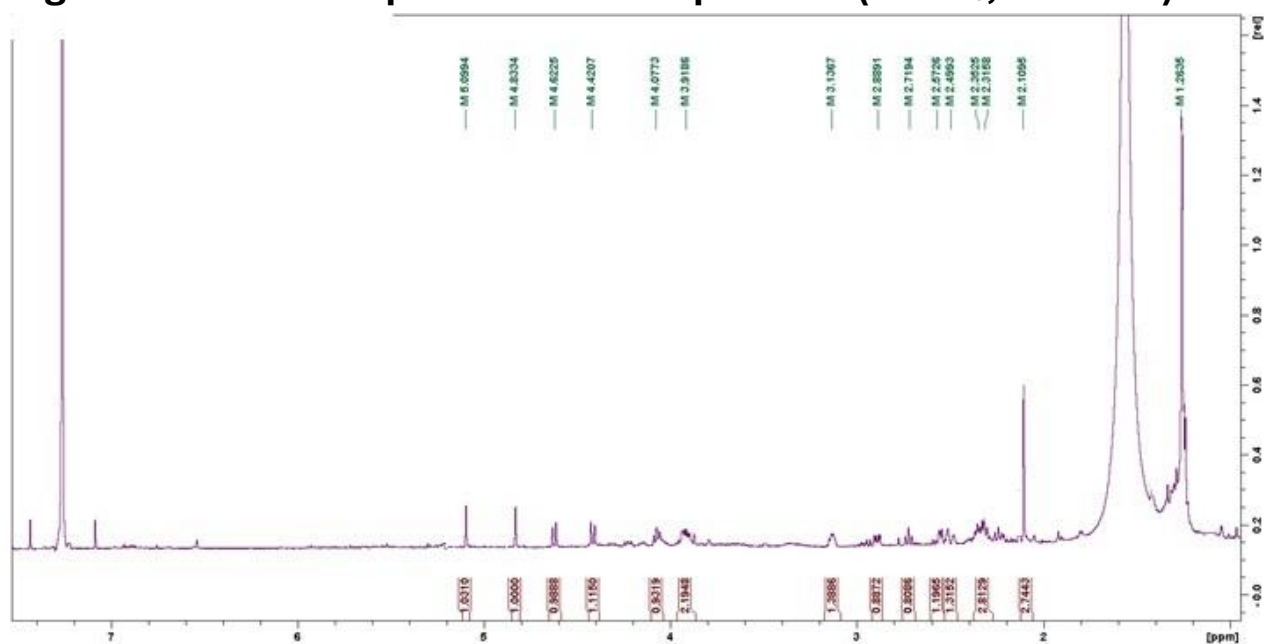
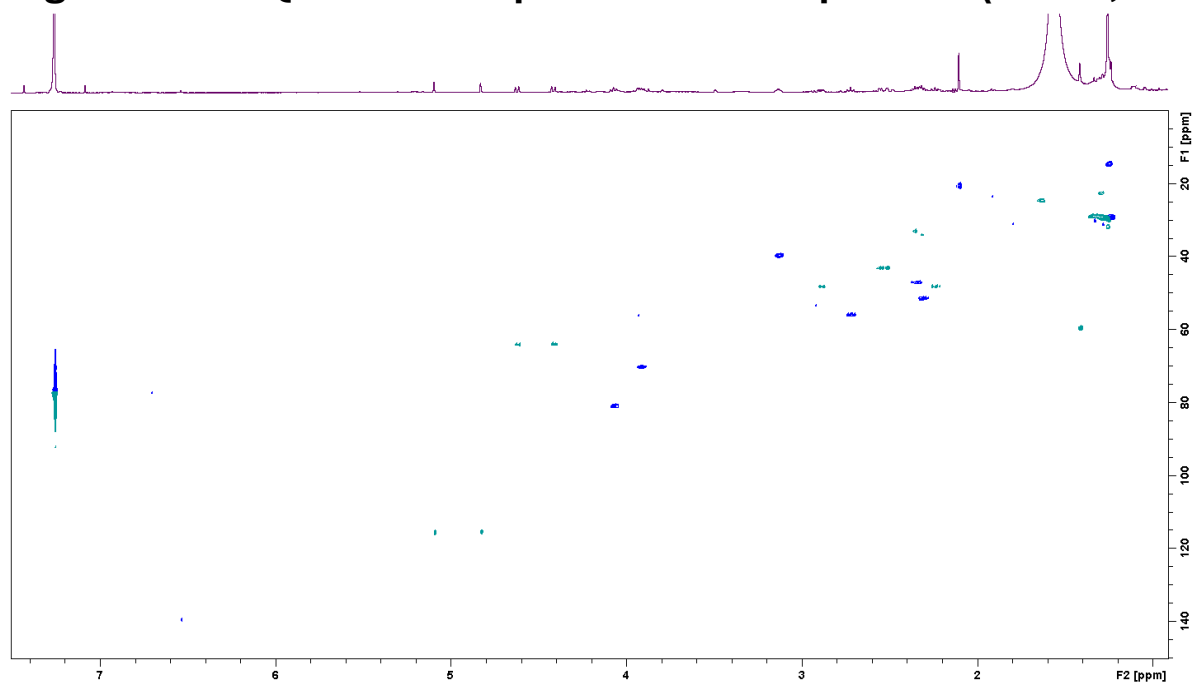
Figure S1: ^1H NMR spectrum of compound 7 (CDCl_3 , 600MHz)**Figure S2: HSQC 2D NMR spectrum of compound 7 (CDCl_3 , 600MHz)**

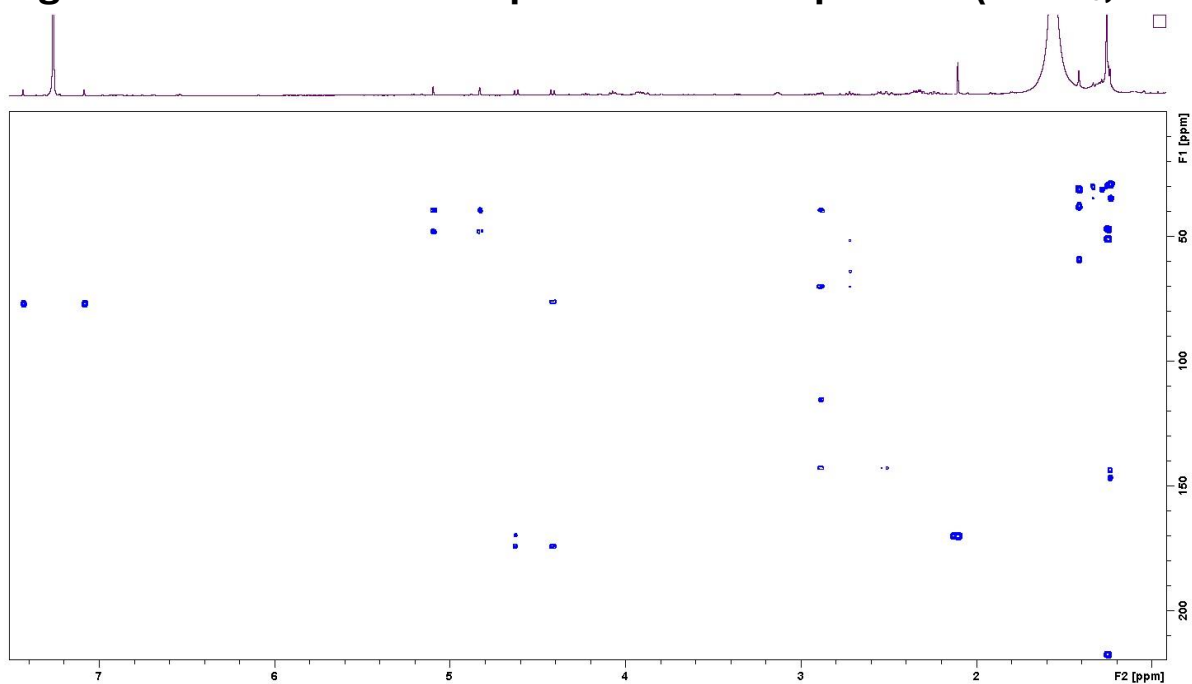
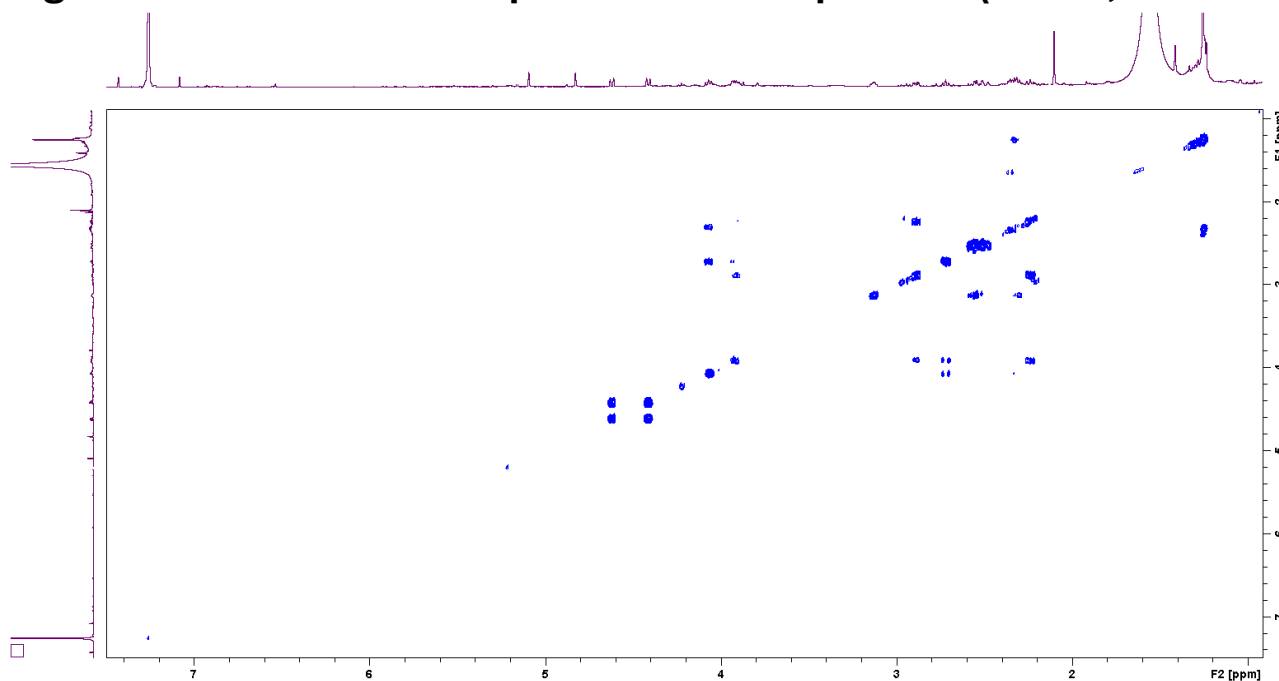
Figure S3: HMBC 2D NMR spectrum of compound 7 (CDCl₃, 600MHz)**Figure S4: COSY 2D NMR spectrum of compound 7 (CDCl₃, 600MHz)**

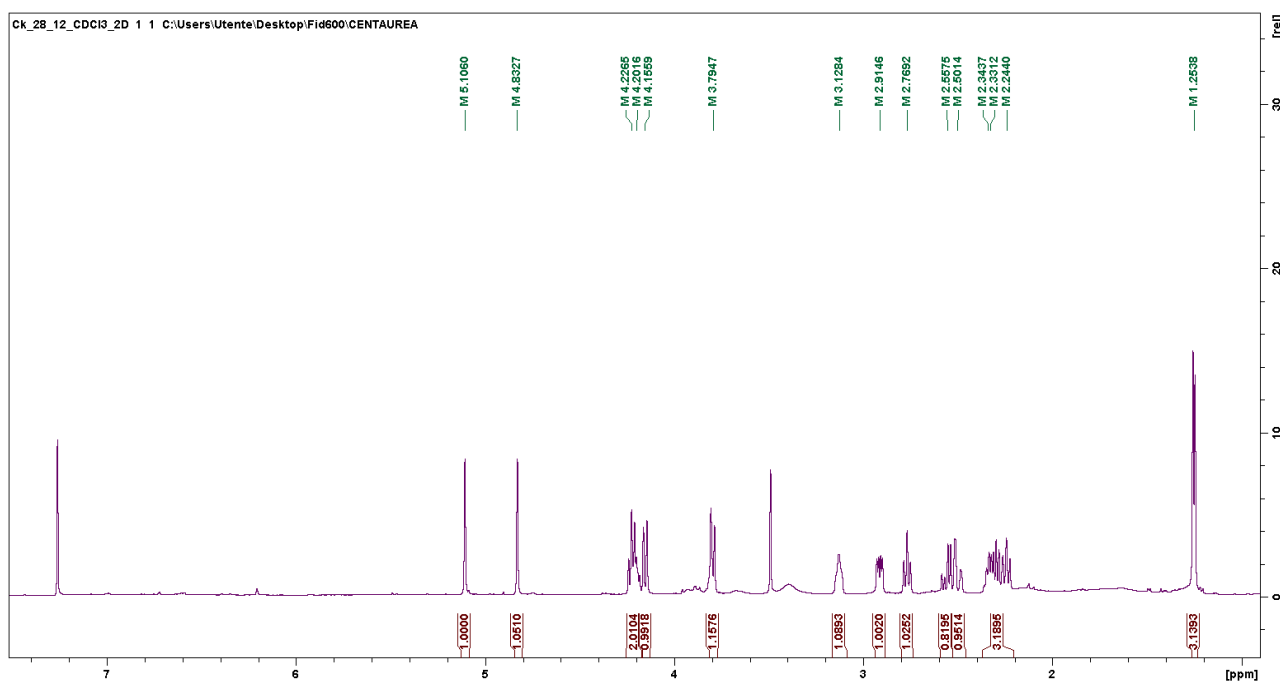
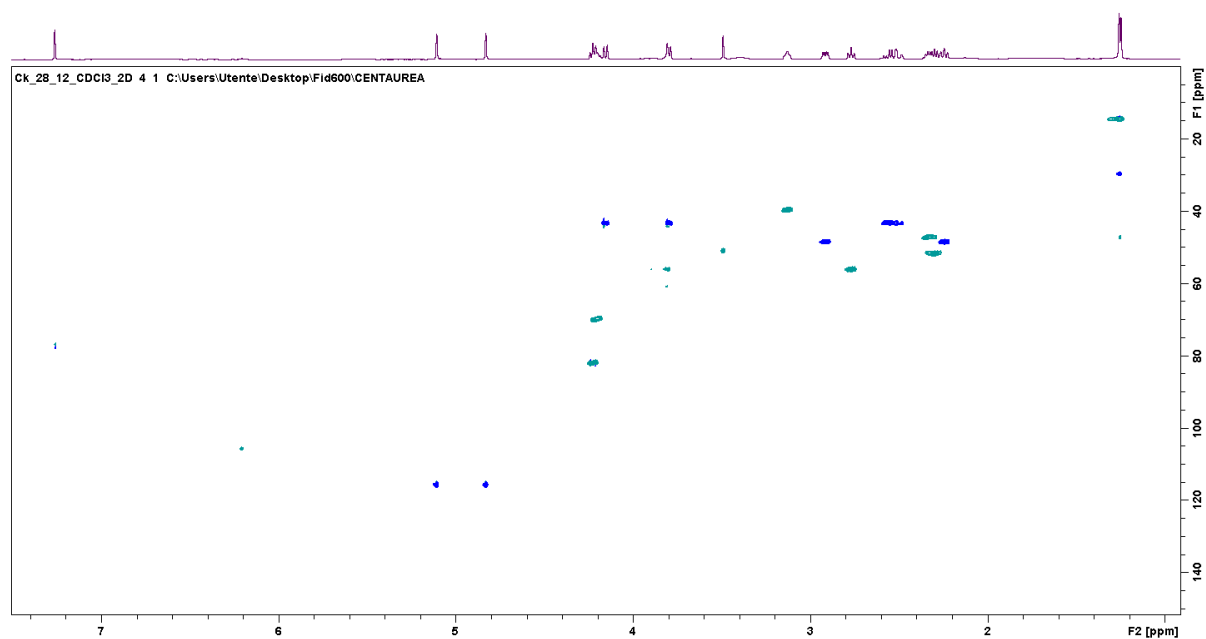
Figure S5: ^1H NMR spectrum of compound 8 (CDCl_3 , 600MHz)**Figure S6: HSQC 2D NMR spectrum of compound 8 (CDCl_3 , 600MHz)**

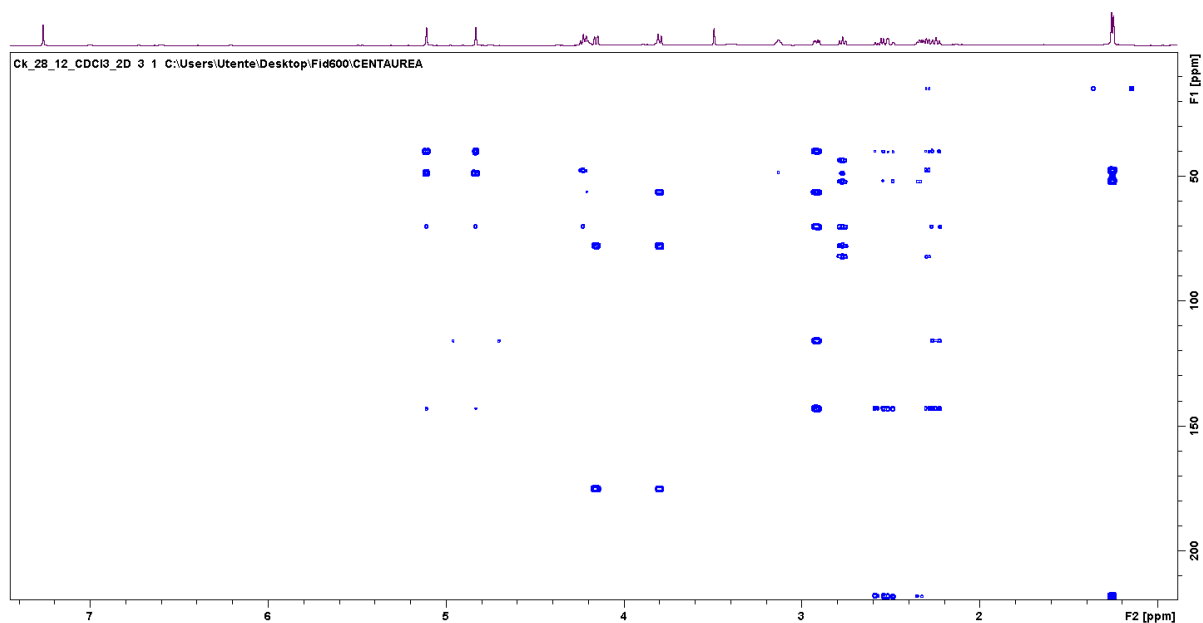
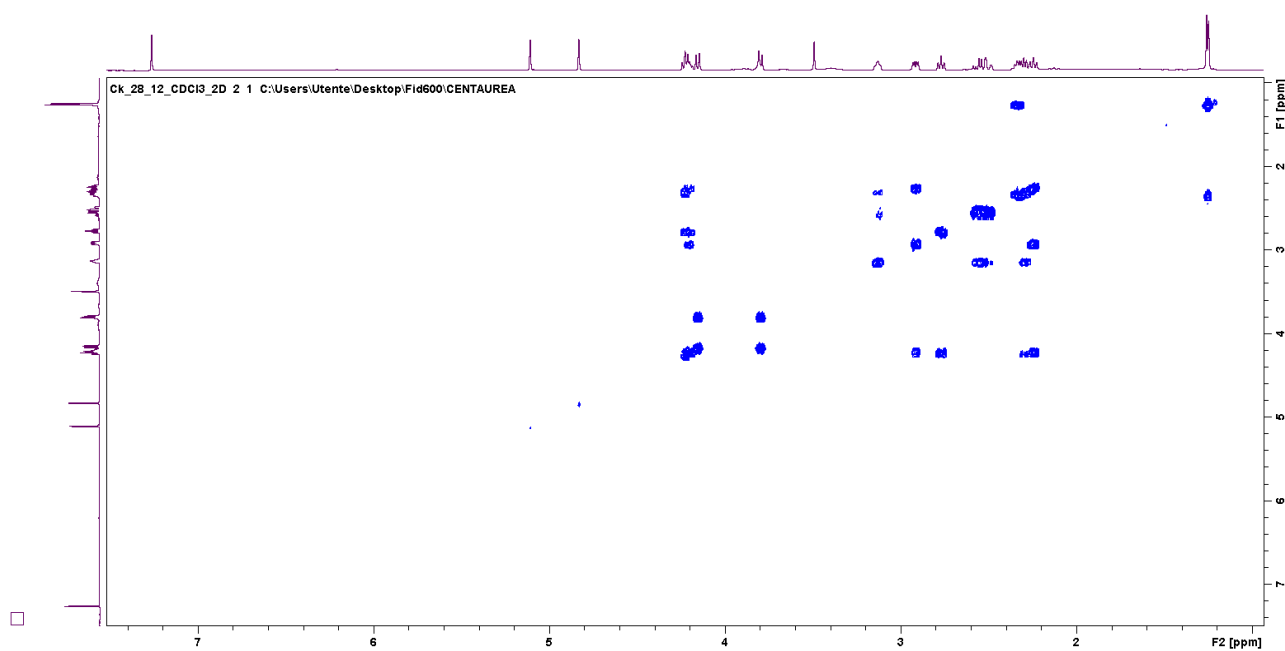
Figure S7: HMBC 2D NMR spectrum of compound 8 (CDCl₃, 600MHz)**Figure S8: COSY 2D NMR spectrum of compound 8 (CDCl₃, 600MHz)**

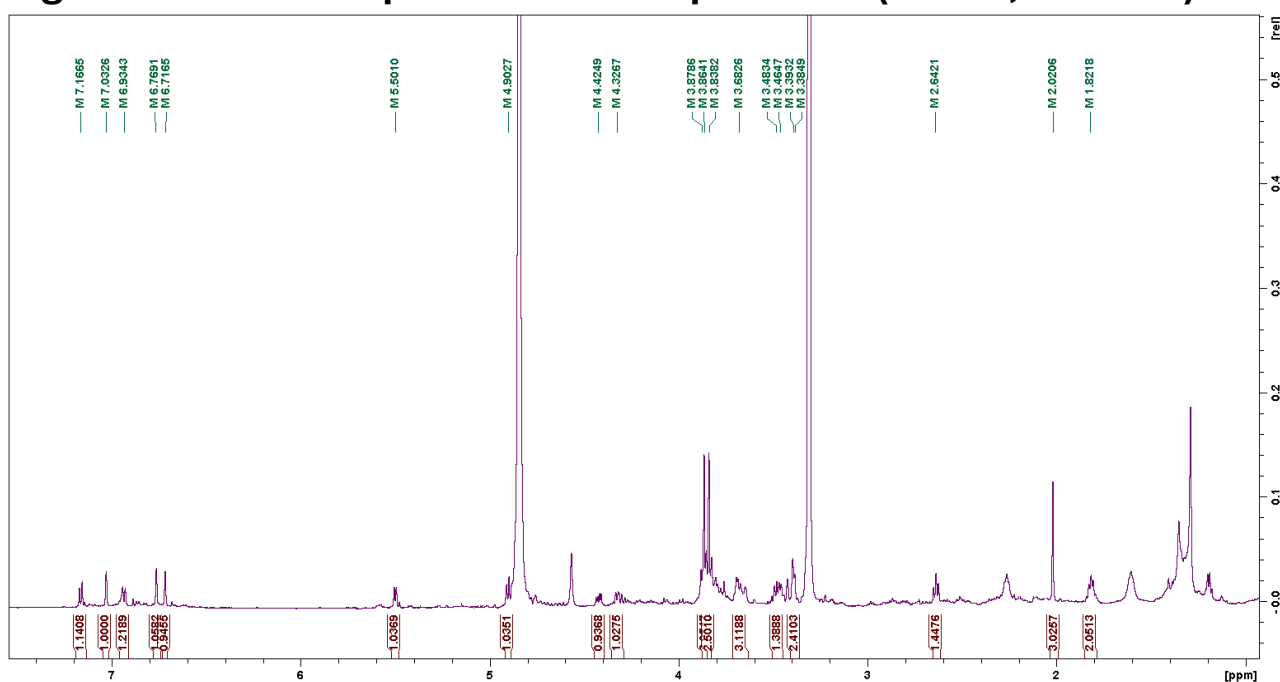
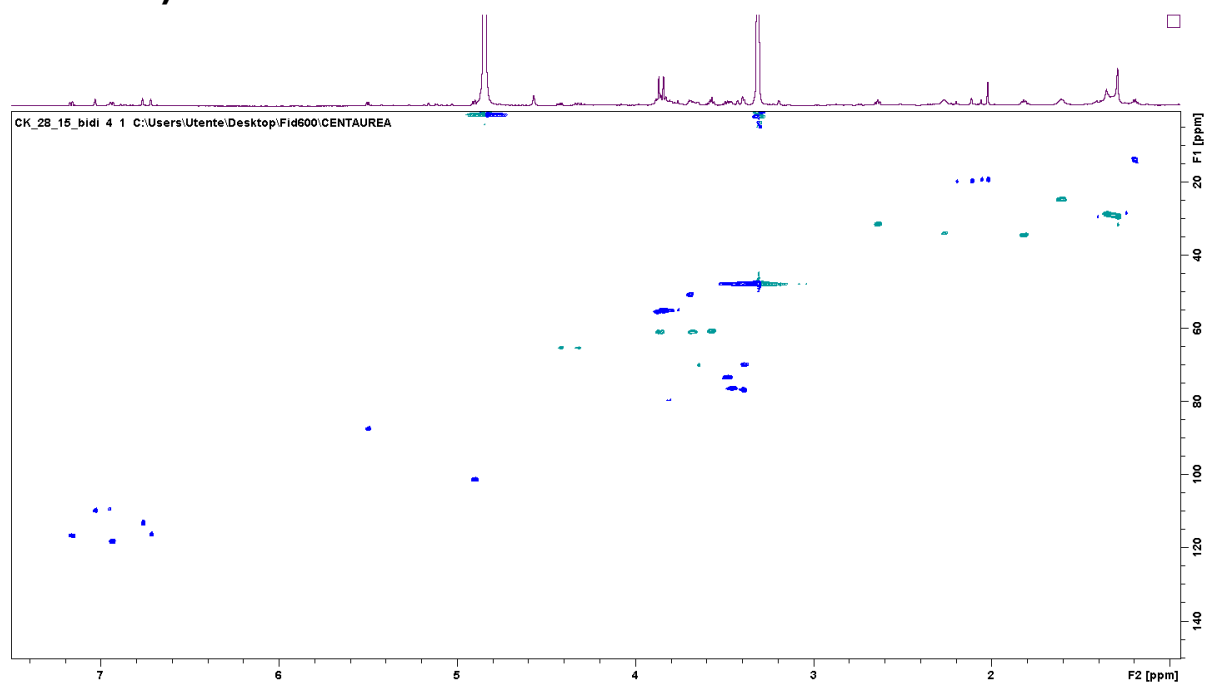
Figure S9: ^1H NMR spectrum of compound 15 (MeOD, 600MHz)**Figure S10: HSQC 2D NMR spectrum of compound 15 (MeOD, 600MHz)**

Figure S11: HMBC 2D NMR spectrum of compound 15 (MeOD, 600MHz)

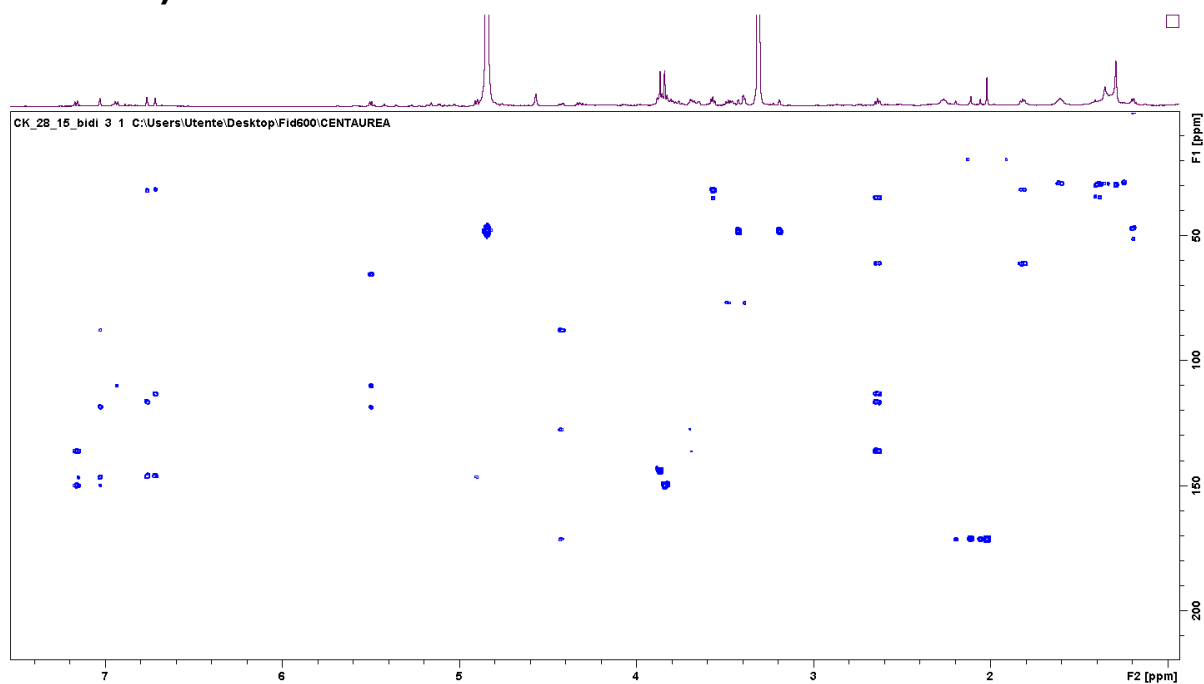


Figure S12: COSY 2D NMR spectrum of compound 15 (MeOD, 600MHz)

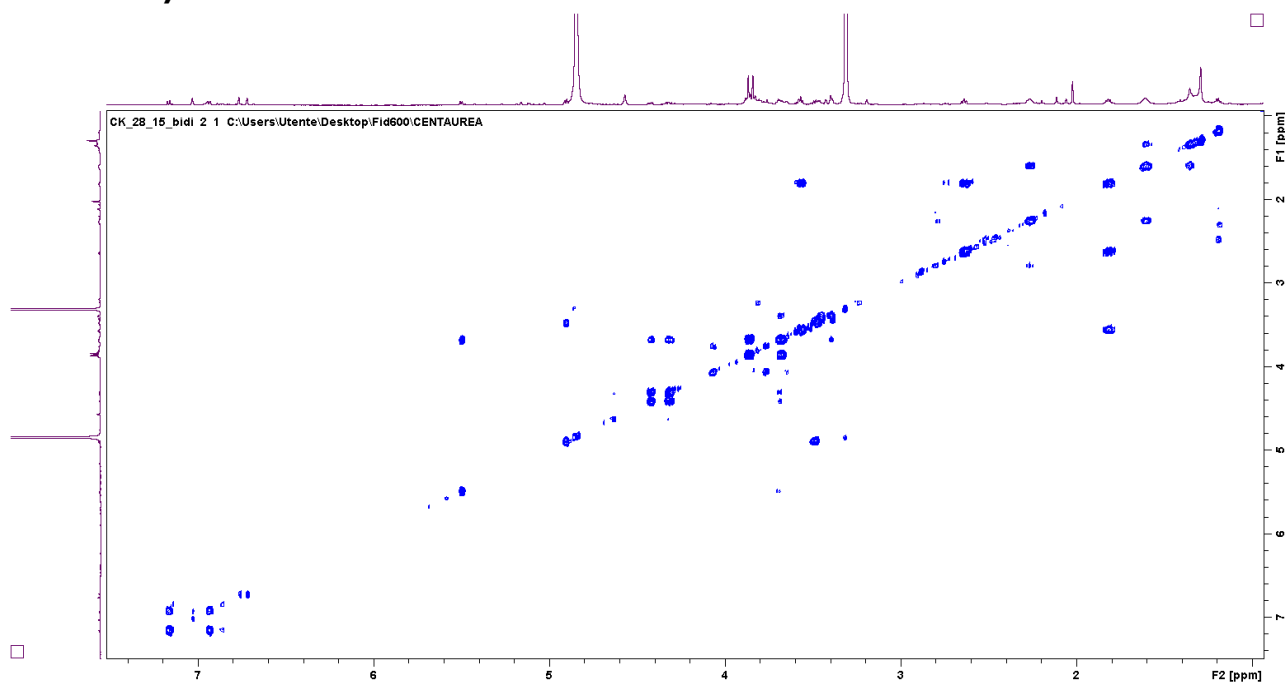


Figure S13: NOESY 2D NMR spectrum of compound 15 (MeOD, 600MHz)

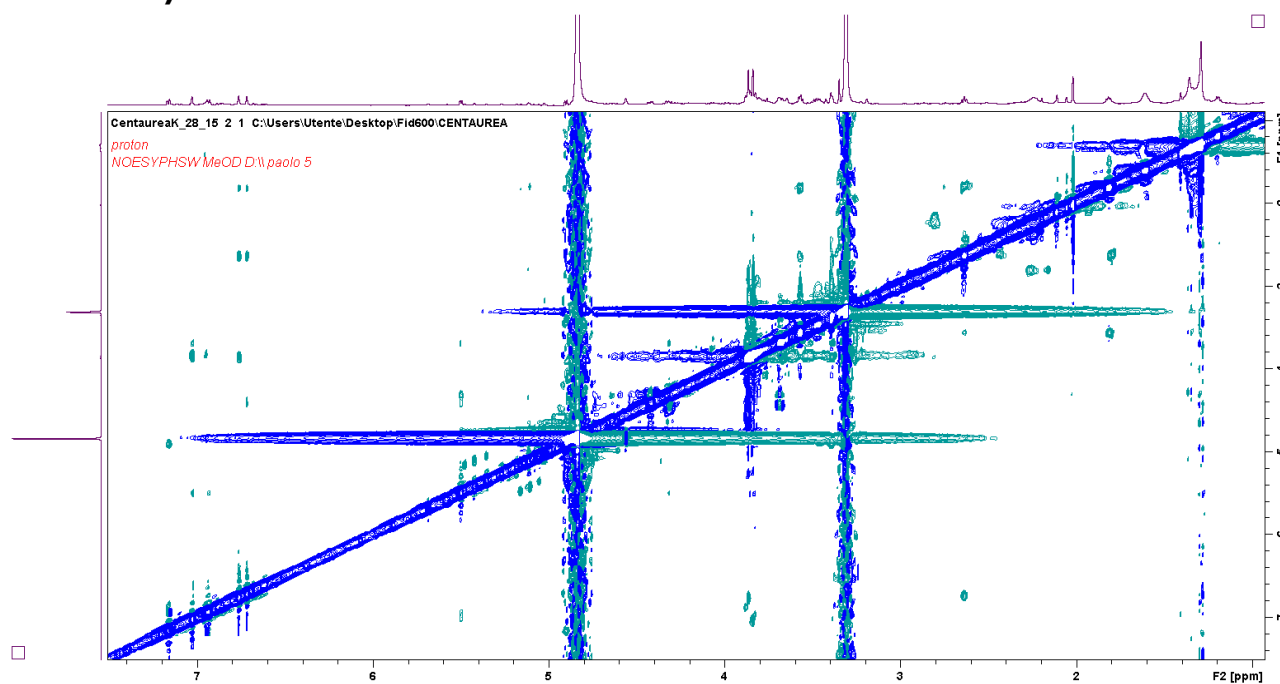


Figure S14: ^1H NMR spectrum of compound 7a (CDCl_3 , 600MHz)

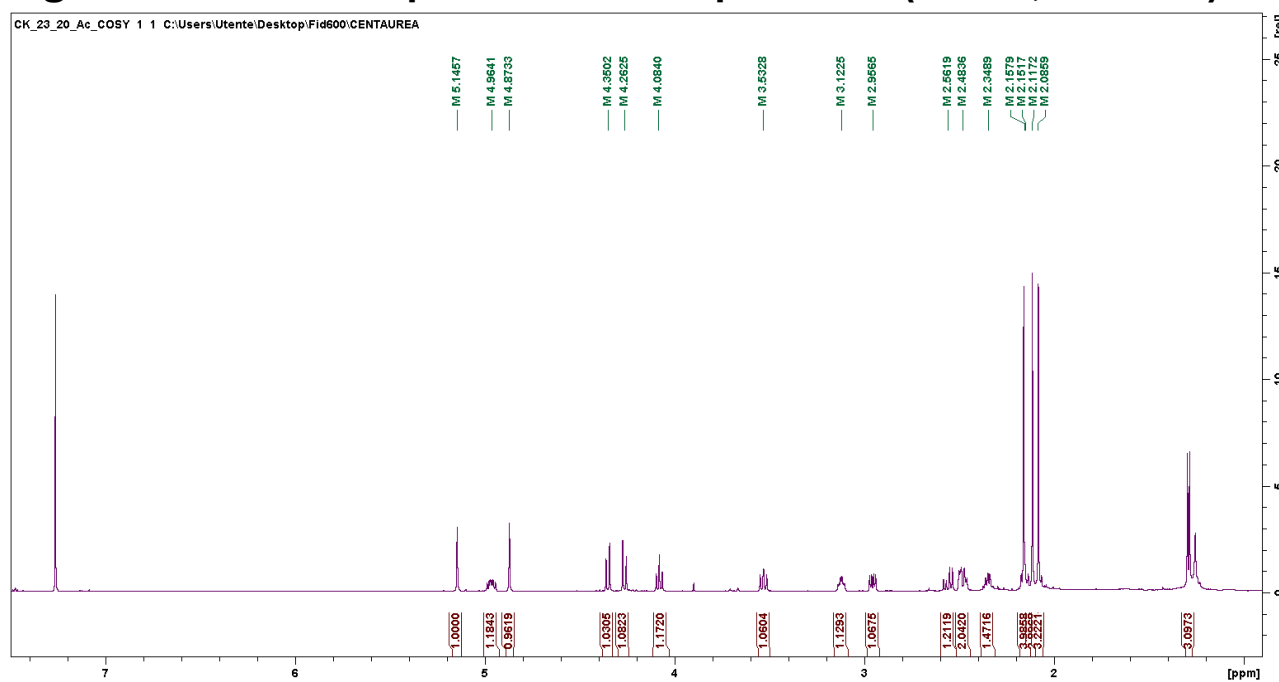
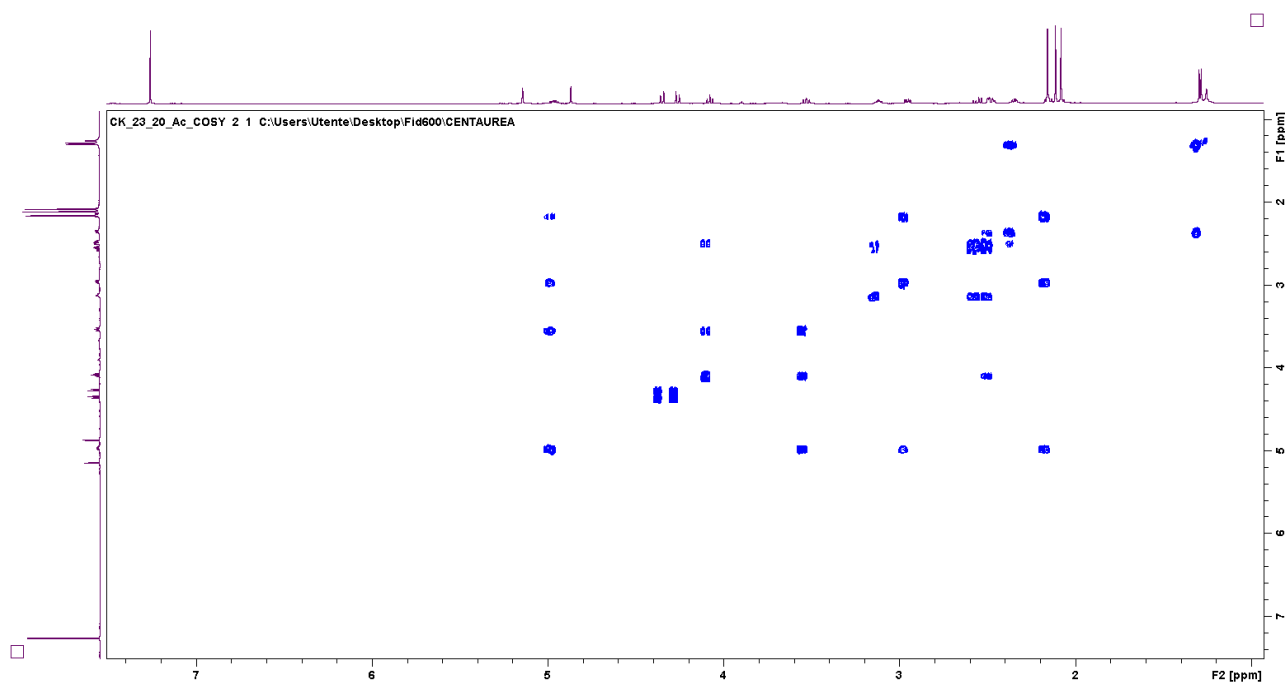


Figure S15: COSY 2D NMR spectrum of compound 7a (CDCl₃, 600MHz)



CHAPTER 3: *Centaurea sicula*

Centaurea sicula L. is a Mediterranean species of genus *Centaurea* that grows in Spain, Southern Italy, Malta, Tunisia, and other countries of Northwestern Africa. It takes name from Sicily, where this species is widespread in the uncultivated lands. Especially in southern Italy, this plant has been used as a food vegetable (after cooking) and it also has a history as a natural remedy against several ailments (Giardinia et al., 2007). *C. sicula* belongs to the section *Seridia* of the *Centaurea* genus and, since all the species belonging to the same section appeared to be devoid of guaianolides (Bruno et al., 2013), we expected a different phytochemical profile than *C. kotschyi* subsp. *perisca*. Indeed, the species in the *Seridia* section seemed to be rich in germacranes (Barrero et al., 1997). *C. sicula* has been previously investigated, revealing the presence of flavonoids (Akkal et al., 2003), acetylenes (Christensen et al., 1990), sesquiterpenes lactones (Medjroubi et al., 2003), and also its essential oil composition was reported (Senatore et al., 2008). Lately, the antioxidant effect of the plant extracts of some Tunisian samples have been studied (Dhouibi et al., 2020). Driven by the antidiabetic potential of *Centaurea* plants, we have conducted a detailed analysis of the apolar extract of the aerial parts of a Sicilian sample of *C. sicula* and, in the frame of my research activity in finding unreported natural leads with activity on metabolic syndrome, all the isolated metabolites have been evaluated for their glucose uptake stimulating activity.

3.1 Phytochemical characterization of the aerial parts

Dried aerial parts of *C. sicula* were defatted with *n*-hexane and then exhaustively extracted with CH₂Cl₂. This latter phase was firstly fractionated on MPLC-DAD, and successively the obtained fractions were further purified on RP-HPLC yielding to 24 pure compounds. Compounds **1-14** belong to terpenoid family, **15-23** are polyphenols and compound **24** has been characterized for the first time as a novel alkaloid. The major class of secondary metabolites in the plant was sesquiterpene lactones that included three different structural families: elemanolides (**1-6**), eudesmanolides (**7** and **8**), and germacranolides (**9-11**) with four previously undescribed secondary metabolites (**5-8**). As we expected, *C. sicula* differs from *C. kotschyi* subsp. *persica*,

that only biosynthesize guaianolides (Serino et al., 2022) and showed a phytochemical composition similar to two other species of *Seridia* section, *C. aspera* ssp. *subinermis* (Cardona et al., 1992) and *C. malacitana* (Barrero et al., 1997), that exhibit the co-occurrence of germacranes, elemanes, and eudesmanes. The known compounds were identified through comparison with their ^1H NMR and HR-ESIMS data with those reported in the literature. The known sesquiterpenoids were identified as the elemanes melitensin (**1**) (Medjroubi et al., 2003), dehydromelitensin (**2**) (Sen et al., 2017), compound **3** (Bruno et al., 1996), compound **4** (Bruno et al., 1988), and the germacranolides 11 β ,13-dihydrosalonitenolide (**9**) (Marco et al., 2005), dihydrocnicin (**10**) (Bicha et al., 2013), and 11-*epi*-dihydroartemisiifolin (**11**) (Porter, 1970). The monoterpenoids daphnauranin E (**12**) (Huang et al., 2018) and isololiolide (**13**) (Kimura et al., 2002) along with abscisic acid (**14**) (Arnone et al., 1990) were also isolated. The structures of the terpenoids are reported in **Figure 3.1**, Chart A. The non-terpenoid compounds included the lignans (-)-salicifoliol (**15**) (González et al., 1989), never reported before for the genus *Centaurea*, pinoresinol (**16**) (Mohamed et al., 2021), and arctigenin (**17**) (Zater et al., 2016) and the flavonoid jaceosidin (**18**) (Aamer et al., 2021). Other purified compounds were the 6'-*trans*-2- butenoyl ester of the cyanogenic glycoside prunasin (**19**) (Cardona et al., 1992), vanillin (**20**) (Cardona et al., 1992), syringaldehyde (**21**) (Jimeno et al., 2004), β -hydroxypropiovanillone (**22**) (Higuchi et al., 2018), and the coumarin scopoletin (**23**) (Higuchi et al., 2018). The structures of these compounds are reported in **Figure 3.1**, Chart B.

PART 2

CHART A

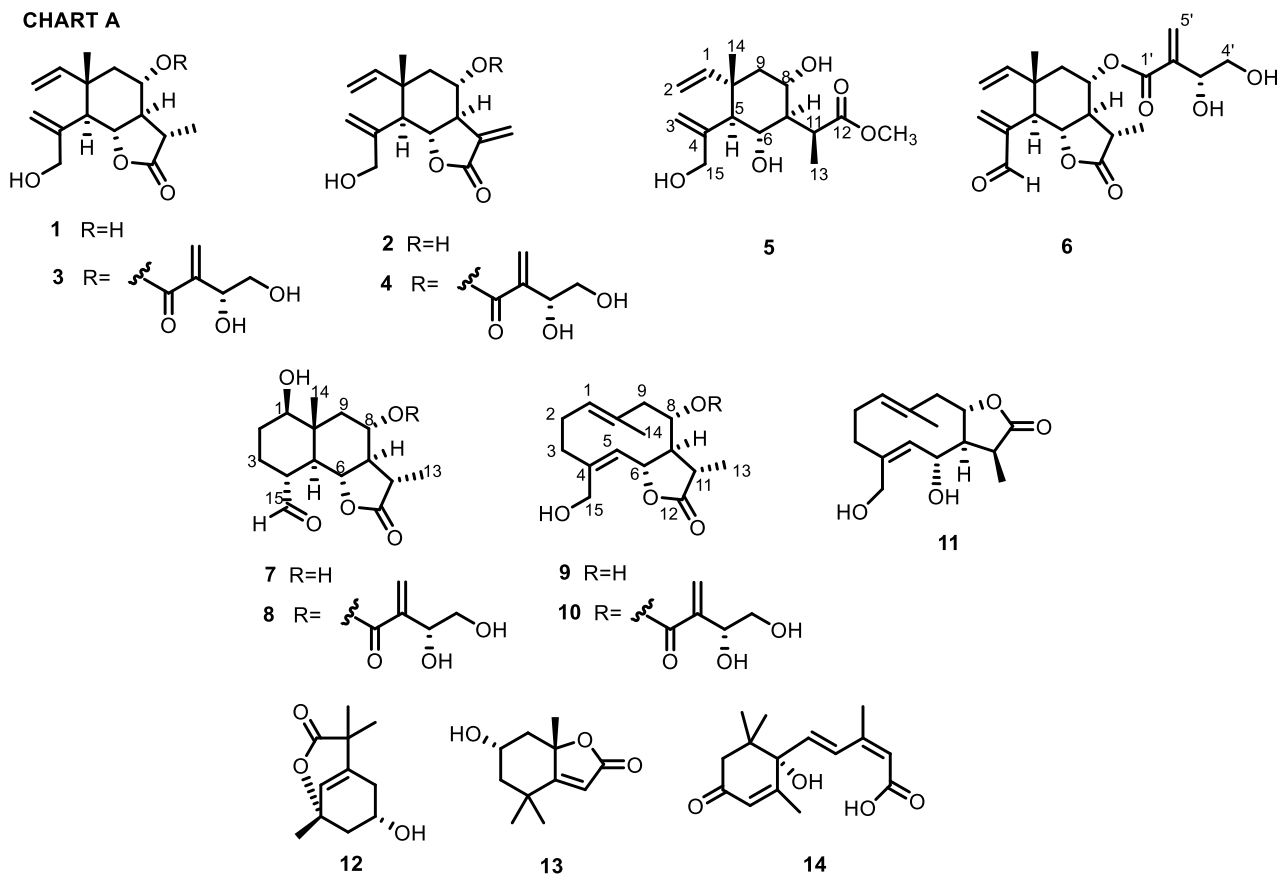


CHART B

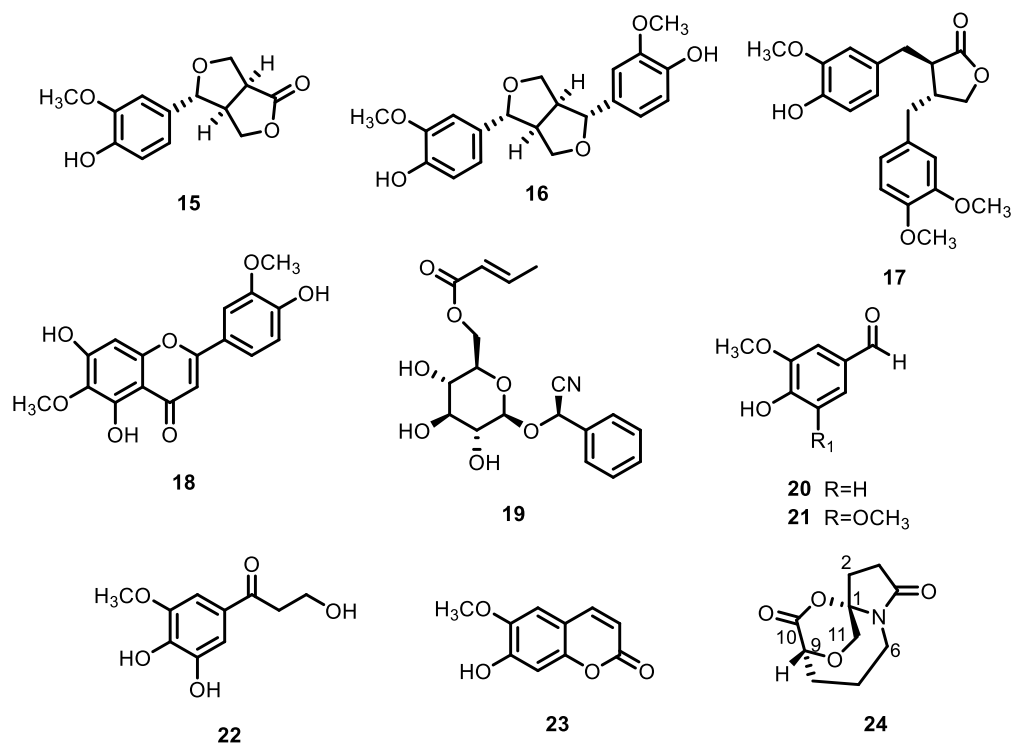


Figure 3. 1 Sesquiterpenes (Chart A) and polyphenols (Chart B) isolated from *C. sicula* aerial parts.

Compound **5** showed a HR-ESIMS peak at m/z 321.1684, indicating $C_{16}H_{26}O_5$ as molecular formula ($C_{16}H_{26}O_5Na$ requires m/z 321.1678) and suggesting the presence of four unsaturation degrees. The analysis of the ^1H NMR spectrum of **5** showed the typical low-field signals at δ_H 5.67 (dd), 4.94 (d), 4.91 (d), 5.01 (s), and 5.42 (s), typical of an elemene skeleton. The two methyl signals, resonating at δ_H 1.02 (s) and δ_H 1.25 (d), were attributable to H_3-14 and H_3-13 , respectively, while the coupled doublets at δ_H 4.09 and 3.97 could be assigned to the oxymethylene group at C-15. Moreover, a methoxy group resonating at δ_H 3.71 (s) was present in the ^1H NMR spectrum of **5**. Detailed analysis of the COSY spectrum of **5** allowed the identification of the spin systems, while the HSQC was functional to the assignment of resonances to protonated carbons. The HMBC spectrum allowed the connection of the subunits, indicating a structure very similar to that of melitensin (**1**), but the marked change in the signal of H-6 (δ_H 3.66; δ_C 68.7) and the HMBC correlation of the methoxy protons with C-12 confirmed the opening of the lactone ring, replaced by methyl esterification at C-12. A similar methyl ester derivative has been found in *C. aspera* var. *subinermis* (Cardona et al., 1992), but it was structurally related to dehydromelitensin (**2**). To determine the configurational relationship between previously undescribed compound **5** and melitensin (**1**), we decided to carry out a basic methanolysis on melitensin (**1**) (**Figure 3.2**). The spectral data of the obtained product, including optical activity, were identical with those of **5**, implying that the two compounds share the same absolute configuration.

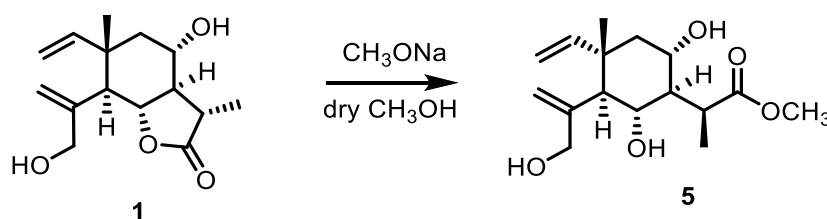


Figure 3. 2 Basic methanolysis of melitensin (**1**) affording compound **5**

Compound **6** was obtained as a white, amorphous solid with a HR-ESIMS peak at m/z 401.1584 [$M+Na$] $^+$ indicating a molecular formula of $C_{20}H_{26}O_7$ ($C_{20}H_{26}O_7Na$ requires m/z 401.1576). From an analysis of the ^1H NMR spectrum, compound **6** appeared to be closely related to compound **3**, co-occurring in this plant. The most significant difference was the absence of the coupled

doublets H-15a/H-15b, replaced by an aldehyde broad singlet resonating at δ_{H} 9.46. An oxidation at C-15 of **6** was in agreement with the molecular formula and strongly supported by the marked downfield shift of H₂-3, now part of a conjugated system and resonating at δ_{H} 6.27 and 6.28. Having associated all the proton signals with those of the directly linked carbon signals via the HSQC spectrum, we could use the HMBC correlations to build up the structure of **6**. Thus, the HMBC correlation of the aldehyde signal with the unprotonated carbon at C-4 (δ_{C} 145.0) and with the methine at C-5 (δ_{C} 46.1) unambiguously placed the aldehyde at position **15**, while the diagnostic HMBC crosspeak between the H-8 (at δ_{H} 5.23) and C1' (δ_{C} 165.0) confirmed the esterification position. The pattern of spatial contacts detected with a NOESY spectrum closely paralleled that of compound **3**, and, for example, correlations of H-8 with CH₃-14 and H-6 indicated the relative configuration around the elemene skeleton. The absolute configuration of **3**, previously isolated from *Centaurea nicaensis*, has been reported in the literature (Bruno et al., 1996), and we can assume that compound **6** shares the same absolute configurational arrangement. NMR data of **6** are reported in the Experimental Section.

Compound **7** showed an HR-ESIMS peak at m/z 305.1355 [$\text{M} + \text{Na}$]⁺ implying a molecular formula of C₁₅H₂₂O₅ (C₁₅H₂₂O₅Na requires m/z 305.1365). The ¹H NMR spectrum of **7** (see Experimental section) showed two methyl signals at δ_{H} 0.98 (s) and 1.37 (d), an aldehyde group resonating at δ_{H} 9.58, and three oxymethines resonating at δ_{H} 3.43, 3.90, and 3.98. All of the proton multiplets could be arranged in a single spin system (see **Figure 3.3**) through the 2D COSY spectrum, while the HMBC cross-peaks exhibited by H₃-14 and H-6 were instrumental to the definition of the planar structure of **7**, belonging to the eudesmanolide family. The NOESY crosspeaks of H₃-14 with H-4 and H-6 placed the aldehyde moiety in the α -orientation on the eudesmane skeleton. Moreover, the NOE correlation H-7/CH₃-13 disclosed the α -orientation of the methyl group on the lactone ring, while the β -orientation of the hydroxy group at C-1 was deduced through the NOESY cross-peak H-1/H-5. An eudesmane sesquiterpenoid sharing the same planar structure but a different orientation of the aldehyde group has been reported from *Onopordon laconicum* (Lazari et al., 1998). Given their evident biogenetic relationships, we have assigned the absolute configuration to **7** as that of the cooccurring elemenes and germacranes.

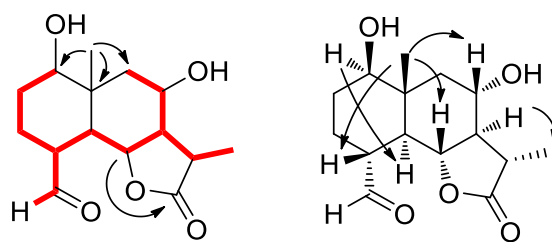


Figure 3.3 Left: COSY (red bold) and diagnostic HMBC (black arrow) correlations of **7**. Right: NOESY correlations of **7**.

Compound **8** HRESI-MS analysis revealed an ion peak at m/z 419.1692 $[M+Na]^+$, suggesting $C_{20}H_{28}O_8$ as the molecular formula ($C_{20}H_{28}O_8Na$ requires m/z 419.1682). Its ^1H NMR spectrum exhibited significant similarities to that of **7**, with the exception of the presence of an acyl moiety, easily assignable as being identical to that present in compounds **3-5**. This was confirmed by combined inspection of 2D NMR spectra. Having built up spin systems and assigned carbon resonances with COSY and HSQC experiments, respectively, the H-8/C-1' HMBC correlation placed the α,β -unsaturated acyl chain at position 8, while the HMBC correlations of H-6 and H₃-13 with C-12 confirmed the structure of the lactone moiety. NOESY correlations of **8** exactly paralleled those shown by **7**, and therefore the two compounds must share the same relative configuration. In summary, compound **8** is a derivative of **7** decorated with a lateral side chain identical with those of other sesquiterpenes found in this plant. The NMR data of **8** are reported in the Experimental Section.

Compound **11** spectral data, including HR-ESIMS (m/z found 289.1407, $C_{20}H_{28}O_8Na$ requires m/z 289.1416) and NMR, appeared in good agreement with those available for 11,13-dihydroartemisiifolin, a germacrane sesquiterpenoid isolated from *Ambrosia artemisiifolia* (Porter et al., 1970). However, some slight differences in the chemical shifts, likely due to configurational change(s), were evident. The diagnostic NOESY correlations between H-6/H-8, H-8/H₃-13, and H-7/H-11 were instrumental to confirm this hypothesis and assign the structure of compound **11** as the epimer at C-11 of 11,13-dihydroartemisiifolin. The presence of an epimer of dihydroartemisiifolin in *Centaurea* species had been postulated in other papers (González et al., 1973), but, to our knowledge, this configurational issue had never been clarified.

Siculamide (**24**) showed a HR-ESIMS peak at m/z 234.0741 $[M+Na]^+$ indicated for this compound the molecular formula $C_{10}H_{13}NO_4$ ($C_{10}H_{13}O_4NNa$ requires m/z 234.0742), implying five degrees of unsaturation. Its ^1H NMR spectrum (see the Experimental Section) displayed a triplet at δ_H 5.18, two mutually coupled doublets at δ_H 4.16 and 4.04, two ^1H multiplets at δ_H 3.75 and 3.41 and eight additional well-resolved multiplets located between δ_H 1.85 and 2.95. The 2D HSQC experiment organized the 13 proton signals as six methylenes and one methine, thus revealing that no heteroatom-linking proton was present in the molecule. Analysis of the 2D NMR COSY spectrum of **24** arranged these multiplets within two spin systems and an isolated (not further coupled) oxymethylene (δ_C 73.3). Correlations present in the HMBC spectrum revealed that three remaining carbon atoms of the molecular formula were assignable as two carbonyl groups, resonating at δ_C 165.8 and 175.9, and an unprotonated carbon at δ_C 80.0, likely linked to heteroatom(s). A detailed analysis of the HMBC correlations guided the assembly of a complex tricyclic architecture for siculamide (**24**) (see **Figure 3.4**). The 2-pyrrolidinone ring was built on the basis of the HMBC correlations of both H_{2-3} and H_{2-6} (δ_H 3.75 and 3.41) with C-4 and unprotonated C-1. Correlations of the isolated oxymethylene H_{2-11} with C-9, C-1, and C-2 implied its attachment at the unprotonated carbon C-1. Finally, correlations of H_{2-8} and H-9 with the carbonyl C-10 completed the structure of the molecule that, in agreement with the molecular formula, should include lactonization at C-1. The structure of siculamide includes two stereogenic centers whose relative orientation was determined based on the detected NOE contact between H-9 and H-11a. Attempts to experimentally determine the absolute configuration of the compound ($[\alpha]_D = -7.5$) were frustrated by the lack of easily derivatizable functional groups and by the almost flat CD spectrum. The structural architecture of siculamide (**24**) is unique, and no related compounds have been reported from natural sources. A certain structural similarity can be found with *Stemona* alkaloids (Wang et al., 2004), many of which include an *N*-alkylated 2- pyrrolidinone moiety. The other alkaloids reported from *Centaurea* species belong to the class of indole derivatives, such as moschamine (Sarker et al., 1997) or schischkiniin (Shoeb et al., 2005).

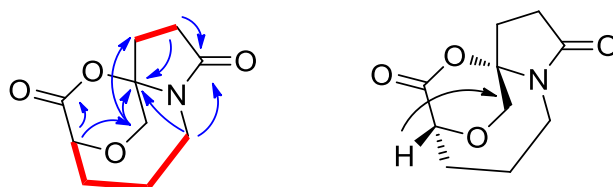


Figure 3. 4 Left: COSY (red bold) and key HMBC correlations (arrows) of compound **24**. Right: Crucial NOESY correlation of **24**.

The unprecedented skeleton of sicularamide could be originated from the two products of transamination of glutamic acid and ornithine (**Figure 3.5**). Formation of an amide bond between 2-oxoglutarate and 5-amino-2-oxovaleric acid should result in the formation of the carbon skeleton of **24**. Then, probably simultaneous formation of the pyrrolidinone ring and of an ester bond give rise to the formation of two cycles. Finally, reduction of the ketone and carboxylic groups and a further ring closure should afford sicularamide (**24**). Noticeably, the two amino acids originating this molecule have the same carbon skeleton and can be, in turn, interconverted by action of the enzyme ornithine aminotransferase, overexpressed in plants in stress conditions.

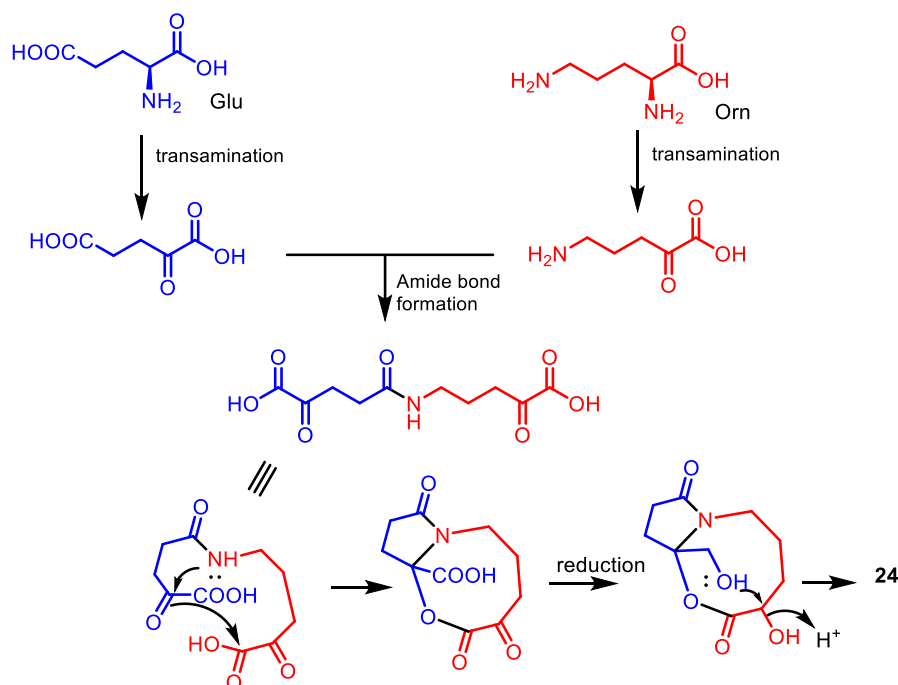


Figure 3. 5 Postulated biogenesis for sicularamide (**24**)

3.2 Bioactivity of the secondary metabolites

Inspired by the traditional use of the plant and in order to investigate how the different sesquiterpene skeletons influence the hypoglycemic activity showed by *C. kotschyi* subsp. *perisca* guaianolides, we decided to test the isolated pure compounds to evaluate their glucose uptake stimulation. To monitor the effects exerted by pure metabolites isolated from *C. sicula* in stimulating cellular glucose uptake, we used mouse embryonic HEK293 cells. The different compounds were tested using the NBDG uptake fluorescence assay, where a fluorescent probe is used to monitor glucose transport through GLUT and sodium-glucose co-transporters (SGLTs). Data reported in **Figure 3.6** clearly indicate that among the isolated compounds **10-12** and **15** exerted a marked effect on glucose uptake. As shown in **Figure 3.7**, the activity of compounds **11** and **15** is inhibited by EDTA (a Na⁺ chelator), suggesting that these compounds promote glucose uptake by modulating the activity of SGLT transporters. Differently, compounds **10** and **12** present an activity not affected by EDTA, suggesting they act by modulating glucose uptake via GLUT transporters. Remarkably, among sesquiterpenoids, only the germacranolides dihydrocnicin (**10**) and 11-*epidihydroartemisiifolin* (**11**) proved to be active. This clearly demonstrates the impact on this pharmacological activity of the carbon skeleton with structural details also being important, as demonstrated by the higher activity of **12** compared to the closely structurally related **13**. Among the phenol-containing compounds, the activity of the lignan salicifoliol (**15**) is remarkable, and, also in this case, the inactivity of the related pinoresinol (**17**) indicates the existence of strict structural requirements that are well worth being investigated in detail. Thus, we have disclosed a remarkable and previously undescribed activity for dihydrocnicin (**10**) and the lignan salicifoliol (**15**).

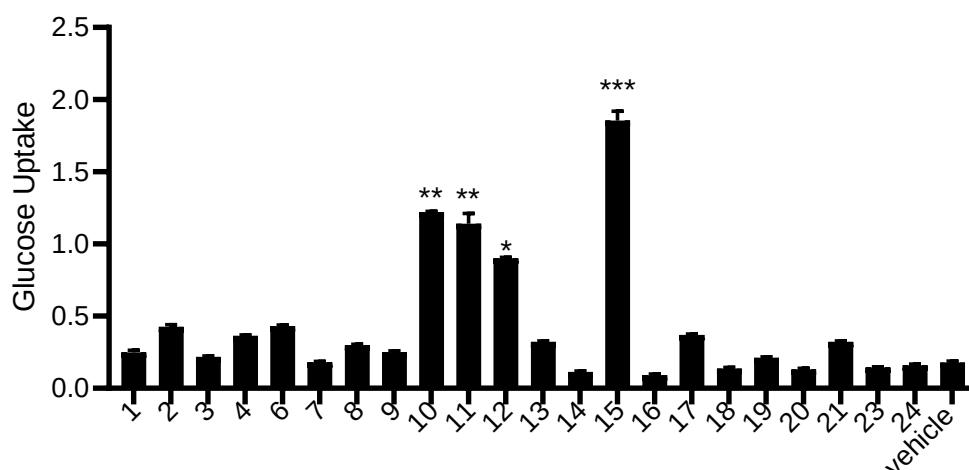


Figure 3.6 Modulatory activity of isolated *Centaurea* compounds on glucose uptake. 2-NBDG normalized fluorescence measured upon incubation of HEK293 cells with 3 μ M of the indicated *Centaurea* molecules for 48 h. Bars indicate means and standard errors of the mean (S.D.); statistical analysis was performed by multiple t test; p values = * < 0.05, ** < 0.01, *** < 0.001.

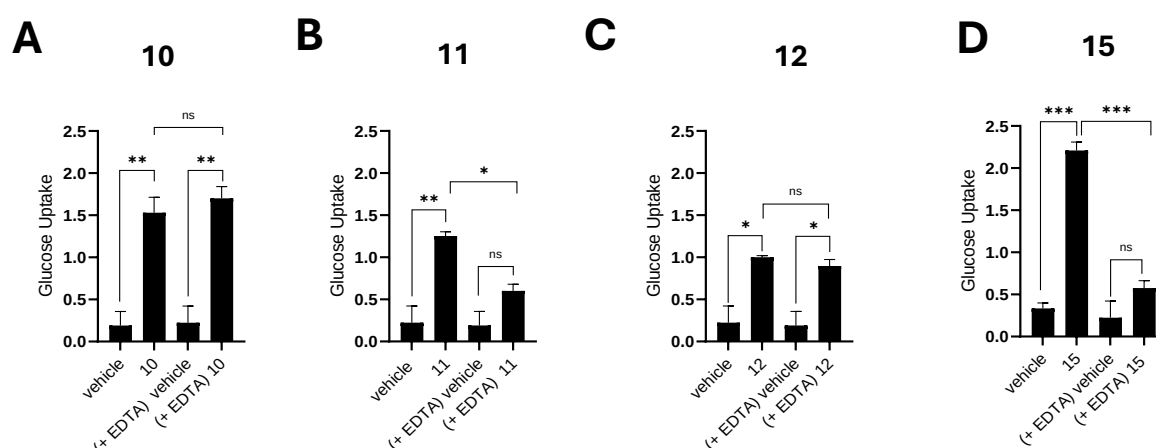


Figure 3.7 Glucose-transporter specificity of compounds **10**, **11**, **12** and **15**. Glucose uptake activity of HEK 293 cells treated for 48 hours with compounds **10**, **11**, **12** and **15** (3 μ M) by 2-NBDG assay. When indicated, EDTA 5 mM (an inhibitors of the Na^+ /Glucose cotransport).

3.3 Conclusions

The phytochemical characterization of the organic extract obtained from *C. sicula* aerial parts led to the isolation of 24 compounds, including four unprecedented sesquiterpenes and siculamide (**24**), an alkaloid bearing a novel skeleton. Its intricate structure can be rationalized through a biogenetic pathway that starts from the amino acids ornithine and glutamate. The

sesquiterpenoids isolated from *C. sicula* belong to the classes of elemanolides, eudesmanolides, and germacranolides, while guaianolides are completely absent. This is not completely surprising if we consider that *C. sicula* belongs to the section *Seridia* of the *Centaurea* genus, and all the species belonging to this section appear to be devoid of guaiane sesquiterpenoids (Bruno et al., 2013). Among the other species of the same section, germacrane dominate the sesquiterpenoids distribution (Barrero et al., 1997; Geppert et al., 1983), while only two species, *C. aspera* ssp. *subinermis* (Cardona et al., 1992) and *C. malacitana* (Barrero et al., 1997) showed, like *C. sicula*, the co-occurrence of germacrane, elemanes, and eudesmanes. The germacranolide dihydrocnicin (**10**) and the lignan salicifoliol (**15**) emerged for their remarkable and previously unknown activity.

3.4 Experimental section

3.4.1 Plant material

The aerial parts of *C. sicula* L. were collected at Campobello di Mazara (Trapani) in June 2021. Plants were identified by Emanuele Schimmenti. Herbarium samples (PAL 88- 155) were deposited at the Palermo Botanical Garden, Italy.

3.4.2 Extraction and isolation

500 g of *C. sicula* L. aerial parts were dried in an oven at 40 °C and defatted at room temperature using *n*-hexane (1.3 L for 3 days, three times) and successively extracted with CH₂Cl₂ (1.3 L for 3 days, three times). After filtration and in vacuo solvent, 2.4 g of the dry crude CH₂Cl₂ extract were obtained. This phase was subjected to MPLC-DAD on a C18 60 Å 50 µm, size 40 (55 g), CV 50 mL cartridge. The mobile phase was a mixture of A (H₂O with 0.1% formic acid) and B (CH₃CN) with the following gradient method: 90% A-10% B for CV 0-2; 80% A-20% B for CV 4-7; 70% A-30% B for CV 9-12; 60% A-40% B for CV 14-17; 50% A-50% B for CV 19-22; 40% A-60% B for CV 24-27; 30% A -70% B for CV 29-32; 20% A-80% B for CV 34-37; 10% A-90% B for CV 39-42; flow rate was set at 20.0 mL/ min. The UV detection wavelength was set at 227 nm. This separation yield to 17 fractions (labeled A-S). All of the fractions were purified on HPLC by semipreparative or analytical RP18 columns, affording the isolation of compounds **1-24**. Fraction E (10.1 mg) was separated by HPLC using H₂O-MeOH (80:20) as mobile phase leading to the isolation of siculamide (**24**, *t*_R = 14.5 min, 1.0 mg), the unprecedented compound **7** (*t*_R =

20.0 min, 2.8 mg), and β -hydroxypropiovanillone (**22**, t_R = 27.0 min, 1.1 mg). Fraction F (38.6 mg) was separated by RP-18 HPLC using H₂O-MeOH (75:25) as isocratic method, at flow rate 0.8 mL/min, to afford (-)-salicifoliol (**15**, t_R = 19.2 min, 1.7 mg), vanillin (**20**, t_R = 24.1 min, 0.5 mg), syringaldehyde (**21**, t_R = 27.7 min, 1.0 mg), isololiolide (**13**, t_R = 30.4 min, 1.7 mg), and the unreported compound **8** (t_R = 36.1 min, 9.0 mg). Fraction G (174.2 mg) was purified on RP-18 HPLC with H₂O-MeOH (60:40) at 3.0 mL/min flow rate, affording scopoletin (**23**, t_R = 12.0 min, 2.8 mg), melitensin (**1**, t_R = 15.0 min, 5.7 mg), daphnauranin E (**12**, t_R = 17.0 min, 2.9 mg), dehydromelitensin (**2**, t_R = 19.0 min, 0.5 mg), 11 β ,13-dihydrosalonitenolide (**9**, t_R = 20.2 min, 8.1 mg), dihydrocnicin (**10**, t_R = 23.0 min, 2.6 mg), **3** (t_R = 10.0 min, 13.8 mg), and **4** (t_R = 35.1 min, 1.8 mg). Fraction H (72.5 mg) was separated by RP-18 HPLC using H₂O- MeOH (60:40) as mobile phase at 2.5 mL/min flow rate, to yield pinoresinol (**16**, t_R = 17.5 min 1.1 mg), the unreported compound **6** (t_R = 18.0 min, 1.1 mg), abscisic acid (**14**, t_R = 20.5 min, 0.8 mg), and 6'-*trans*-2- butenoyl ester of prunasin (**19**, t_R = 23.0 min, 2.9 mg). Fraction I (89.9 mg) was purified with the same conditions used for fraction G to afford the previously undescribed compound **5** (t_R = 17.5 min, 0.8 mg) and 11-*epi*-dihydroartemisiifolin (**11**, t_R = 25.6 min, 4.8 mg). Fraction L was purified under the following conditions: H₂O- MeOH, 40:60, flow rate 3.0 mL/min, to afford arctigenin (**17**, t_R = 11.4 min, 1.9 mg) and jaceosidin (**18**, t_R = 19.1 min, 4.8 mg).

3.4.3 Spectral Data

Compound **5** : Colorless amorphous solid, $[\alpha]_D^{25} + 3.6$ (c 0.08, MeOH); ¹H NMR (700 MHz, CDCl₃): 5.67 (1H, dd, J 17.5, 10.6 Hz, H-1), 5.42 (1H, s, H-3b), 5.01 (1H, s, H-3a), 4.94 (1H, d, J 10.6 Hz, H-2a), 4.91 (1H, d, J 17.5 Hz, H-2b), 4.09 (1H, d, J 13.2 Hz, H-15b), 3.97 (1H, d, J 13.2 Hz, H-15a), 3.88 (1H, m, H-8), 3.71 (3H, OMe), 3.66 (1H, t, J 5.1 Hz, H-6), 3.29 (1H, m, H-11), 2.07 (1H, overlapped, H-5), 2.05 (1H, overlapped, H-7), 1.79 (1H, m, H-9b), 1.55 (1H, overlapped, H-9a), 1.25 (3H, d, J 7.3 Hz, H-13), 1.02 (3H, s, H-14). ¹³C NMR (150 MHz, CDCl₃): 179.0 (C-12), 147.0 (C-1), 146.9 (C-4), 114.3 (C-3), 111.6 (C-2), 68.7 (C-6), 67.7 (C-15), 65.5 (C-8), 55.1 (C-5), 53.3 (C-7), 52.1 (OMe), 47.9 (C-9), 39.9 (C-10), 36.4 (C-11), 18.7 (C-14), 10.6 (C-13). HRESIMS m/z found 321.1684 (C₁₆H₂₆O₅Na requires m/z 321.1678).

Compound **6** : Colorless amorphous solid, $[\alpha]_D^{25} + 5.7$ (c 0.11, MeOH); ¹H NMR (700 MHz, CDCl₃): 9.46 (1H, s, H-15), 6.37 (1H, s, H-5'b), 6.28 (1H, s, H-3b), 6.27 (1H, s, H-3a), 6.05 (1H, s,

PART 2

H-5'a), 5.67 (1H, dd, J 17.3, 10.3 Hz, H-1), 5.23 (1H, td, J 10.9, 4.2 Hz, H-8), 4.96 (1H, d, J 10.3 Hz, H-2b), 4.87 (1H, d, J 17.3 Hz, H-2a), 4.61 (1H, m, H-3'), 4.47 (1H, t, J 11.4 Hz, H-6), 3.84 (1H, dd, J 11.1, 3.6 Hz, H-4'b), 3.62 (1H, dd, J 11.1, 6.8 Hz, H-4'a) 3.05 (1H, d, J 11.4 Hz, H-5), 2.64 (1H, dd, J 12.2, 6.9 Hz, H-11), 2.05 (1H, q, J 11.4 Hz, H-7), 1.96 (1H, dd, J 12.2, 6.9 Hz, H-9b), 1.76 (1H, J 12.2, 4.2 Hz, H-9a), 1.24 (3H, d, J 6.9 Hz, H-13), 1.10 (3H, s, H-14). ^{13}C NMR (150 MHz, CDCl_3): 193.8 (C-15), 177.4 (C-12), 165.2 (C-1'), 145.1 (C-1), 145.0 (C-4), 138.8 (C-2'), 137.8 (C-3), 127.6 (C-5'), 113.0 (C-2), 77.0 (C-6), 71.2 (C-3'), 70.8 (C-8), 65.7 (C-4'), 56.0 (C-7), 44.7 (C-9), 46.1 (C-5), 41.8 (C-10), 41.0 (C-11), 18.2 (C-4), 14.0 (C-13). HRESIMS m/z found 401.1584 ($\text{C}_{20}\text{H}_{26}\text{O}_7\text{Na}$ requires m/z 401.1576).

Compound **7** : Colorless amorphous solid, $[\alpha]_{\text{D}}^{25} +7.5$ (c 0.12, MeOH). ^1H NMR (700 MHz, CDCl_3): 9.58 (1H, d, J 4.1 Hz, H-15), 3.98 (1H, dt, J 10.7 Hz, H-8), 3.90 (1H, t, J 3.90 Hz, H-6), 3.43 (1H, dd, J 11.3, 4.4 Hz), 2.52 (1H, m, H-11), 2.47 (1H, m, H-4), 2.34 (1H, dd, J 12.7, 4.2 Hz, H-9b), 1.87 (1H, m, H-2a), 1.76 (1H, overlapped, H-3b), 1.75 (1H, overlapped, H-5), 1.71 (1H, overlapped, H-7), 1.63 (1H, overlapped, H-2b), 1.59 (1H, overlapped, H-3a), 1.37 (3H, d, J 6.9 Hz, H-13), 1.26 (1H, overlapped, H-9a), 0.98 (3H, s, H-14); ^{13}C NMR (150 MHz, CDCl_3): 202.6 (C-15), 177.8 (C-12), 78.7 (C-6), 77.2 (C-1), 68.8 (C-8), 58.7 (C-7), 47.9 (C-5), 47.7 (C-4), 47.2 (C-9), 41.4 (C-11), 40.2 (C-10), 28.3 (C-2), 24.5 (C-3), 14.3 (C-13), 12.9 (C-14). HRESIMS m/z found 305.1355 ($\text{C}_{15}\text{H}_{22}\text{O}_5\text{Na}$ requires m/z 305.1365).

Compound **8** : Colorless amorphous solid, $[\alpha]_{\text{D}}^{25} + 24.8$ (c 0.10, MeOH); ^1H NMR (700 MHz, CDCl_3): 9.60 (1H, d, J 4.1 Hz, H-15), 5.20 (1H, m, H-8), 4.61 (1H, m, H-3'), 4.01 (1H, t, J 11.0 Hz, H-6), 3.44 (1H, dd, J 11.1, 4.5 Hz), 2.55 (1H, m, H-11), 2.53 (1H, m, H-4), 2.46 (1H, dd, J 12.6, 4.4 Hz, H-9b), 1.95 (1H, q, J 11.0 Hz, H-7), 1.87 (1H, m, H-2b), 1.78 (1H, t, J 11.0 Hz, H-7), 1.76 (1H, overlapped, H-3b), 1.63 (1H, overlapped, H-2a), 1.59 (1H, overlapped, H-3a), 1.37 (1H, t, J 12.6 Hz, H-9a), 1.21 (3H, d, J 6.9 Hz, H-13), 1.05 (3H, s, H-14). ^{13}C NMR (150 MHz, CDCl_3): 202.1 (C-15), 176.7 (C-12), 165.2 (C-1'), 138.8 (C-2'), 127.7 (C-5'), 78.3 (C-6), 77.1 (C-1), 71.3 (C-3'), 70.7 (C-8), 65.8 (C-4'), 58.4 (C-7), 48.0 (C-5), 47.9 (C-4), 42.8 (C-9), 40.7 (C-11), 40.4 (C-10), 28.4 (C-2), 24.4 (C-3), 14.1 (C-13), 12.8 (C-14). ESIMS m/z 419 $[\text{M}+\text{Na}]^+$; HRESIMS m/z found 419.1692 ($\text{C}_{20}\text{H}_{28}\text{O}_8\text{Na}$ requires 419.1682).

Siculamide (**24**) : Colorless amorphous solid, $[\alpha]_D^{25} -7.5$ (c 0.02, MeOH); ^1H NMR (700 MHz, CDCl_3): 5.18 (1H, t, J 6.7 Hz, H-9), 4.16 (1H, d, J 11.9 Hz, H-11a), 4.04 (1H, d, J 11.6 Hz, H-11b), 3.75 (1H, m, H-6b), 3.41 (1H, m, H-6a), 2.95 (1H, dt, J 18.7, 9.8 Hz, H-3b), 2.76 (1H, m, H-2a), 2.56 (1H, m, H-3a), 2.26 (1H, m, H-8a), 2.06 (1H, overlapped, H-2b), 2.03 (1H, overlapped, H-7a), 1.85 (1H, overlapped, H-8b); ^{13}C NMR (150 MHz, CDCl_3): 175.9 (C-4), 165.8 (C-10), 89.0 (C-9), 80.0 (C-1), 73.3 (C-11), 43.8 (C-6), 31.9 (C-8), 30.9 (C-2), 28.3 (C-3), 20.2 (C-7). HRESIMS m/z found 234.0741 ($\text{C}_{10}\text{H}_{13}\text{O}_4\text{NNa}$ requires m/z 234.0742).

3.4.4 Methanolysis of Melitensin

4.0 mg amount of melitensin (**1**) was treated with 3 equivalents of sodium methoxide in dry MeOH (1 mL) and left stirring overnight at room temperature. TLC was used to monitor the reaction, which was stopped after the detection of compound **5**.

3.4.5 NBDG Glucose Uptake Assay

Embryonic kidney cells (HEK293) were grown in DMEM (Gibco, Thermo Scientific, Waltham, MA, USA), supplemented with 10% FBS (Gibco), glutamine (Gibco), penicillin, and streptomycin (Gibco) and subcultivated every 2 days to avoid overconfluency. For the experiment, cells were seeded (5×10^3 /well) on a 96-well plate (PerkinElmer, Waltham, MA, USA) in a final volume of 100 μL /well of growth media. When confluency reached 20%, cells were treated for 48 h with the indicated compounds at 3 μM (working solution, prepared in DMEM with 10% FBS) for 48 h. Upon incubation, culture media was carefully removed and replaced with 100 μL of HBSS (Gibco) containing 100 μM 2-DG (2-deoxyglucose), 0.4 g/L BSA, and 1.27 mM CaCl_2 . When indicated, 5 mM EDTA was added to measure the contribution of SGLT to the readout of the assay. For 1 h, plates were incubated at 37 $^\circ\text{C}$. The medium of each well was then further supplemented with 2-(*N*-(7-nitrobenz-2-oxa-1,3-diazol-4-yl)amino)-2-deoxyglucose (2-NBDG) and 64 μM working solution (Invitrogen, Thermo Scientific, Waltham, MA, USA) diluted in DMEM containing 10% FBS. After 45 min of incubation with the fluorescent probe, plates were washed twice in PBS. Uptake of 2-NBDG was determined using an Envision 2105 multiplate reader (PerkinElmer) with the following parameters: $\lambda_{\text{exc}} = 471$ nm, $\lambda_{\text{ems}} = 529$ nm, and monochromator cutoff at 360 nm. Following the determination of 2-NBDG, cells were fixed in 3.7% paraformaldehyde for 30 min, permeabilized in 0.1% Triton X-100 in PBS, and stained with the nuclear dye DAPI (100 μM) to measure the number of cells in each well (λ_{exc}

PART 2

= 351 nm and $\lambda_{\text{exc}} = 450$ nm). Values of normalized glucose uptake are indicated as the ratio between intracellular 2- NDBG fluorescence and DAPI fluorescence (Serino et al., 2021).

References

- Aamer, M., Iqbal, S., Jabeen, A., Ahmad, M.S., Asrar, M., Yousuf, S., Choudhary, M.I. and Wang, Y., **2021**. Bioassay-guided isolation of anti-inflammatory constituents from *Erigeron canadensis*. *Chemistry of Natural Compounds*, 57, pp.749-751.
- Akkal, S., Benayache, F., Bentamene, A., Medjroubi, K., Seguin, E. and Tillequin, F., **2003**. Flavonoid aglycones from *Centaurea napifolia*. *Chemistry of natural compounds*, 39, pp.219-220.
- Arnone, A., Cardillo, R., Nasini, G. and de Pava, O.V., **1990**. Enantioselective reduction of racemic abscisic acid by *Aspergillus niger* cultures. *Journal of the Chemical Society, Perkin Transactions 1*, (11), pp.3061-3063.
- Barrero, A.F., Oltra, J.E., Morales, V., Álvarez, M. and Rodríguez-García, I., **1997**. Biomimetic cyclization of cnicin to malacitanolide, a cytotoxic eudesmanolide from *Centaurea malacitana*. *Journal of natural products*, 60(10), pp.1034-1035.
- Bicha, S., Amrani, A., Benaissa, O., León, F., Zama, D., Brouard, I., Benayache, S., Bentamene, A. and Benayache, F., **2013**. A flavonoid with high antioxidant effect from *Centaurea acaulis* L. *Der Pharmacia Lettre*, 5(6), pp.24-30.
- Bruno, M. and Herz, W., **1988**. Sesquiterpene lactones and flavones from *Centaurea cineraria* subsp. *umbrosa*. *Phytochemistry*, 27(6), pp.1873-1875.
- Bruno, M., Paternostro, M.P., Gedris, T.E. and Herz, W., **1996**. Sesquiterpene lactones and other constituents of *Centaurea nicaensis*. *Phytochemistry*, 41(1), pp.335-336.
- Bruno, M., Bancheva, S., Rosselli, S. and Maggio, A., **2013**. Sesquiterpenoids in subtribe *Centaureinae* (Cass.) Dumort (tribe *Cardueae*, *Asteraceae*): Distribution, ¹³C NMR spectral data and biological properties. *Phytochemistry*, 95, pp.19-93.
- Cardona, L., Fernández, I., Pedro, J.R. and Vidal, R., **1992**. Polyoxygenated terpenes and cyanogenic glucosides from *Centaurea aspera* var. *subinermis*. *Phytochemistry*, 31(10), pp.3507-3509.
- Christensen, L.P. and Lam, J., **1990**. Acetylenes and related compounds in *Cynareae*. *Phytochemistry*, 29(9), pp.2753-2785.
- Euro+Med PlantBase. <https://ww2.bgbm.org/EuroPlusMed/query.asp> (accessed July 28, **2023**).
- Dhouibi, N., Manuguerra, S., Arena, R., Mahdhi, A., Messina, C.M., Santulli, A. and Dhaouadi, H., **2020**. Screening of antioxidant potentials and bioactive properties of the extracts

- obtained from two *Centaurea* L. Species (*C. kroumirensis* Coss. and *C. sicula* L. subsp *sicula*). *Applied Sciences*, 10(7), p.2267.
- Geppert, B., Drożdż, B., Kietczewski, M. and Holub, M., **1983**. Sesquiterpene lactones. XXIII. Isolation of sesquiterpene lactones from *Centaurea* L. species. *Acta Societatis Botanicorum Poloniae*, 52(1), pp.23-34.
- Giardina, G., Raimondo, F.M. and Spadaro, V., **2007**. *A catalogue of plants growing in Sicily*. Herbarium Mediterraneum Panormitanum.
- González, A.G., Arteaga, J.M. and Bretón, J.L., **1973**. Germacranolides from *Centaurea seridis*. *Phytochemistry*, 12(12), pp.2997-2999.
- González, A.G., Estévez-Reyes, R. and Mato, C., **1989**. Salicifoliol, a new furo lactone-type lignan from *Bupleurum salicifolium*. *Journal of Natural Products*, 52(5), pp.1139-1142.
- Higuchi, Y., Aoki, S., Takenami, H., Kamimura, N., Takahashi, K., Hishiyama, S., Lancefield, C.S., Ojo, O.S., Katayama, Y., Westwood, N.J. and Masai, E., **2018**. Bacterial catabolism of β -hydroxypropiovanillone and β -hydroxypropiosyringone produced in the reductive cleavage of arylglycerol- β -aryl ether in lignin. *Applied and Environmental Microbiology*, 84(7), pp.e02670-17.
- Huang, S.Z., Ma, Q.Y., Wang, Q., Dai, H.F., Liu, Y.Q., Zhou, J. and Zhao, Y.X., **2018**. Daphnauranins C–E, three new antifeedants from *Daphne aurantiaca* roots. *Molecules*, 23(10), p.2429.
- Jimeno, M.L., Aprea-Rojas, M.D.C., Cano, F.H. and Rodríguez, B., **2004**. NMR and x-ray conformational study of artemisiifolin and three other related germacranolides. *Magnetic Resonance in Chemistry*, 42(5), pp.474-483.
- Kimura, J. and Maki, N., **2002**. New loliolide derivatives from the brown alga *Undaria pinnatifida*. *Journal of Natural Products*, 65(1), pp.57-58.
- Lazari, D., Garcia, B., Skaltsa, H., Pedro, J. and Harvala, C., **1998**. Sesquiterpene lactones from *Onopordon laconicum* and *O. sibthorpium*. *Phytochemistry*, 47(3), pp.415-422.
- Marco, J.A., Sanz-Cervera, J.F., Yuste, A., Sancenón, F. and Carda, M., **2005**. Sesquiterpenes from *Centaurea aspera*. *Phytochemistry*, 66(14), pp.1644-1650.
- Medjroubi, K., Bouderdara, N., Benayache, F., Akkal, S., Seguin, E. and Tillequin, F., **2003**. Sesquiterpene lactones of *Centaurea nicaensis*. *Chemistry of natural compounds*, 39, pp.506-507.
- Mohamed, T.A., Elshamy, A.I., Abd-ElGawad, A.M., Hussien, T.A., El-Toumy, S.A., Efferth, T. and Hegazy, M.E.F., **2021**. Cytotoxic and chemotaxonomic study of isolated metabolites

- from *Centaurea aegyptiaca*. *Journal of the Chinese Chemical Society*, 68(1), pp.159-168.
- Porter, T.H., Mabry, T.J., Yoshioka, H. and Fischer, N.H., **1970**. The isolation and structure determination of artemisiifolin, a new germacranolide from *Ambrosia artemisiifolia* L.(Compositae). *Phytochemistry*, 9(1), pp.199-204.
- Sarker, S.D., Savchenko, T., Whiting, P., Šik, V. and Dinan, L.N., **1997**. Moschamine, cis-moschamine, moschamindole and moschamindolol: four novel indole alkaloids from *Centaurea moschata*. *Natural Product Letters*, 9(3), pp.189-199.
- Sen, A., Ozbas Turan, S. and Bitis, L., **2017**. Bioactivity-guided isolation of anti-proliferative compounds from endemic *Centaurea kilaea*. *Pharmaceutical biology*, 55(1), pp.541-546.
- Senatore, F., Formisano, C., Raio, A., Bellone, G. and Bruno, M., **2008**. Volatile components from flower-heads of *Centaurea nicaeensis* All., *C. parlatoris* Helder and *C. solstitialis* L. ssp. *schouwii* (DC.) Dostál growing wild in southern Italy and their biological activity. *Natural Product Research*, 22(10), pp.825-832.
- Serino, E., Chahardoli, A., Badolati, N., Sirignano, C., Jalilian, F., Mojarab, M., Farhangi, Z., Rigano, D., Stornaiuolo, M., Shokoohinia, Y. and Taglialatela-Scafati, O., **2021**. Salvigenin, a trimethoxylated flavone from *Achillea wilhelmsii* C. Koch, exerts combined lipid-lowering and mitochondrial stimulatory effects. *Antioxidants*, 10(7), p.1042.
- Serino, E., Chianese, G., Musto, G., Zengin, G., Rigano, D., Stornaiuolo, M., Formisano, C. and Taglialatela-Scafati, O., **2022**. Guaiane-rich phytochemical profile of *Centaurea kotschyi* subsp. *persica* (Boiss.) Wagenitz and identification of hypoglycaemic metabolites. *Phytochemistry*, 199, p.113189.
- Shoeb, M., Celik, S., Jaspars, M., Kumarasamy, Y., MacManus, S.M., Nahar, L., Thoo-Lin, P.K. and Sarker, S.D., **2005**. Isolation, structure elucidation and bioactivity of schischkiniin, a unique indole alkaloid from the seeds of *Centaurea schischkinii*. *Tetrahedron*, 61(38), pp.9001-9006.
- Wang, F.P. and Chen, Q.H., **2014**. Stemona alkaloids: biosynthesis, classification, and biogenetic relationships. *Natural Product Communications*, 9(12), p.1934578X1400901238.
- Zater, H., Huet, J., Fontaine, V., Benayache, S., Stévigny, C., Duez, P. and Benayache, F., **2016**. Chemical constituents, cytotoxic, antifungal and antimicrobial properties of *Centaurea*

PART 2

diluta Ait. subsp. algeriensis (Coss. & Dur.) Maire. *Asian Pacific Journal of Tropical Medicine*, 9(6), pp.554-561.

APPENDIX B

Here are reported the spectra of the previously undescribed compounds isolated from *Centaurea sicula* (described in PART 2 – Chapter 2 of this manuscript).

Figures S1-S5: NMR spectra of compound 5 (CDCl ₃ , 600MHz).....	84
Figures S6-S10: NMR spectra of compound 6 (CDCl ₃ , 600MHz).....	86
Figures S11-S15: NMR spectra of compound 7 (CDCl ₃ , 700MHz)	89
Figures S16-20: NMR spectra of compound 8 (CDCl ₃ , 700MHz)	91
Figures S21-S25: NMR spectra of compound 9 (CDCl ₃ , 700MHz).....	94
Figures S26-S30: NMR spectra of compound 12 (CDCl ₃ , 600MHz)	96
Figures S31-S35: ¹ H NMR spectrum of compound 24 (CDCl ₃ , 700MHz).....	99

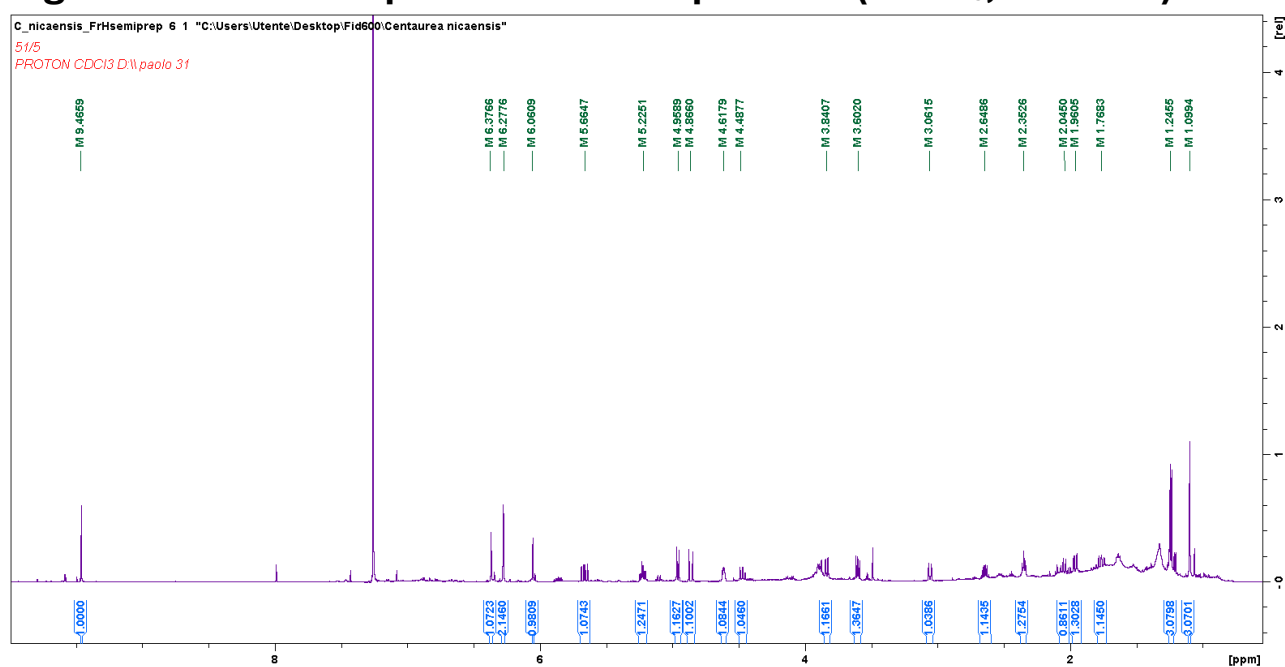
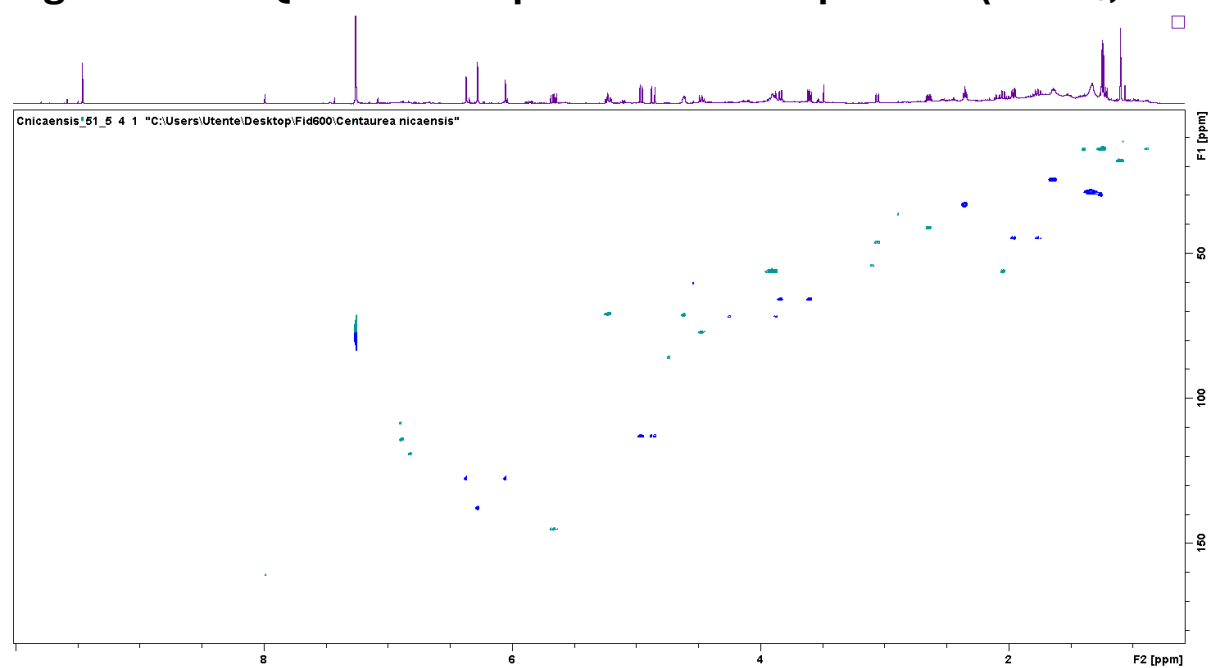
Figure S1: ^1H NMR spectrum of compound 5 (CDCl_3 , 600MHz)**Figure S2: HSQC 2D NMR spectrum of compound 5 (CDCl_3 , 600MHz)**

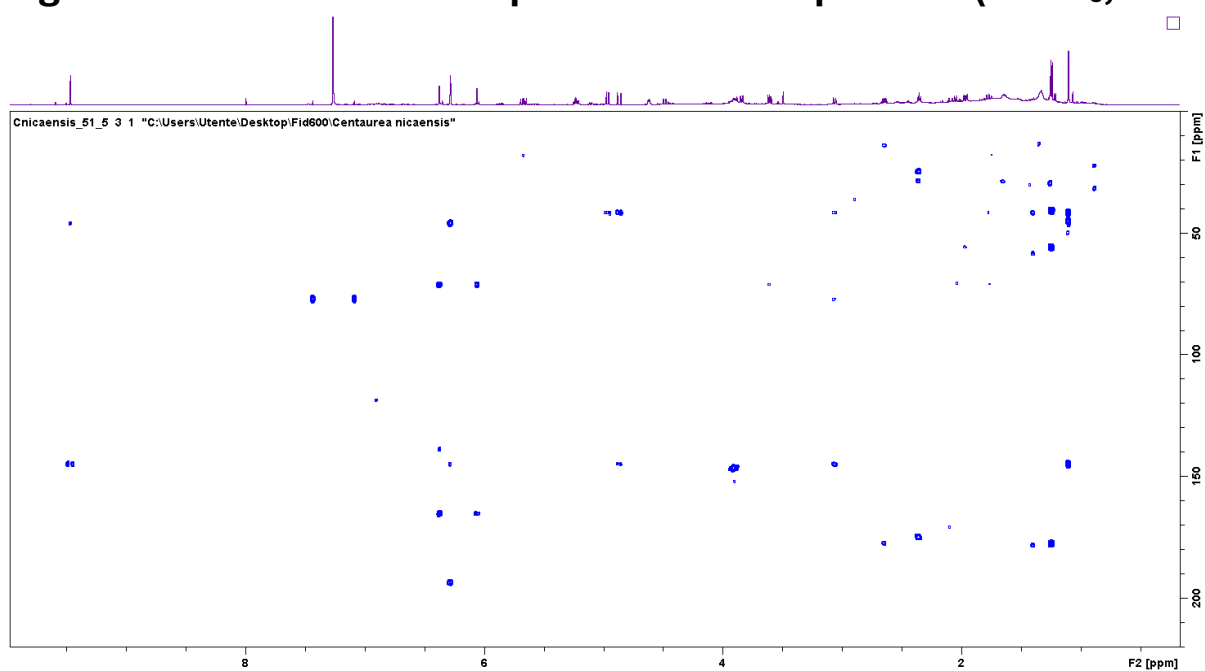
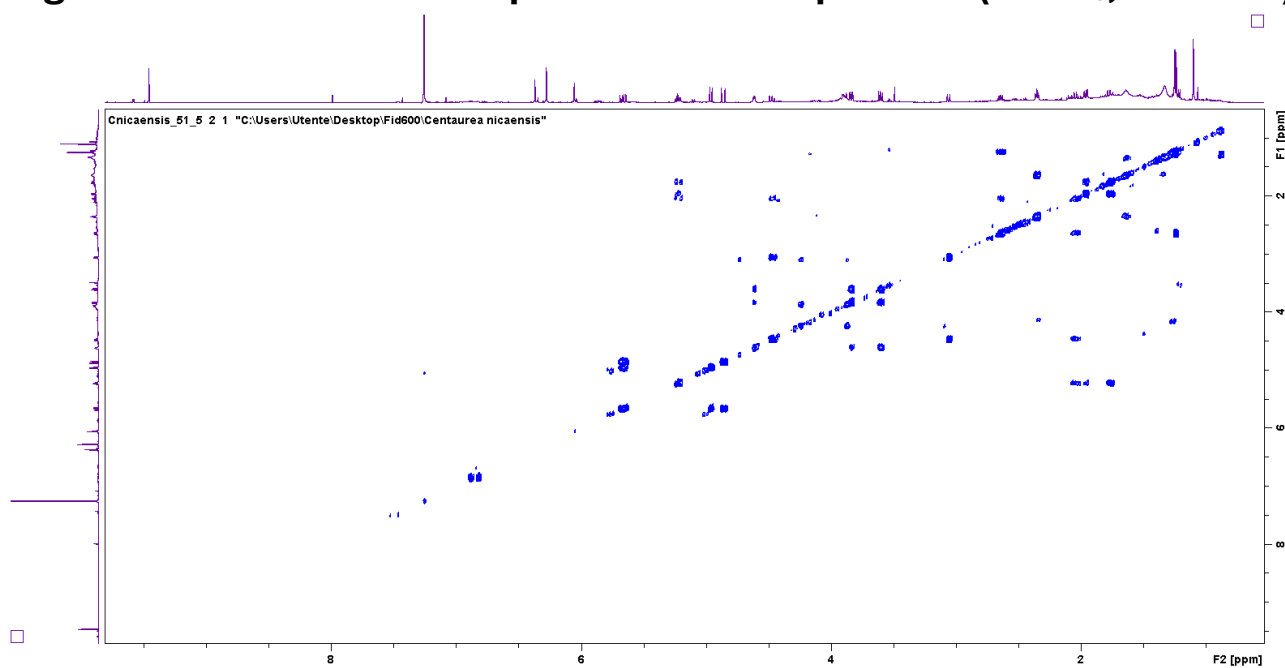
Figure S3: HMBC 2D NMR spectrum of compound 5 (CDCl₃, 600MHz)**Figure S4: COSY 2D NMR spectrum of compound 5 (CDCl₃, 600MHz)**

Figure S5: NOESY 2D NMR spectrum of compound 5 (CDCl₃, 600MHz)

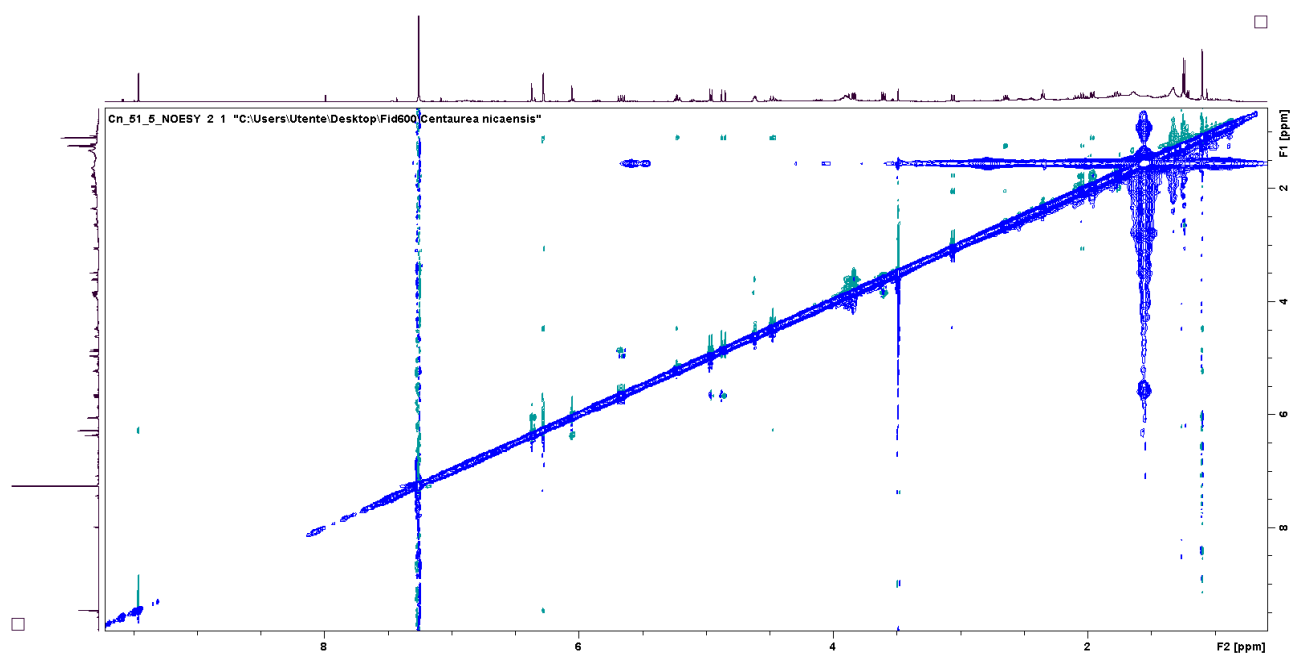


Figure S6: ¹H NMR spectrum of compound 6 (CDCl₃, 600MHz)

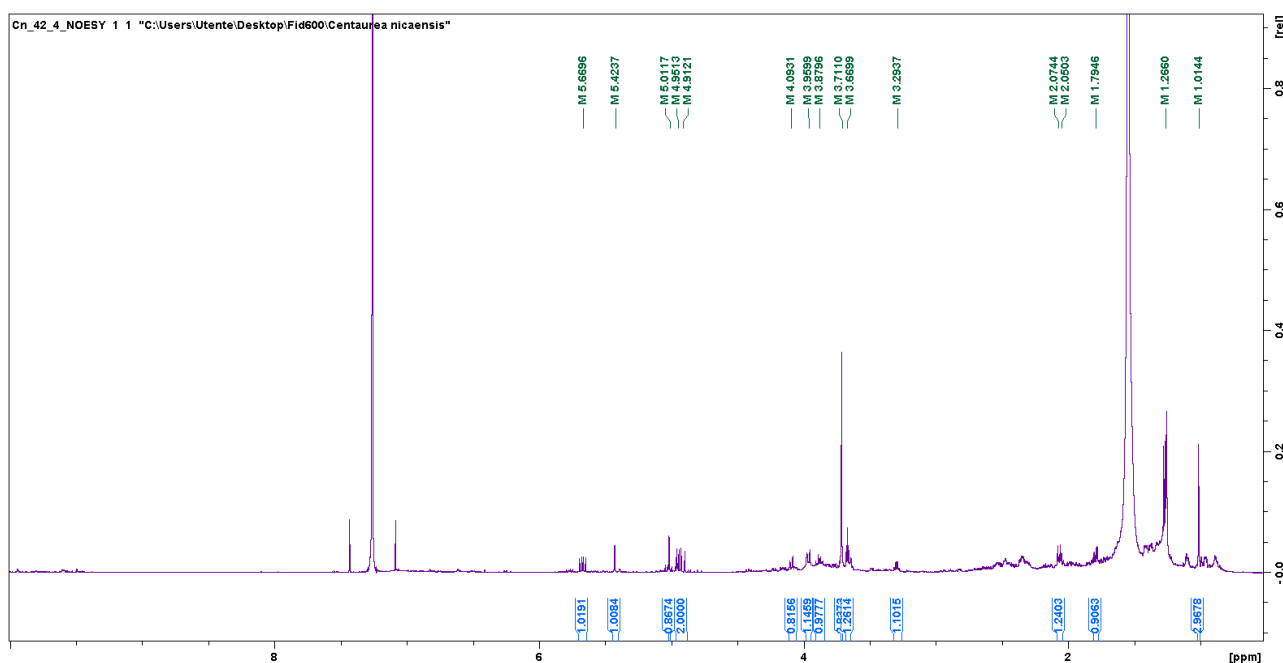


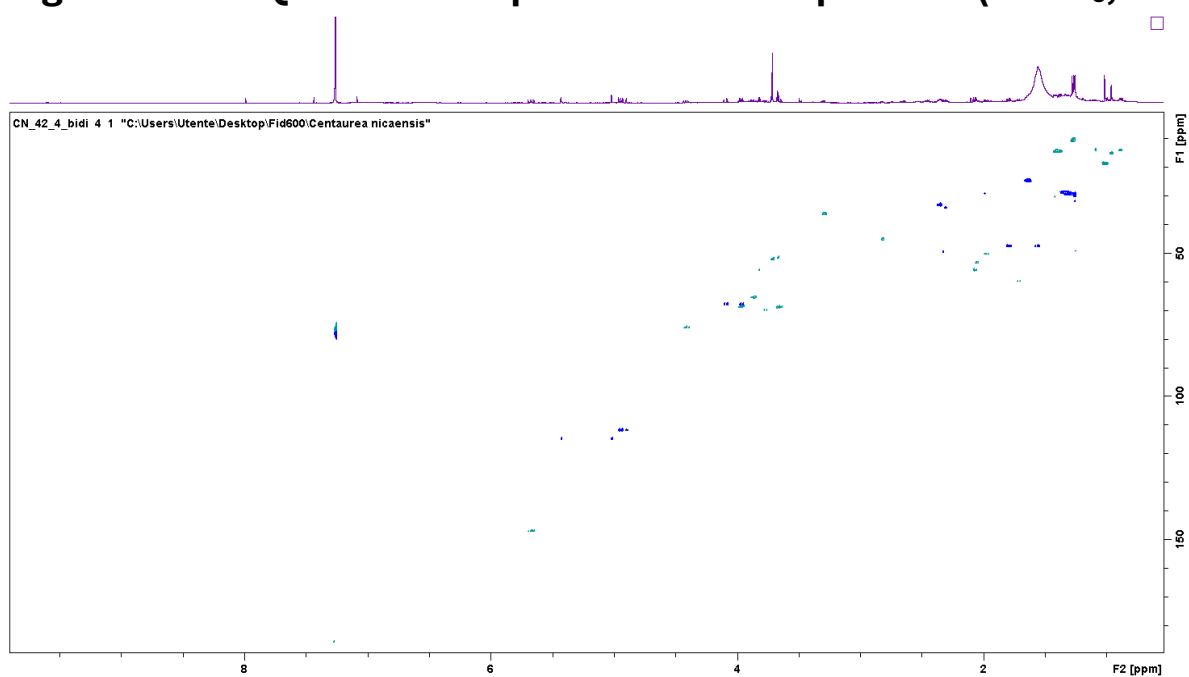
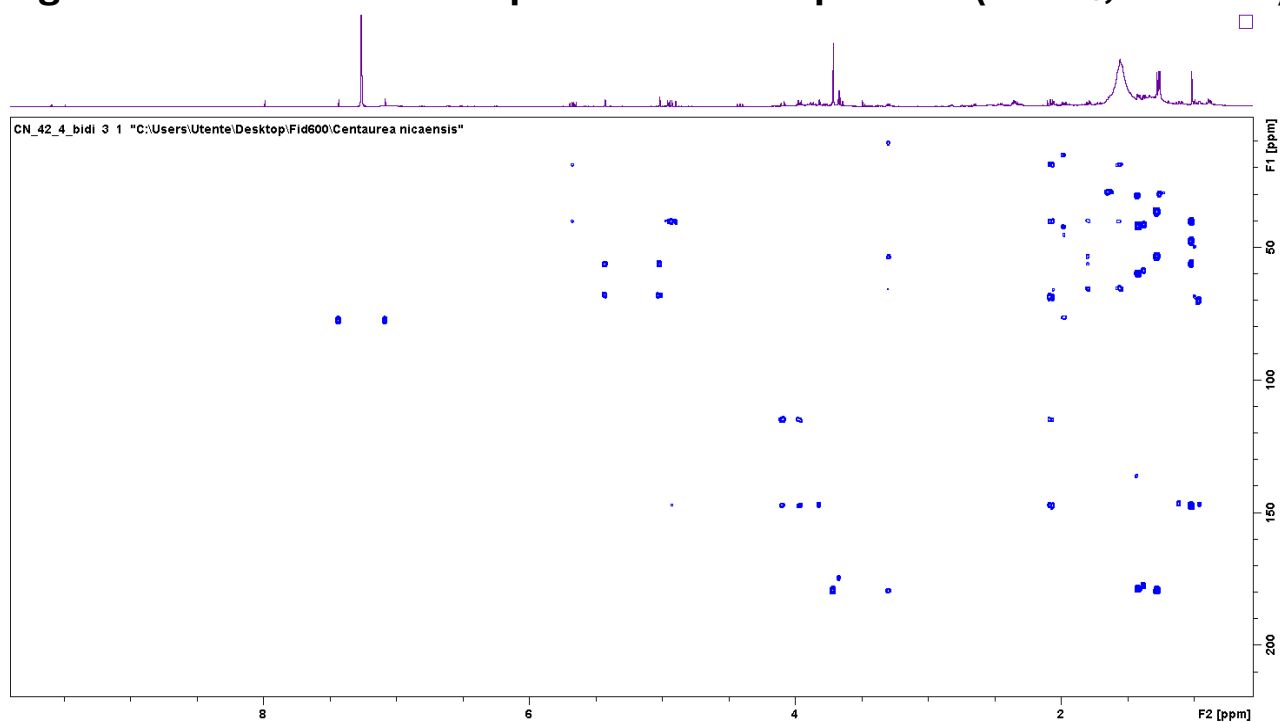
Figure S7: HSQC 2D NMR spectrum of compound 6 (CDCl₃, 600MHz)**Figure S8: HMBC 2D NMR spectrum of compound 6 (CDCl₃, 600MHz)**

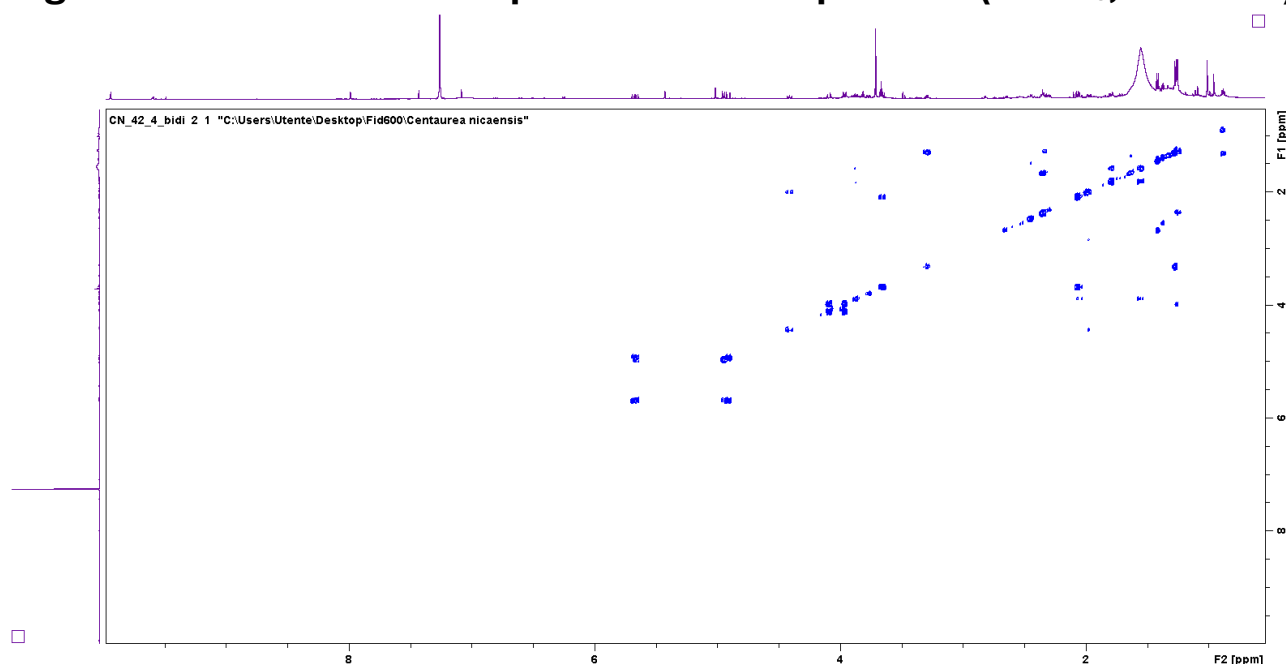
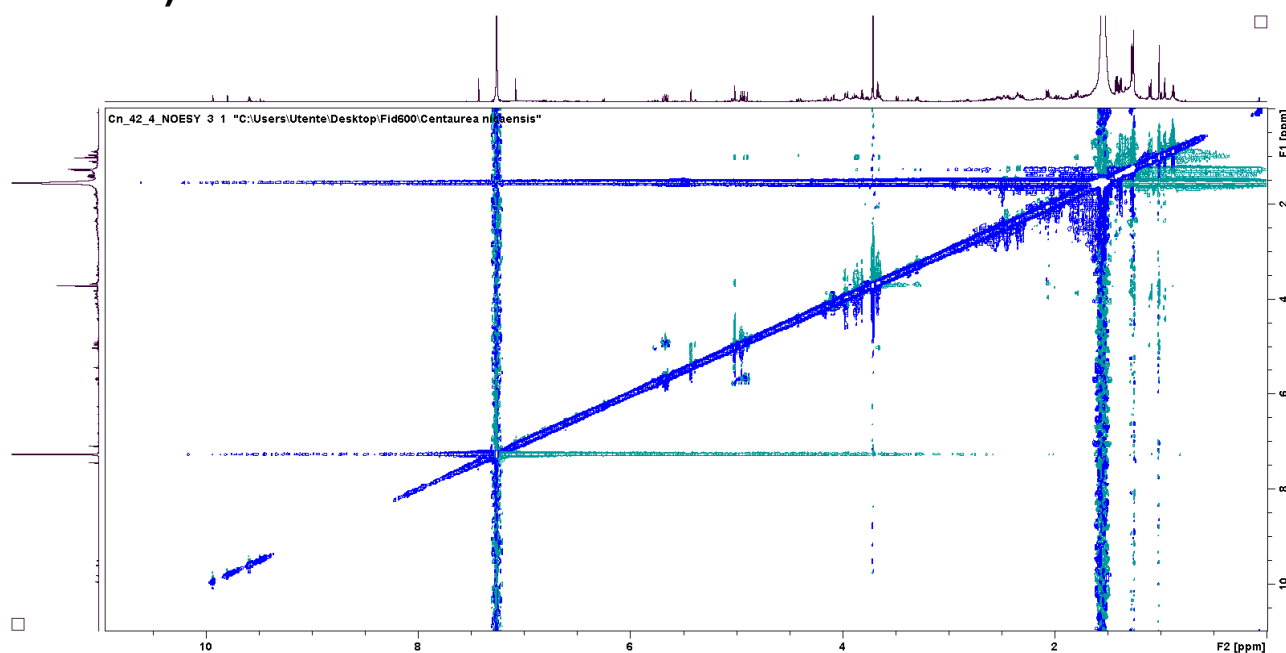
Figure S9: COSY 2D NMR spectrum of compound 6 (CDCl₃, 600MHz)**Figure S10: NOESY 2D NMR spectrum of compound 6 (CDCl₃, 600MHz)**

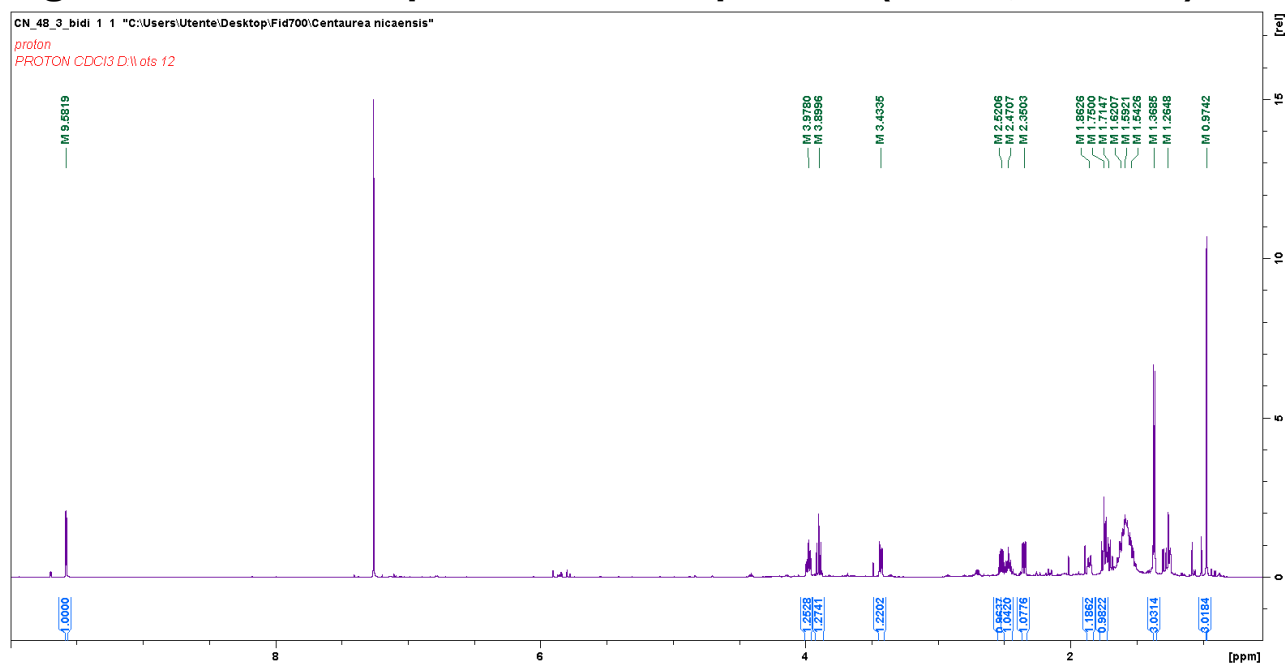
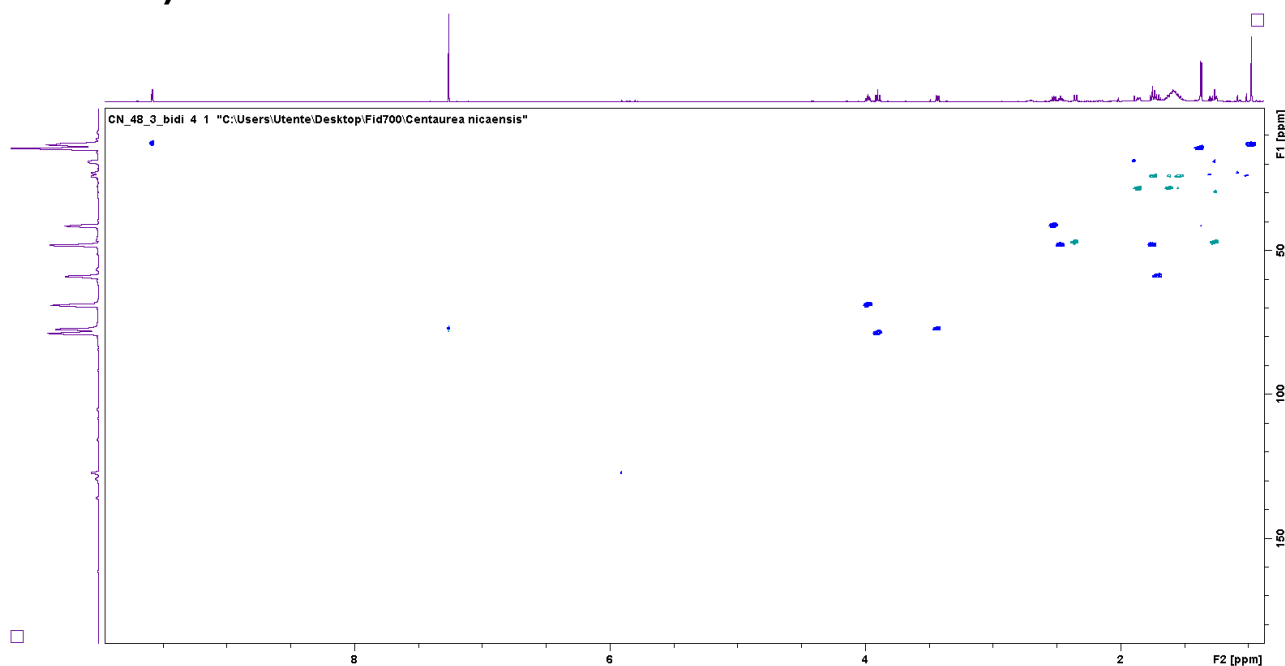
Figure S11: ^1H NMR spectrum of compound 7 (CDCl_3 , 700MHz)**Figure S12: HSQC 2D NMR spectrum of compound 7 (CDCl_3 , 700MHz)**

Figure S13: HMBC 2D NMR spectrum of compound 7 (CDCl₃, 700MHz)

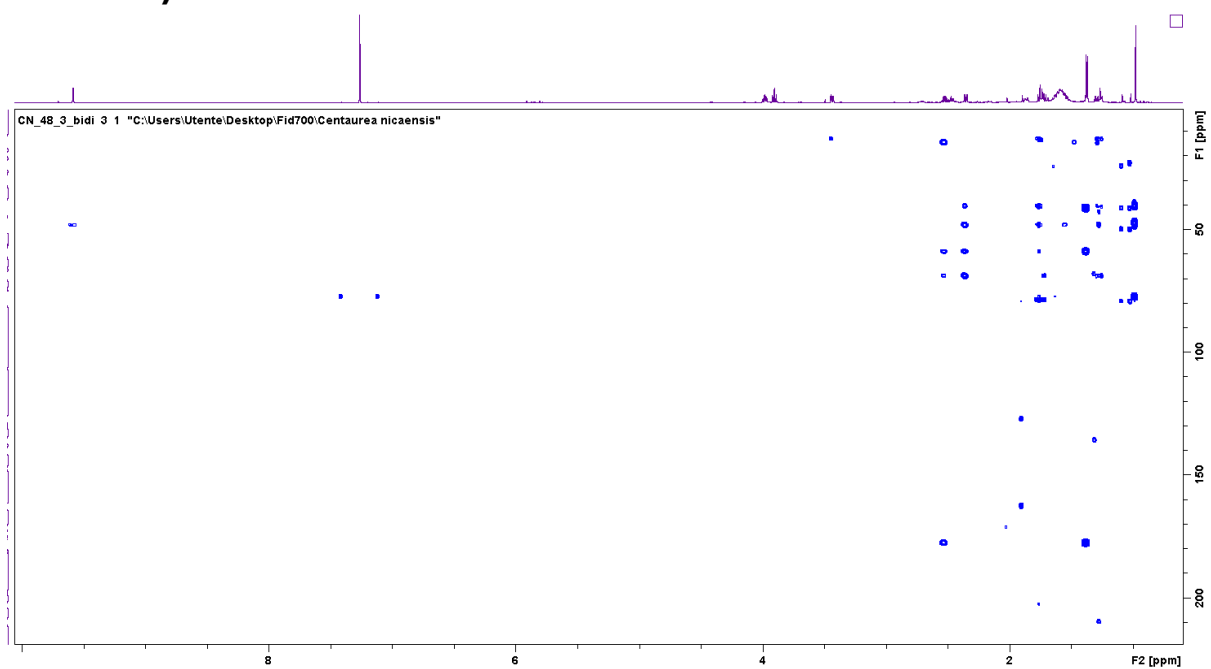


Figure S14: COSY 2D NMR spectrum of compound 7 (CDCl₃, 700MHz)

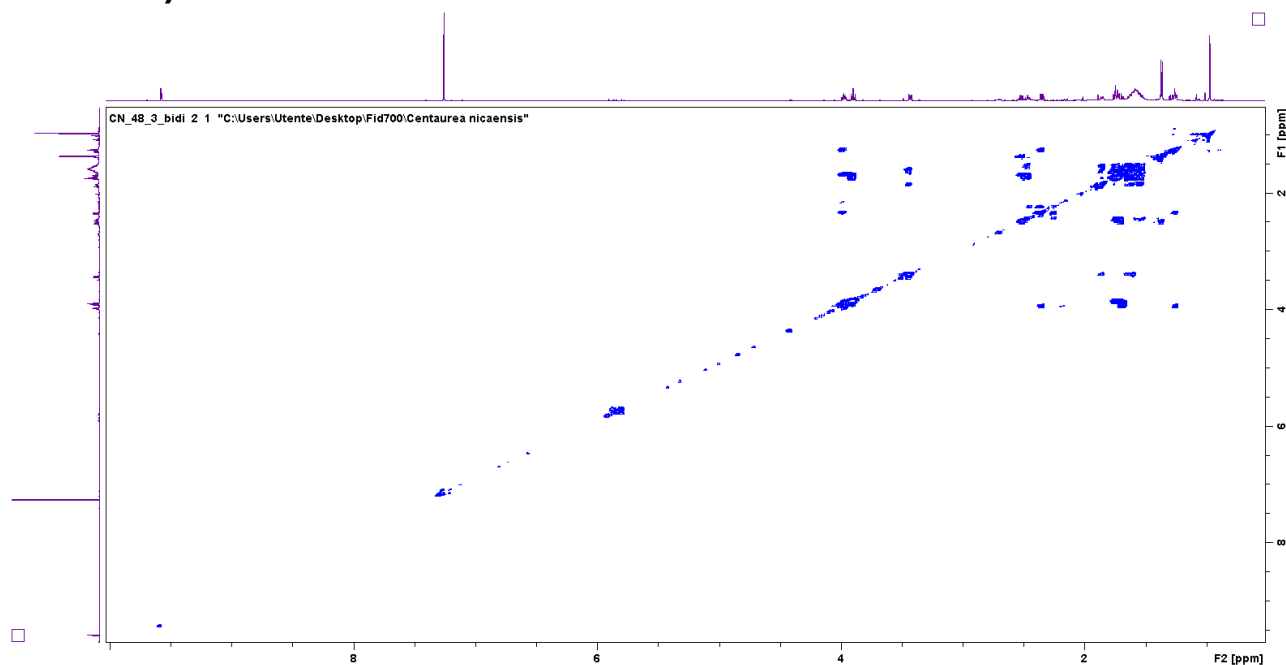


Figure S15: NOESY 2D NMR spectrum of compound 7 (CDCl₃, 700MHz)

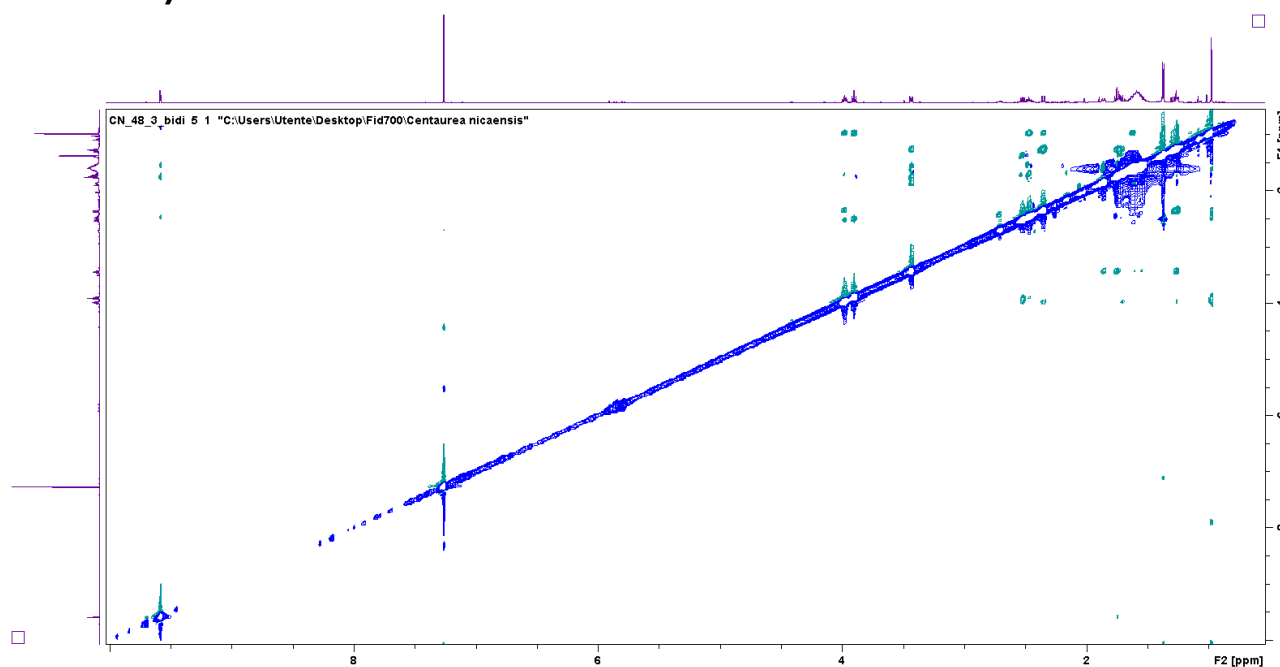


Figure S16: ¹H NMR spectrum of compound 8 (CDCl₃, 700MHz)

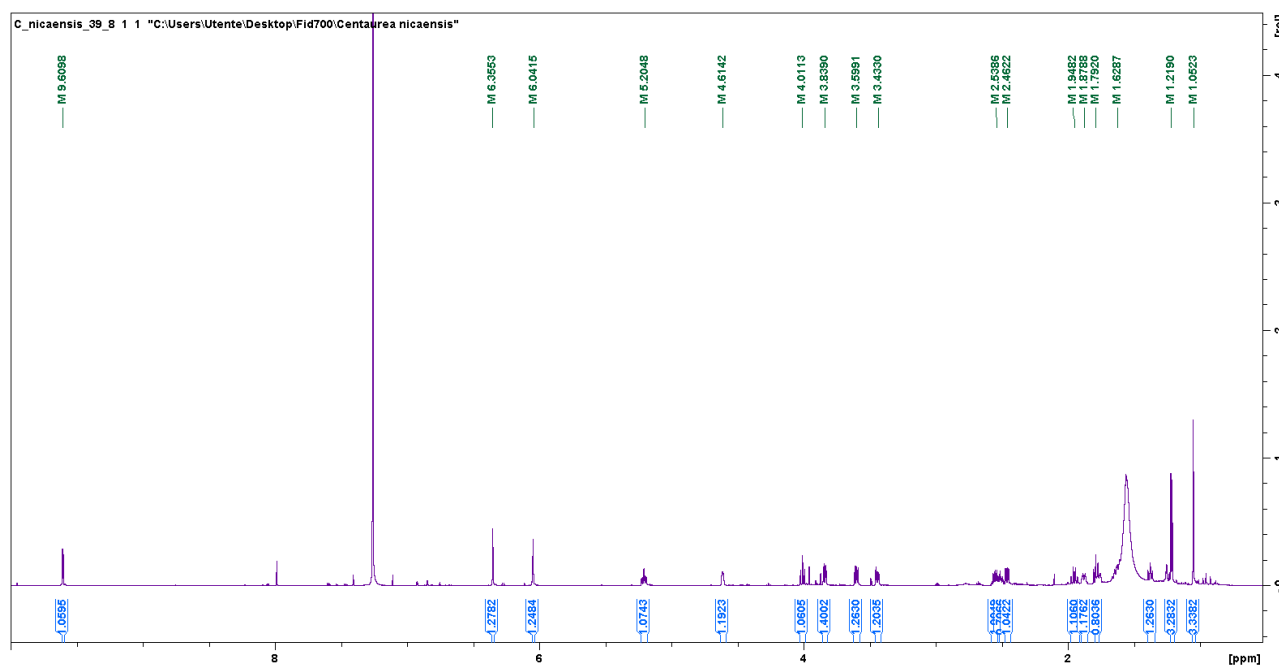


Figure S17: HSQC 2D NMR spectrum of compound 8 (CDCl₃, 700MHz)

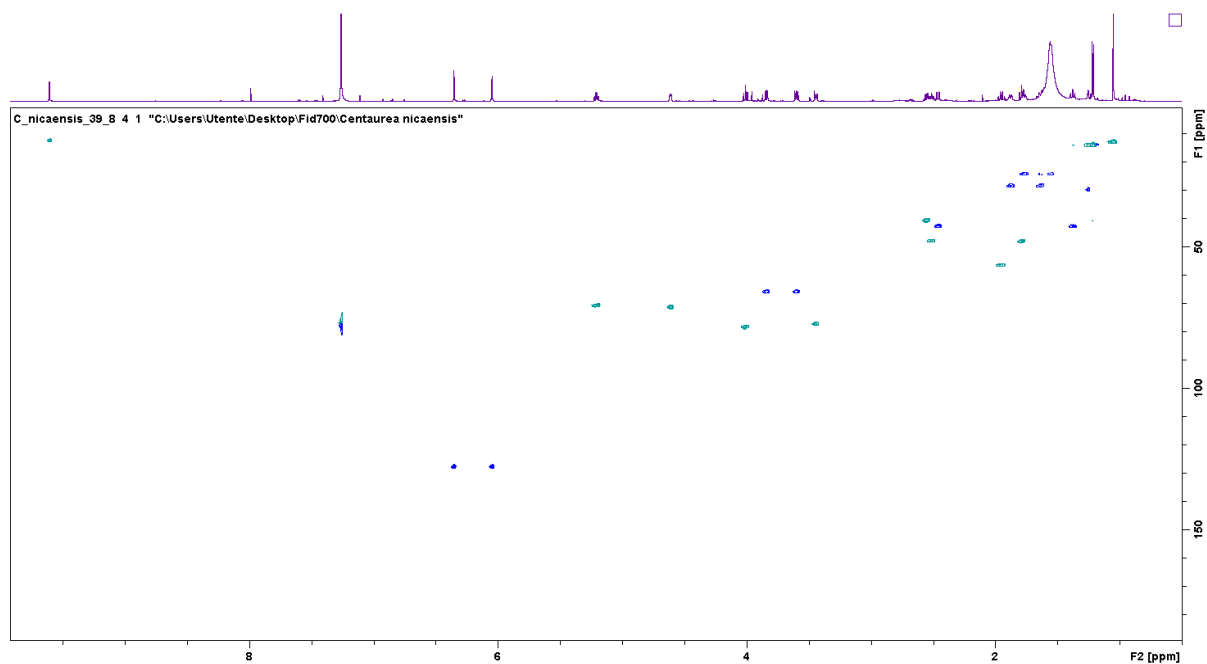


Figure S18: HMBC 2D NMR spectrum of compound 8 (CDCl₃, 700MHz)

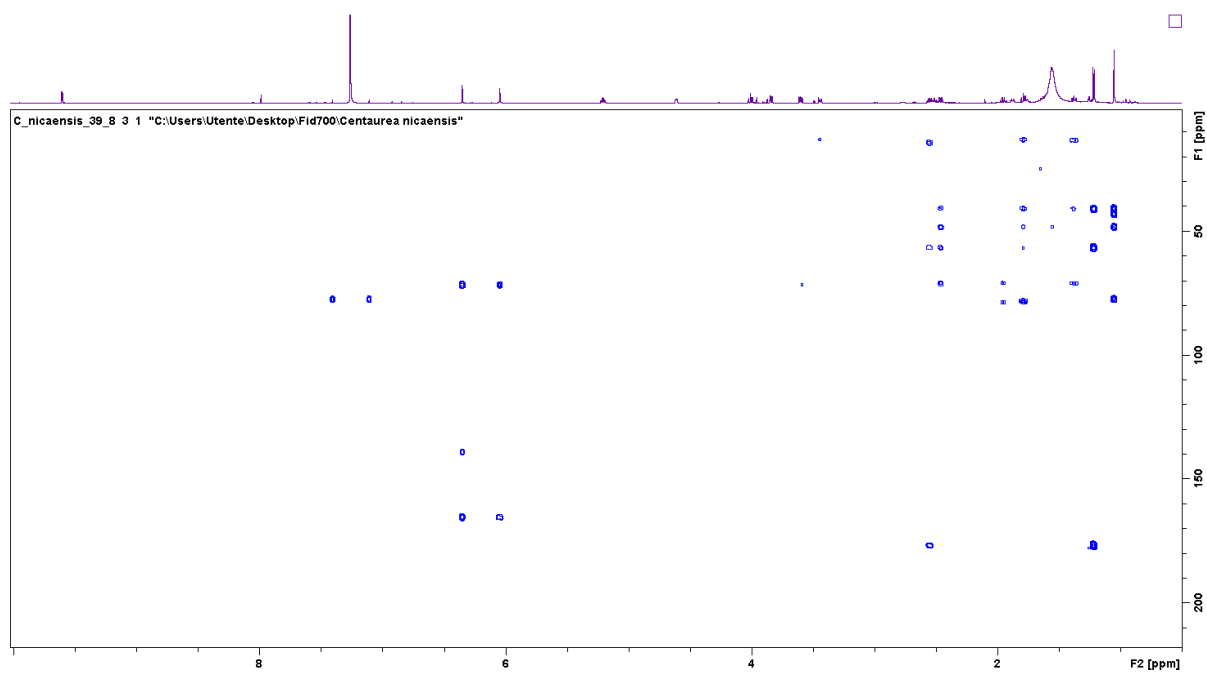


Figure S19: COSY 2D NMR spectrum of compound 8 (CDCl₃, 700MHz)

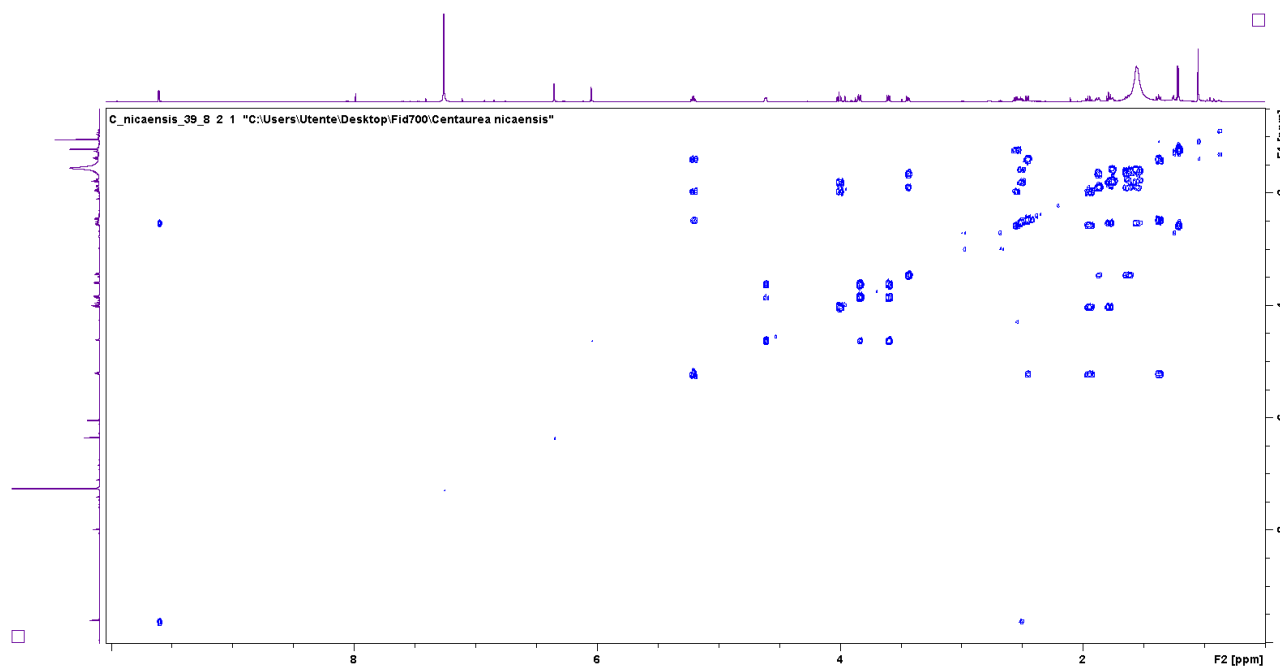


Figure S20: NOESY 2D NMR spectrum of compound 8 (CDCl₃, 700MHz)

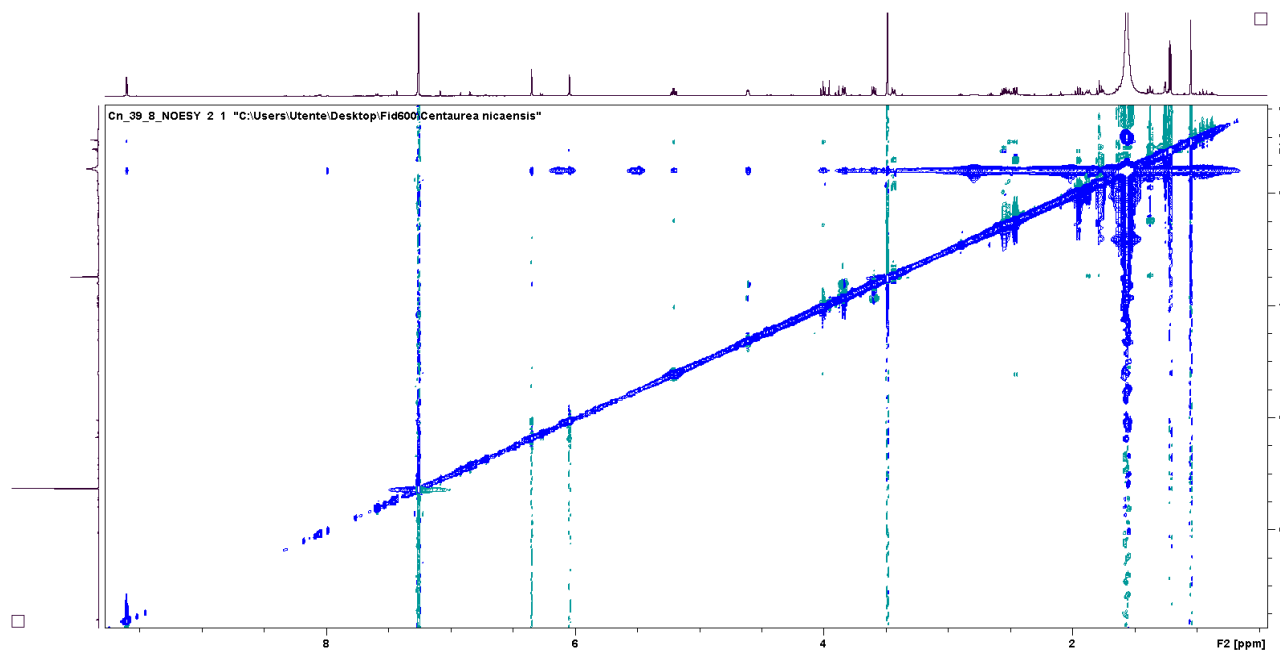


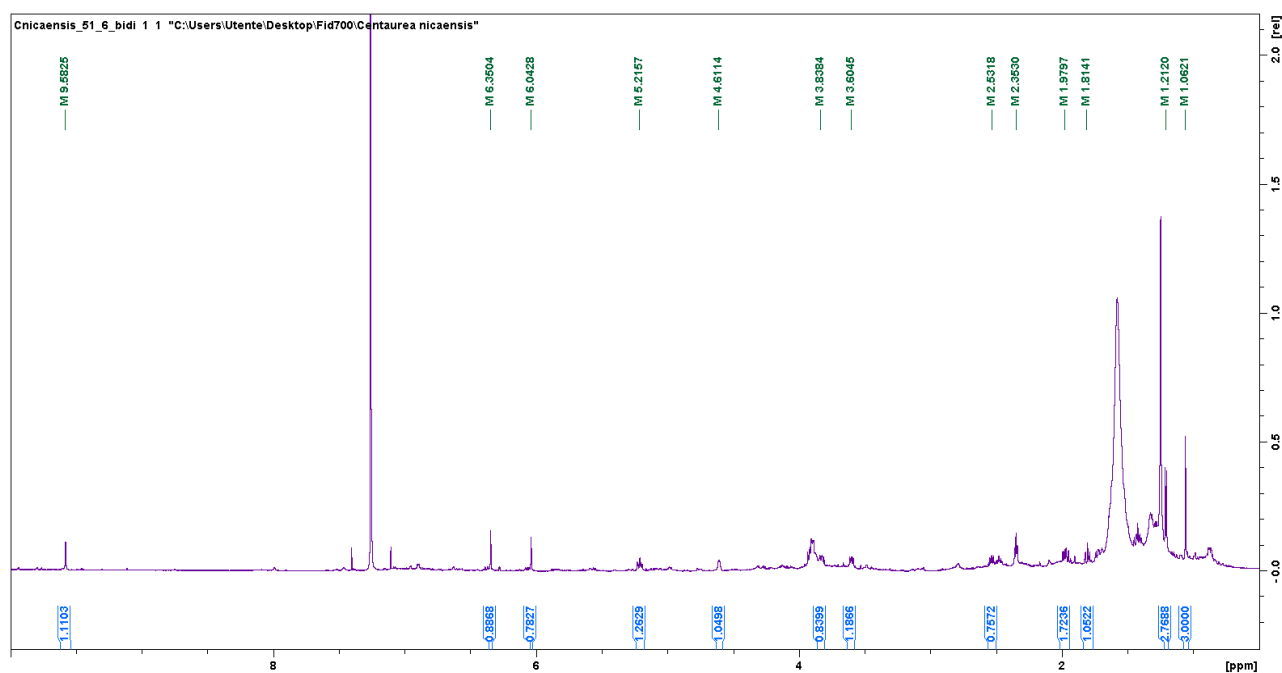
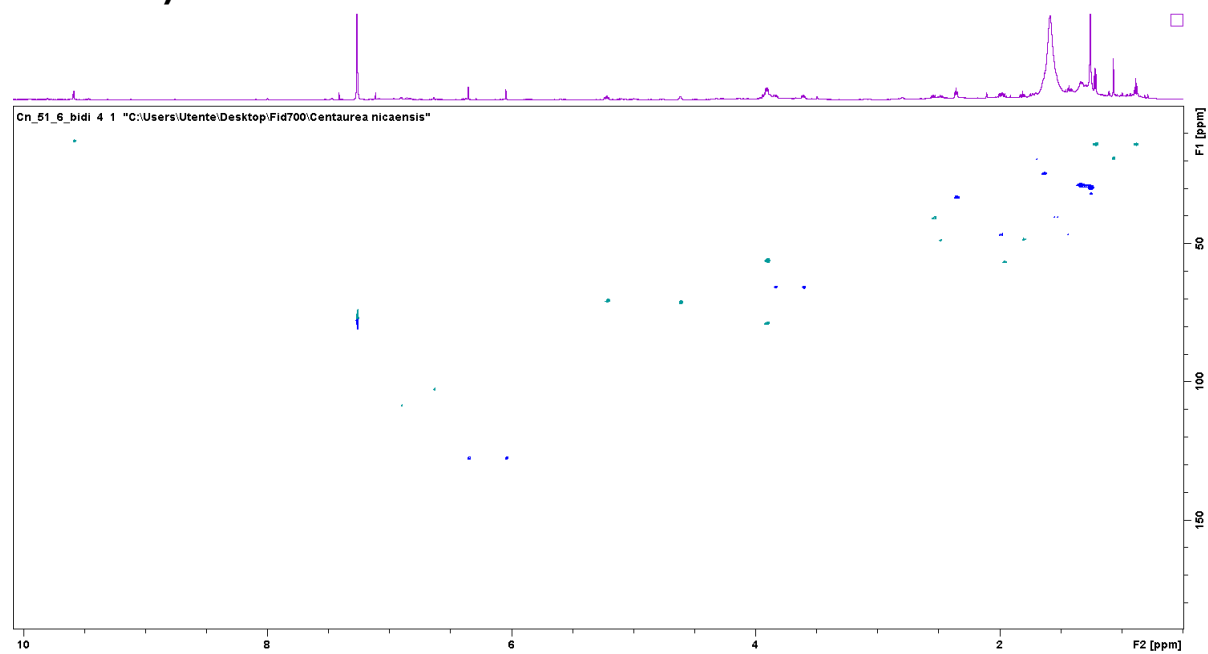
Figure S21: ^1H NMR spectrum of compound 9 (CDCl_3 , 700MHz)**Figure S22: HSQC 2D NMR spectrum of compound 9 (CDCl_3 , 700MHz)**

Figure S23: HMBC 2D NMR spectrum of compound 9 (CDCl₃, 700MHz)

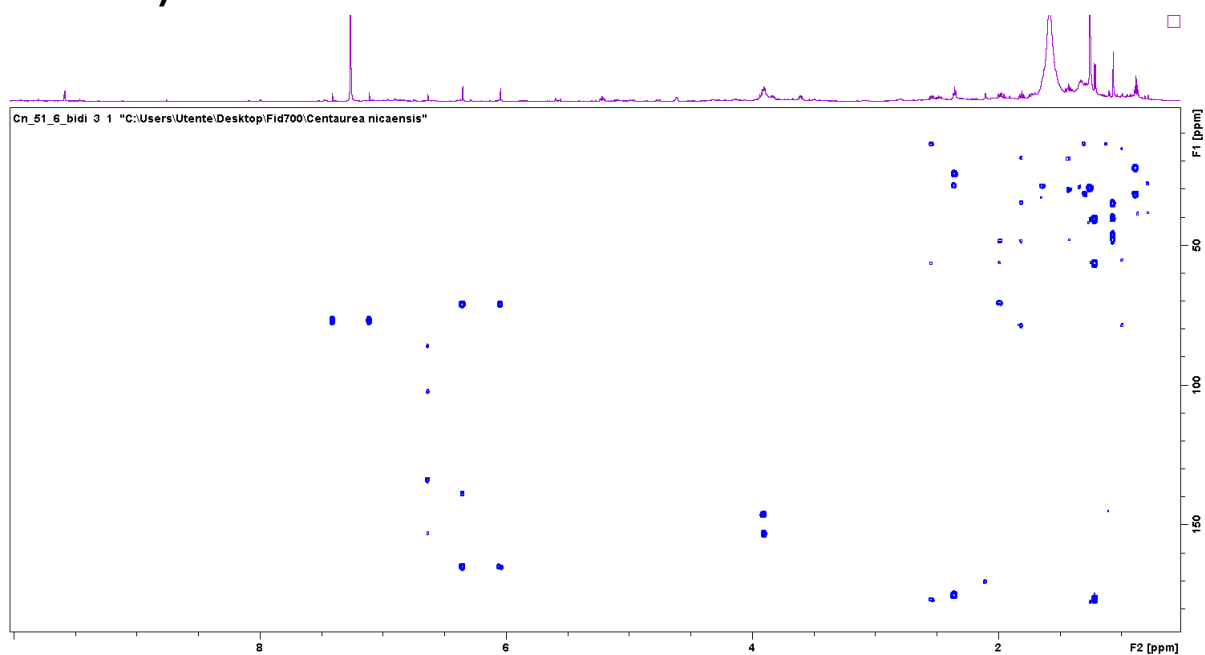


Figure S24: COSY 2D NMR spectrum of compound 9 (CDCl₃, 700MHz)

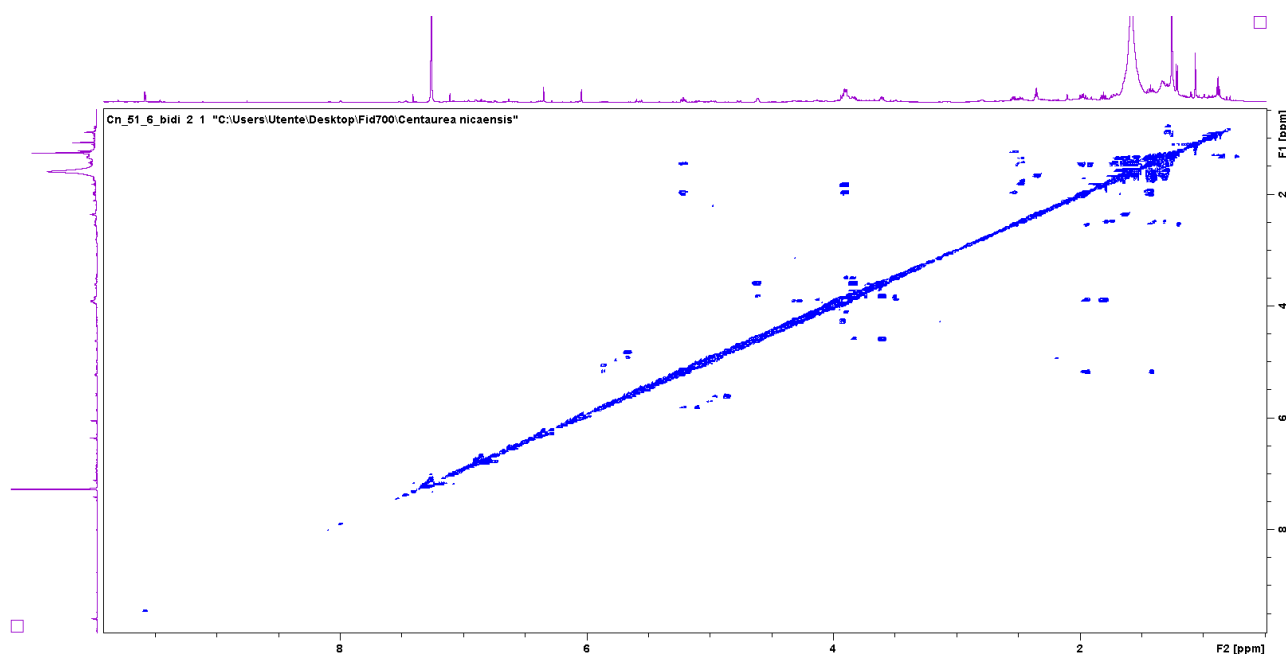


Figure S25: NOESY 2D NMR spectrum of compound 9 (CDCl₃, 700MHz)

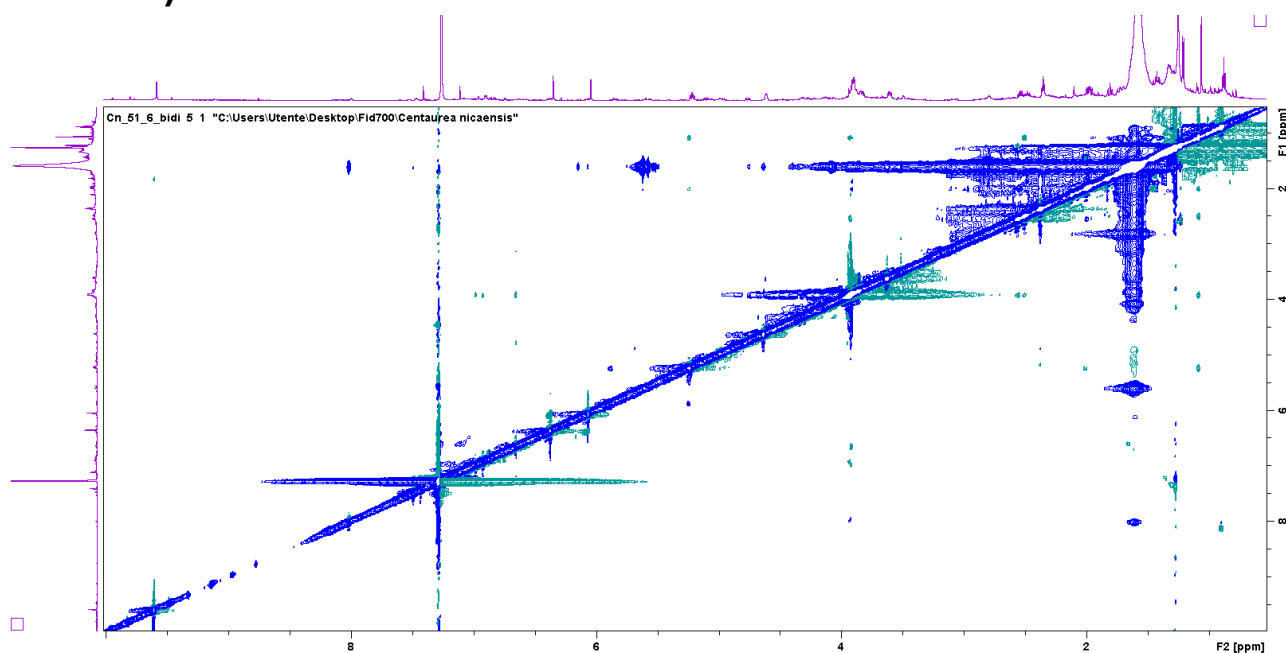


Figure S26: ¹H NMR spectrum of compound 12 (CDCl₃, 600MHz)

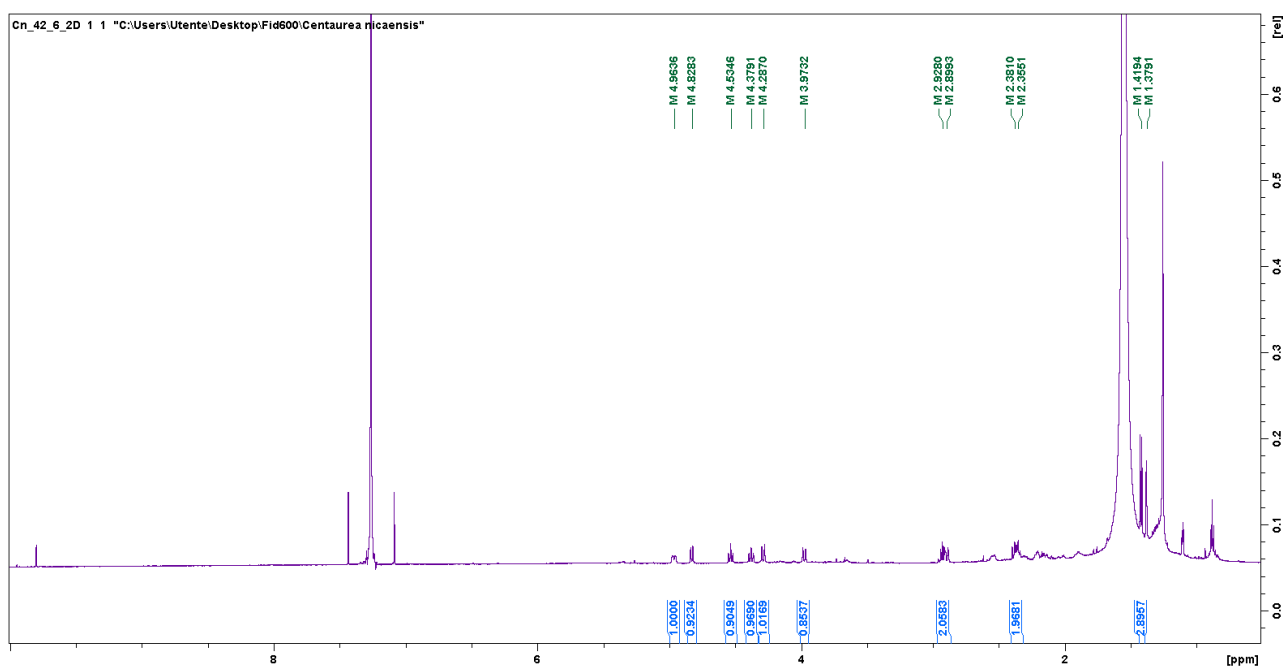


Figure S27: HSQC 2D NMR spectrum of compound 12 (CDCl₃, 600MHz)

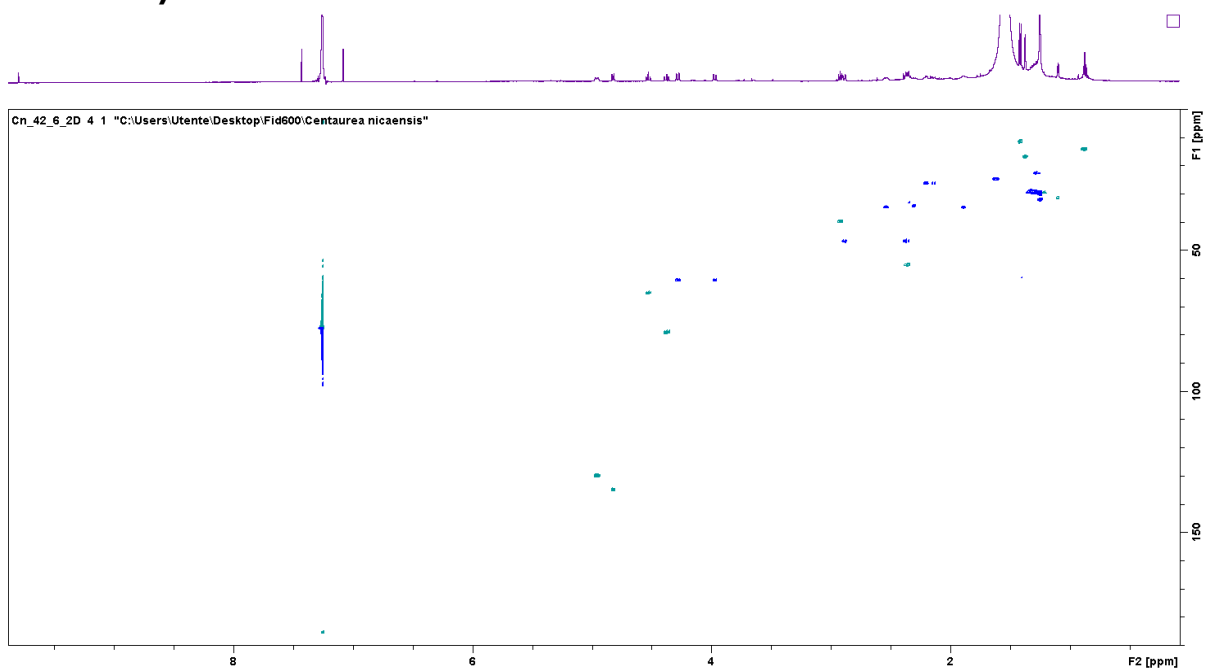


Figure S28: HMBC 2D NMR spectrum of compound 12 (CDCl₃, 600MHz)

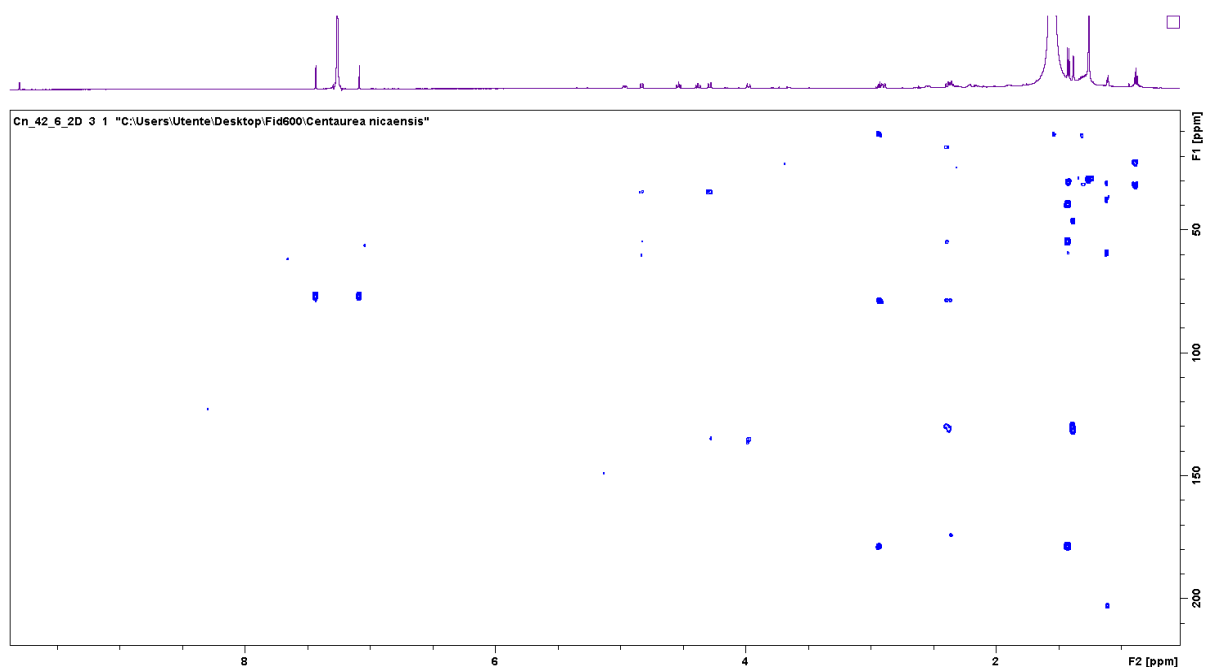


Figure S29: COSY 2D NMR spectrum of compound 12 (CDCl₃, 600MHz)

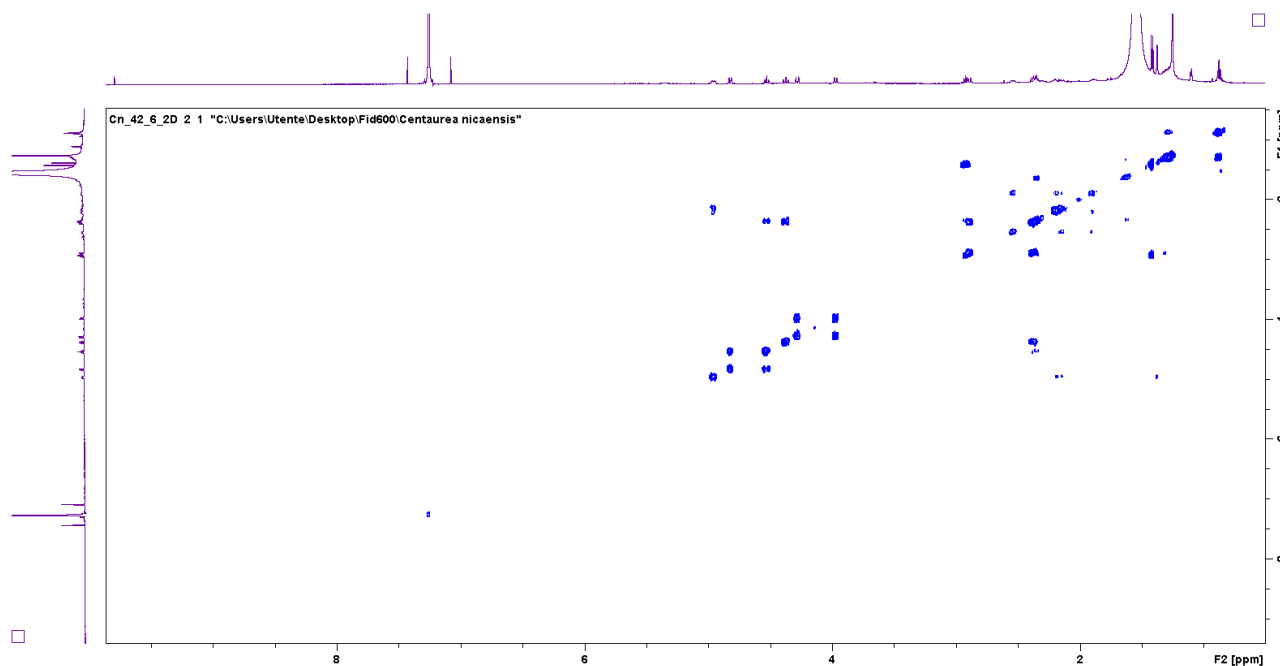


Figure S30: NOESY 2D NMR spectrum of compound 12 (CDCl₃, 600MHz)

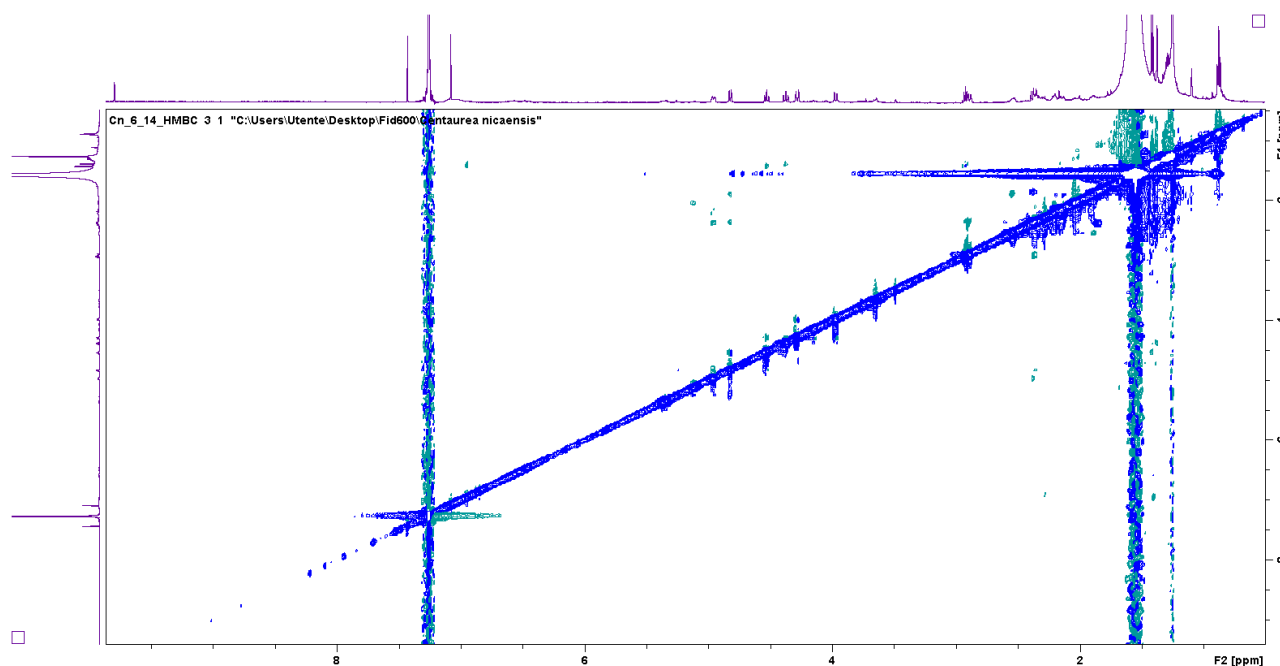


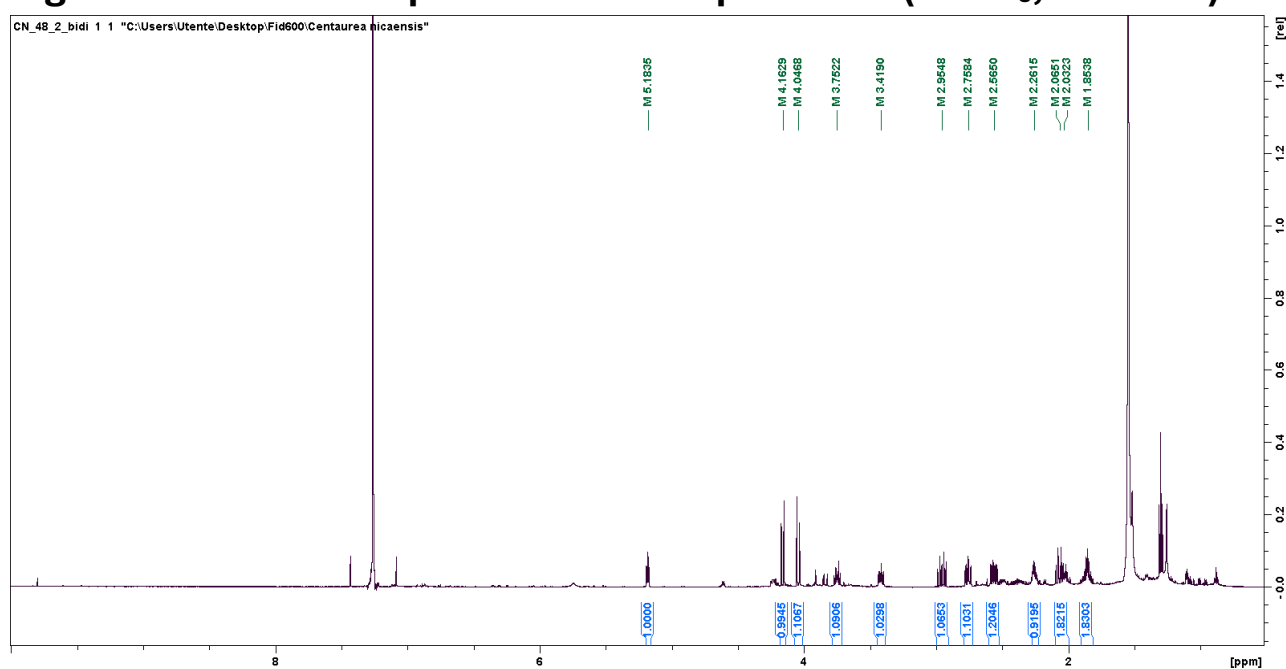
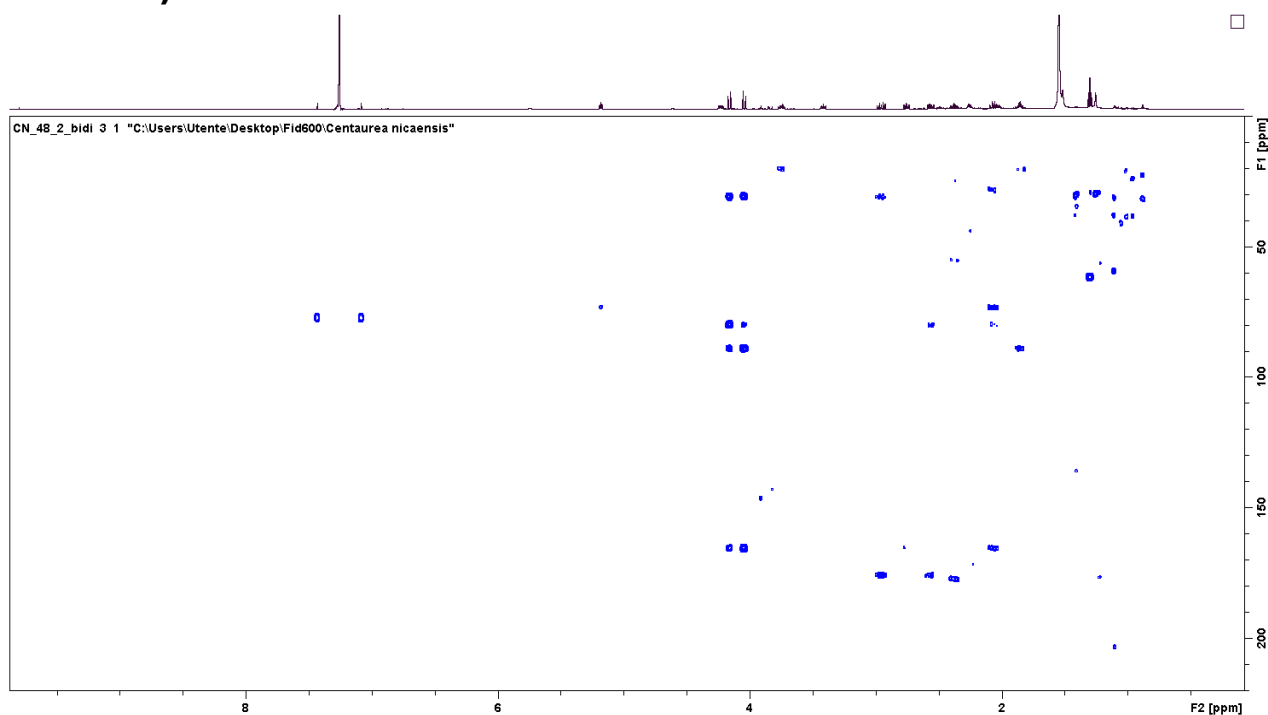
Figure S31: ^1H NMR spectrum of compound 24 (CDCl_3 , 700MHz)**Figure S32: HSQC 2D NMR spectrum of compound 24 (CDCl_3 , 600MHz)**

Figure S33: HMBC 2D NMR spectrum of compound 24 (CDCl₃, 600MHz)

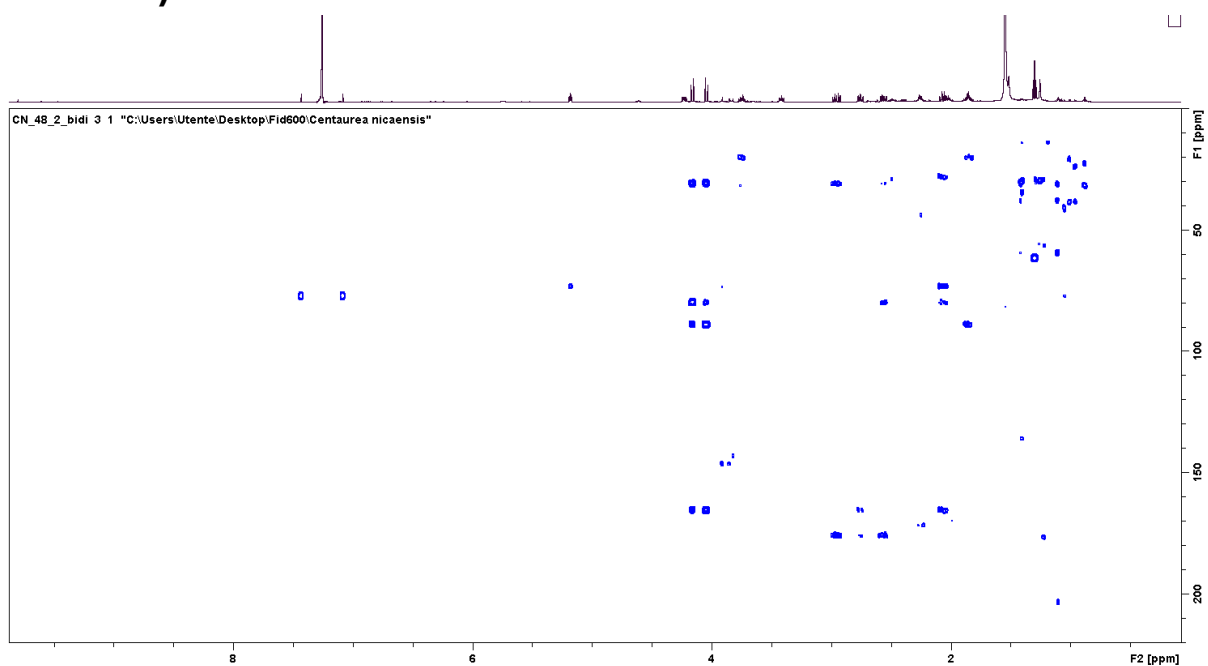


Figure S34: COSY 2D NMR spectrum of compound 24 (CDCl₃, 700MHz)

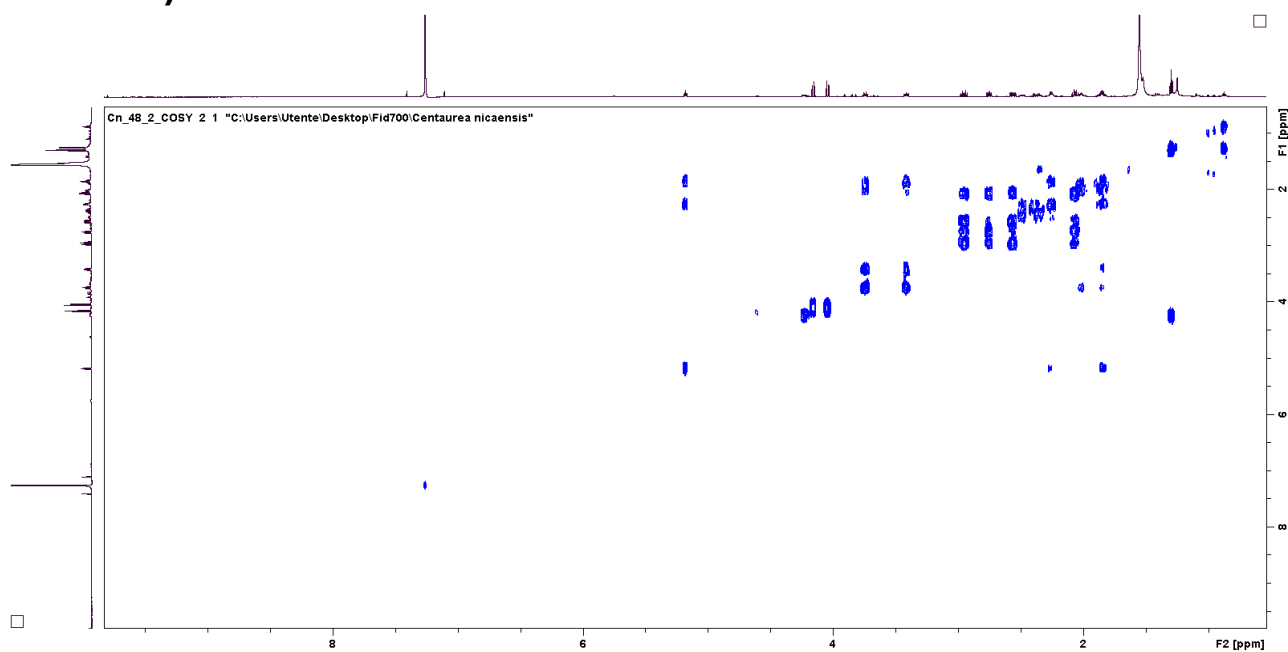
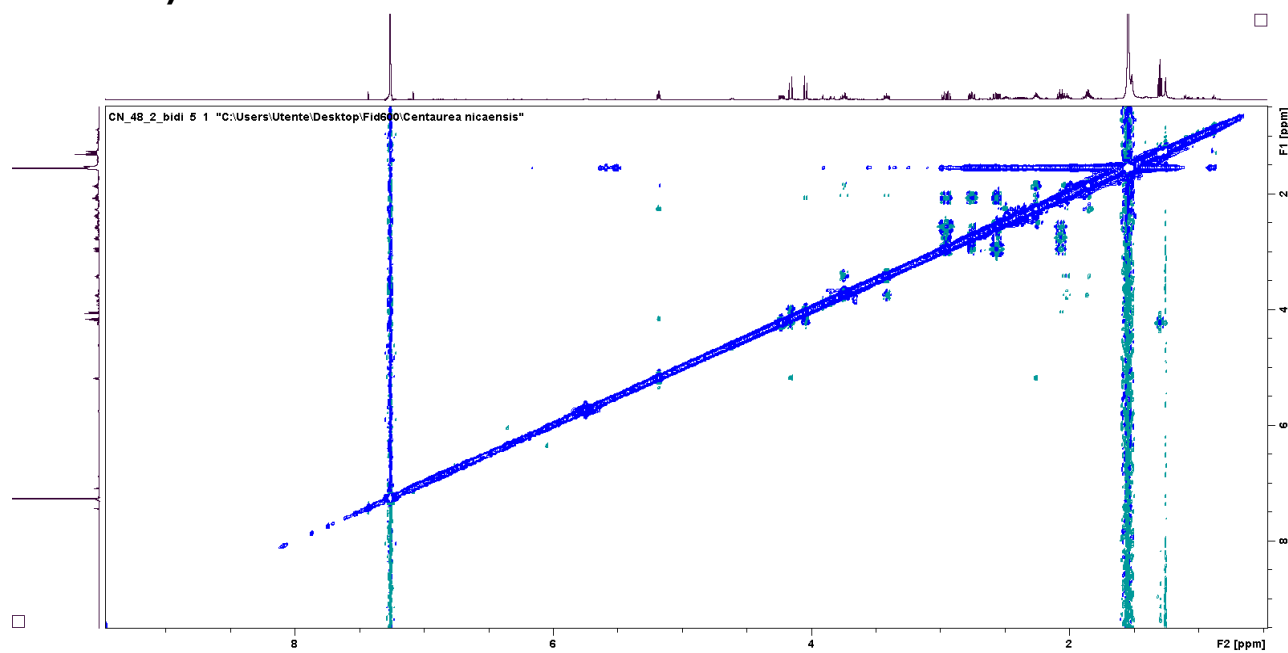


Figure S35: NOESY 2D NMR spectrum of compound 24 (CDCl₃, 600MHz)



CHAPTER 4: Alpha-glucosidase inhibitory profile of sesquiterpene lactones from *Centaurea* species

As mentioned in the previous chapters, the organic extracts of several *Centaurea* species displayed an inhibitory activity on α -glucosidase (Fattaheian-Dehkordi et al., 2021), but the bioactive compounds responsible for this activity have not been identified yet. In fact, besides flavonoids and phenolic acids, of which antioxidant and anti-inflammatory effects are well described in literature, the main metabolites of the bioactive *Centaurea* extracts are generally sesquiterpene lactones and lignans, but their role in MetS modulation has not been clearly understood yet. Alpha-glucosidase is a carbohydrate-hydrolase located in the brush border cells of the intestine, which acts on the O-glucosidic α (1 $\rightarrow$ 4) linkage. The enzyme breaks down starch and glucose oligosaccharides into single α -glucose monosaccharides. Its inhibition reduces the digestion rate of carbohydrates, which are not broken down into glucose molecules, reducing glucose levels in the blood. Therefore, the α -glucosidase inhibitors represent a class of drugs used to control the glycemic level in non-insulin-dependent diabetes mellitus (type 2), particularly postprandial hyperglycemia (Bashkin et al., 2021). Thus, in order to deepen the potential use of *Centaurea* species to treat metabolic syndrome and identify the bioactive compounds, a screening of the major sesquiterpenes lactones and lignans from the apolar extract of *C. kotschy* and *C. sicula* has been carried out. In this chapter, are reported the preliminary results obtained, but other analyses are still ongoing.

4.1 Alpha-glucosidase activity of sesquiterpene lactones

The inhibitory effect of the sesquiterpene lactones and lignans isolated from the aerial parts of the previously described *Centaurea* species against α -glucosidase was evaluated in comparison with the antidiabetic compound acarbose used as reference compound. The structures of the tested compounds are shown in **Figure 4. 1**. The α -glucosidase enzyme is involved in the carbohydrate metabolism and the release of glucose. Therefore, its inhibition is an effective strategy of delaying glucose absorption and lowering the postprandial blood glucose level, which can potentially suppress the progression of diabetes mellitus (Kilinc et al., 2020; Masullo et al., 2022). The results obtained showed the exhibited IC₅₀ values (**Table 4.1**). The two glycosylated lignans **17** and **18** showed the highest activity, displaying an IC₅₀ of 82.19

and 70.03 μM , respectively. Among the sesquiterpenes, compounds **5**, **8** and **15** displayed the highest activity, with IC_{50} values ranging from 141.94 to 144.02 μM , followed by compounds **2**, **11**, **12** and **14**.

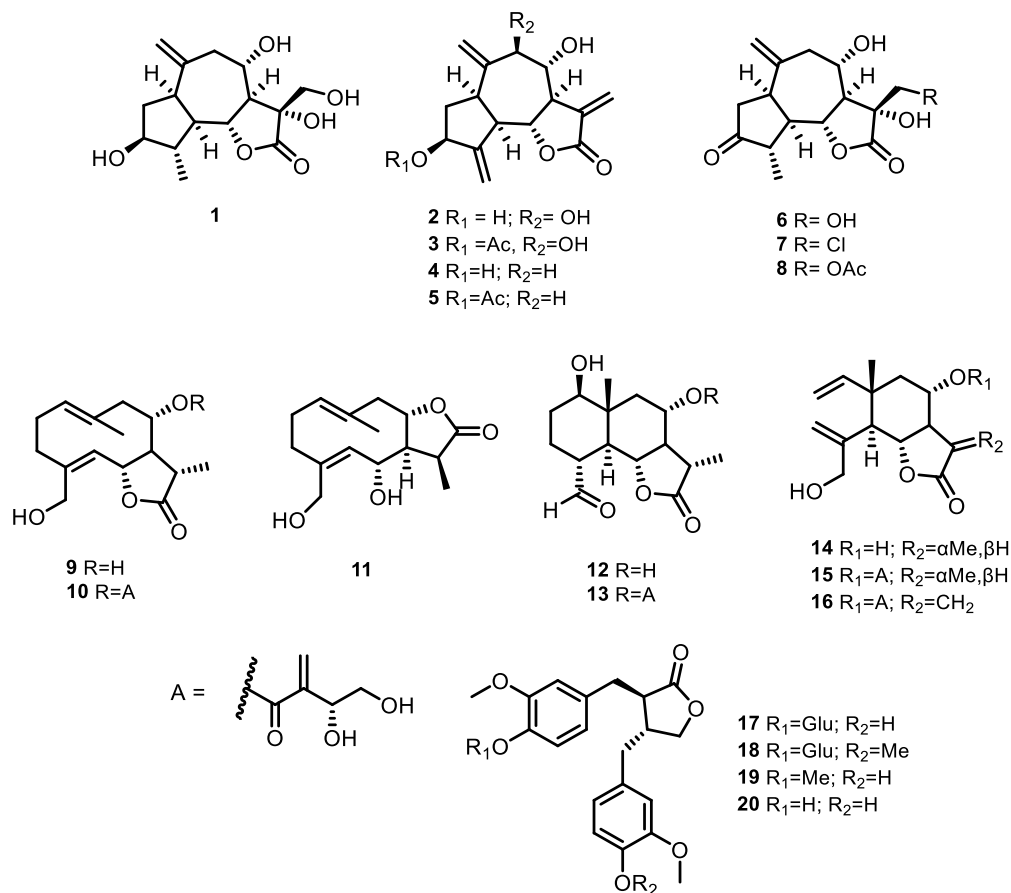


Figure 4. 1 Chemical structures of the tested sesquiterpenes and lignans isolated from *C. kotschyi* and *C. sicula*.

4.2 Molecular docking experiments

The detected promising inhibitory activities of the metabolites prompted us to investigate the basis behind the observed biological behaviors at a molecular level through molecular docking experiments. Here are reported the preliminary results of compounds **2**, **5**, **8**, **11**, **12** and **14**. Starting from the *Saccharomyces cerevisiae* α -glucosidase model, generated by us through homology modelling experiments (Masullo et al., 2022) since the absence of experimentally determined protein structure, molecular docking calculations were performed on the enzyme binding site involved in its catalytic activity. In particular, the putative interactions of the investigated compounds with D214, E276, D349 key residues, as accounted in previous studies

(Jabeen et al., 2016; Kilinc et al., 2020; Ma et al., 2015), were considered for the subsequent analysis. Specifically, the visual inspection of the sampled poses disclosed the ability of the tested compounds to establish extensive interaction networks with the above reported residues, selecting those featured contacts with D214 and/or E276 and/or D349 (**Figure 4.2**). All the six investigated compounds but **5** showed a H-bond with at least one of these aminoacids, while compound **5** showed polar contacts with E276 and H-bond with F157 and R312 (**Figure 4.3**, panel C), corroborating the promising data emerged from the obtained biological data (**Table 4.1**). In addition, the screened compounds showed an extensive interaction network with other residues belonging to the α -glucosidase catalytic site (e.g., F157, F158, H239, E304, P309, R312, Q350) (**Figure 4.4**) stabilizing their binding with the protein counterpart.

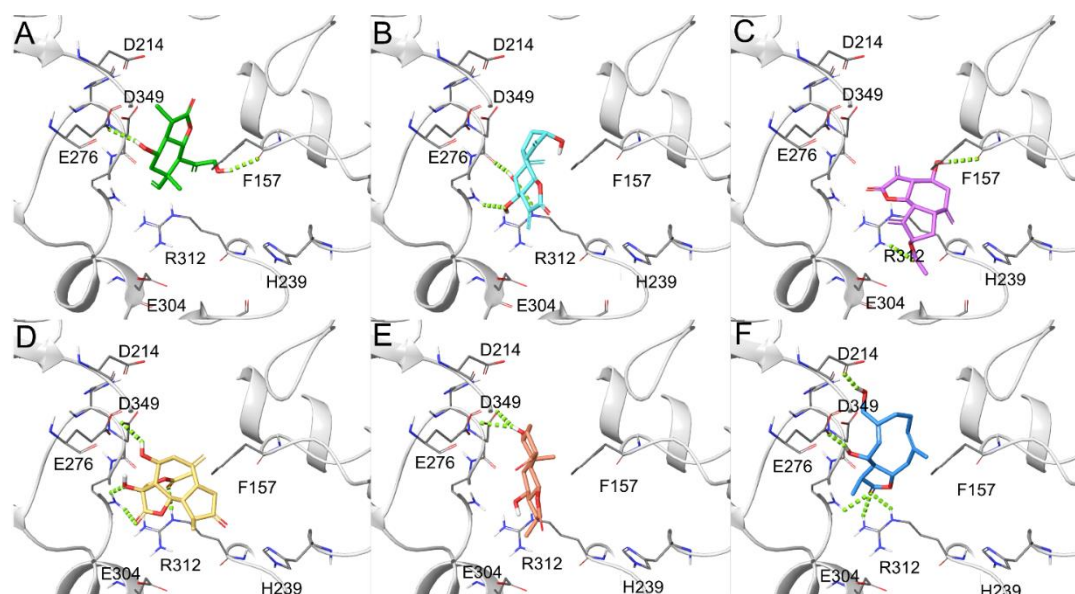


Figure 4.2 Docking poses of A) **14** (colored by atom type: C green, O red, polar H light grey); B) **2** (colored by atom type: C cyan, O red, polar H light grey); C) **5** (colored by atom type: C light violet, O red, polar H light grey); D) **8** (colored by atom type: C yellow, O red, polar H light grey); E) **12** (colored by atom type: C orange, O red, polar H light grey); F) **11** (colored by atom type: C light blue, O red, polar H light grey) in the *Saccharomyces cerevisiae* α -glucosidase binding site (represented with light grey ribbons; key residues are reported as sticks and colored by atom type: C grey, O red, N blue, polar H light grey). H-bonds are represented with dotted green lines.

4.3 Enzyme kinetics

To determine the mechanism of inhibition by the most active sesquiterpenes, compounds **2**, **5**, **12** and **14** against α -glucosidase enzyme, the analysis of the enzymatic kinetics was carried out using Lineweaver-Burk plots to determine the type of inhibition. The results are reported in

Figure 4.3 and **Table 4.2**. In this study, the maximum velocity ($V_{\max}$) value was determined as 204.08 $\mu\text{M}/\text{min}$ and Michaelis-Menten constant (K_m) as 1.59 mg/mL for 4-pNPG. The presence of **14** decreased $V_{\max}$ value to 147.05, but K_m value remained unaltered (1.60 mg/mL). This indicated that **14** inhibited α -glucosidase enzyme in a non-competitive manner. On the contrary **2** increased K_m value to 1.73 mg/mL but did not affect $V_{\max}$ (204.10 $\mu\text{M}/\text{min}$). This result showed that **2** was a competitive inhibitor of α -glucosidase. The kinetic analysis results showed that the inhibition mode of compounds **5** and **12** was a mixed-typed inhibition because the regression lines did not intersect both the x- and y-axes in the Lineweaver-Burk plot. In both cases, the same behaviors were observed resulting in a decrease of V_m and a change of K_m . The inhibition kinetics were determined by Dixon plots, which were obtained by plotting $1/V$ versus $[I]$ varying concentrations of 4-pNPG. Dixon plots gave families of lines passing through the same point at the second quadrant, giving the inhibition constant K_i . The K_i of the compounds against α -glucosidase were evaluated using Dixon plots and summarized in **Table 4.2**.

Table 4. 1 Glucosidase inhibitory activity of secondary metabolites by *Centaurea* spp. aerial parts

Compounds	IC ₅₀ $\pm$ SD (μM)
1	1000.00 $\pm$ 7.14
2	149.46 $\pm$ 1.97
3	197.53 $\pm$ 3.90
4	282.33 $\pm$ 36.60
5	141.94 $\pm$ 8.35
6	236.59 $\pm$ 17.70
7	188.21 $\pm$ 7.51
8	144.67 $\pm$ 12.50
9	577.78 $\pm$ 17.00
10	163.21 $\pm$ 3.39
11	180.94 $\pm$ 28.10
12	167.05 $\pm$ 19.90
13	567.72 $\pm$ 9.24
14	149.81 $\pm$ 14.81
15	144.02 $\pm$ 11.90
16	234.71 $\pm$ 6.85
17	82.19 $\pm$ 2.15
18	70.03 $\pm$ 2.34
19	133.01 $\pm$ 2.71
20	190.50 $\pm$ 2.70
Acarbose	43.34 $\pm$ 4.46

Data are expressed as the mean values $\pm$ SD of three experiments.

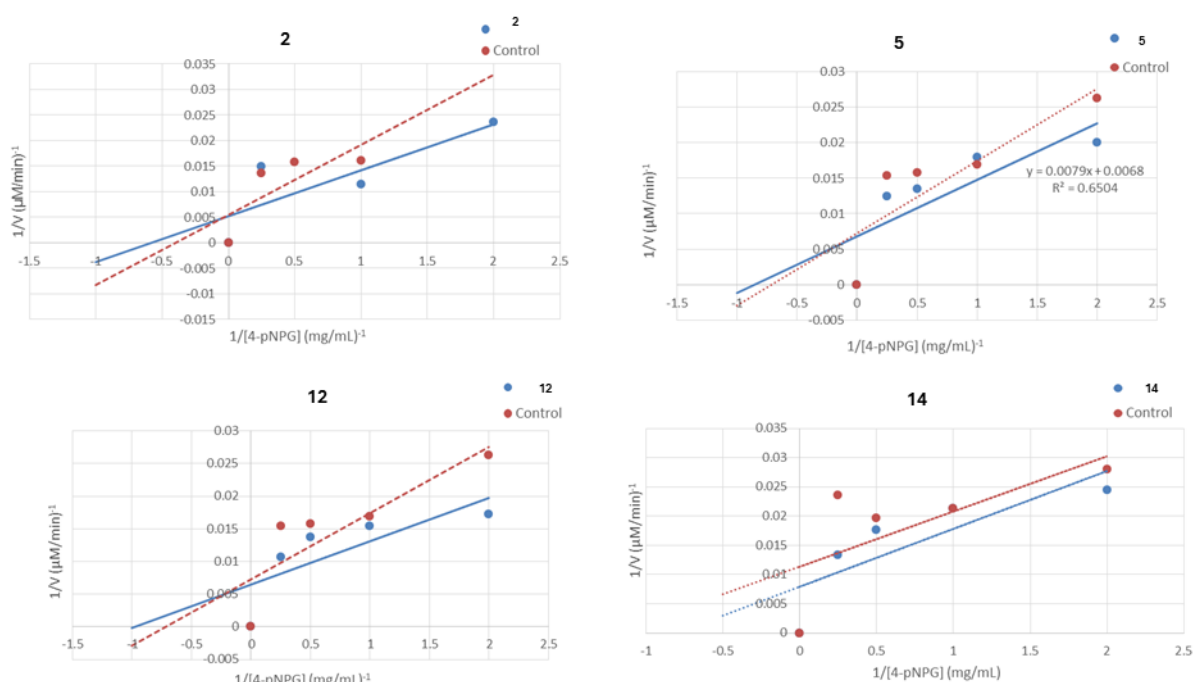


Figure 4.3 Lineweaver plots for the inhibition of α -glucosidase by **2**, **5**, **12** and **14**.

Table 4. 2 Kinetic parameters of α -glucosidase in the presence of the compounds.

Inhibitors	Inhibitions type	$V_{\max}$ ($\mu\text{M}/\text{min}$) K_m (mg/mL)	K_i values ($\mu\text{g}/\text{min}$)
2	Competitive	204.1, 1.73	78.21
5	Mix	147.06, 1.16	289.00
12	Mix	151.51, 0.98	300.21
14	Non-competitive	147.05, 1.60	74.11

4.4 Conclusions

The α -glucosidase inhibitory activity reported for organic extracts of *Centaurea* sp. aerial parts prompted us to disclose the bioactive constituents and deepen in their interaction with the target. Therefore, given the library of sesquiterpenes and lignans isolated in the previously described *Centaurea* species, we tested the main metabolites isolated from *C. kotschy* and *C. sicula*. These preliminary results showed that lignans are highly active, with an IC_{50} between 82.19 and 190.06 μM . Their interaction with the enzyme counterpart is currently being evaluated. However, compounds **17** and **18**, both bearing a sugar unit, are the most active, suggesting that, as expected, this moiety plays an important role in the interaction with the target. Among the sesquiterpenes, the guaianolide **5** showed the highest activity. Moreover, in the light of the results obtained in the previous studies, it is interesting to note that guaianolide

2 and germacranolide **11**, isolated from *C. kotschy* and *C. sicula* respectively, showed a multi-target mechanism in blood glucose control, modulating both GLUT and α -glucosidase.

4.5 Experimental Section

4.5.1 Plant Material and isolation of compounds

The compounds and lignans reported were isolated as described in the previous chapters (Serino et al., 2024; Serino et al., 2022).

4.5.2 α -Glucosidase Inhibition Assay

α -Glucosidase inhibitory activity was evaluated as previously reported (Kilinc et al., 2020). Briefly, 0.1 M phosphate buffer at pH 7.0 (150 μ L) was added to each well; successively 10 μ L of samples, dissolved in MeOH at different concentrations (final concentrations 8, 12, 16 μ g/mL) were added. The reaction was initiated by the addition of 15 μ L of the α -glucosidase enzyme (from *Saccharomyces cerevisiae*) (2 U/ml) in each well, and the plate was incubated for 5 minutes at 37°C. After 5 minutes, 75 μ L of the substrate (2.5 mM) 4-p-nitrophenyl- α -D-glucopyranoside (4-pNPG) were added, and successively the plate was incubated for 10 minutes at 37°. The absorbance was measured at 405 nm before (T0') and after the incubation with the enzyme (T10'). The positive control was acarbose. Negative control absorbance (phosphate buffer in place of the sample) was also recorded. Inhibition of the enzyme was calculated, and the results were expressed as IC₅₀ values. All experiments were carried out in triplicate. IC₅₀ values were calculated through non-linear regression using Graphpad Prism 5 Software and expressed as means $\pm$ SD (standard deviation).

4.5.3 Inhibitory kinetic analysis

Kinetic analysis of α -glucosidase enzyme for **2**, **5**, **12** and **14** was carried out using the Lineweaver-Burk plot. The experiment was conducted at different 4-pNPG concentrations (0.25, 0.50, 1.00 and 2.00 mg/mL) in the absence and presence of the sample solution. The data obtained were plotted as 1/activity (1/V) against 1/substrate concentration (1/[S]), according to Lineweaver-Burk graph and the type of inhibition for compounds **2**, **5**, **12** and **14** (Şöhretoğlu et al., 2018) was determined. In addition, inhibition constant (K_i) was deduced from the Dixon plot which is a graphical method (plot of 1/enzyme velocity (1/V) versus inhibitor

concentration [I] with varying concentrations of the substrate (0.25, 1.00 and 2.00 mg/mL) for each inhibitor.

4.5.4 Molecular modeling

To be consistent with the performed biological assays (see α -Glucosidase Inhibition Assay paragraph), the 3D protein model of α -glucosidase from *Saccharomyces cerevisiae* was considered for molecular modeling experiments. In details, the 3D protein model of *Saccharomyces cerevisiae* α -glucosidase previously built through homology modeling (Masullo, Lauro, Cerulli, Bifulco & Piacente, 2022) was used for molecular docking calculations, focusing on the enzyme region responsible of the catalytic activity and comprising D214, E276, D349 key residues (Jabeen et al., 2016; Ma et al., 2015; Masullo et al., 2022). The library of investigated sesquiterpene lactones (see Results and Discussion sections) was prepared with LigPrep software (Schrodinger Suite) (LigPrep, 2023). All the possible tautomers and protonation states ($\text{pH} = 7.4 \pm 1.0$) for each compound were generated and then minimised using the OPLS 2005 force field. Glide software (Schrödinger Suite) was used for molecular docking experiments in the Standard Precision [SP] mode (Friesner et al., 2004; Friesner et al., 2006; Glide, 2023; Halgren et al., 2004). In details, the following parameters were set: 10,000 poses for the initial phase for energy minimisation, 5.0 kcal/mol as energy window for ring sampling, 500.0 as scoring window for keeping the initial poses, 0.8 as scaling factor of related to van der Waals radii, 0.15 as a partial charge cutoff, basing on a 0.5 kcal/mol rejection cutoff for the obtained minimised poses. Eventually, 50 maximum number of poses for each compound were saved in the output file.

References

- Bashkin, A., Ghanim, M., Abu-Farich, B., Rayan, M., Miari, R., Srouji, S., Rayan, A. and Falah, M., **2021**. Forty-one plant extracts screened for dual antidiabetic and antioxidant functions: Evaluating the types of correlation between α -amylase inhibition and free radical scavenging. *Molecules*, 26(2), p.317.
- Fattaheian-Dehkordi, S., Hojjatifard, R., Saeedi, M. and Khanavi, M., **2021**. A review on antidiabetic activity of *Centaurea* spp.: A new approach for developing herbal remedies. *Evidence-Based Complementary and Alternative Medicine*, 2021(1), p.5587938.
- Formisano, C., Sirignano, C., Rigano, D., Chianese, G., Zengin, G., Seo, E.J., Efferth, T. and Taglialatela-Scafati, O., **2017**. Antiproliferative activity against leukemia cells of sesquiterpene lactones from the Turkish endemic plant *Centaurea drabifolia* subsp. *detonsa*. *Fitoterapia*, 120, pp.98-102.
- Friesner, R.A., Banks, J.L., Murphy, R.B., Halgren, T.A., Klicic, J.J., Mainz, D.T., Repasky, M.P., Knoll, E.H., Shelley, M., Perry, J.K. and Shaw, D.E., **2004**. Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy. *Journal of medicinal chemistry*, 47(7), pp.1739-1749.
- Friesner, R.A., Murphy, R.B., Repasky, M.P., Frye, L.L., Greenwood, J.R., Halgren, T.A., Sanschagrin, P.C. and Mainz, D.T., **2006**. Extra precision glide: Docking and scoring incorporating a model of hydrophobic enclosure for protein– ligand complexes. *Journal of medicinal chemistry*, 49(21), pp.6177-6196.
- Glide, S., LLC, New York, NY. **2023**
- Halgren, T.A., Murphy, R.B., Friesner, R.A., Beard, H.S., Frye, L.L., Pollard, W.T. and Banks, J.L., **2004**. Glide: a new approach for rapid, accurate docking and scoring. 2. Enrichment factors in database screening. *Journal of medicinal chemistry*, 47(7), pp.1750-1759.
- Jabeen, F., Shehzadi, S.A., Fatmi, M.Q., Shaheen, S., Iqbal, L., Afza, N., Panda, S.S. and Ansari, F.L., **2016**. Synthesis, in vitro and computational studies of 1, 4-disubstituted 1, 2, 3-triazoles as potential α -glucosidase inhibitors. *Bioorganic & Medicinal Chemistry Letters*, 26(3), pp.1029-1038.

- Kılınç, H., Masullo, M., D'Urso, G., Karayildirim, T., Alankus, O. and Piacente, S., **2020**. Phytochemical investigation of *Scabiosa sicula* guided by a preliminary HPLC-ESIMSⁿ profiling. *Phytochemistry*, *174*, p.112350.
- LigPrep, S. R., LLC, New York, NY. **2023**
- Ma, H., Wang, L., Niesen, D.B., Cai, A., Cho, B.P., Tan, W., Gu, Q., Xu, J. and Seeram, N.P., **2015**. Structure activity related, mechanistic, and modeling studies of gallotannins containing a glucitol-core and α -glucosidase. *RSC advances*, *5*(130), pp.107904-107915.
- Masullo, M., Lauro, G., Cerulli, A., Bifulco, G. and Piacente, S., **2022**. *Corylus avellana*: A Source of Diarylheptanoids With α -Glucosidase Inhibitory Activity Evaluated by in vitro and in silico Studies. *Frontiers in Plant Science*, *13*, p.805660.
- Serino, E., Chianese, G., Musto, G., Zengin, G., Rigano, D., Stornaiuolo, M., Formisano, C. and Taglialatela-Scafati, O., **2022**. Guaiane-rich phytochemical profile of *Centaurea kotschyi* subsp. *persica* (Boiss.) Wagenitz and identification of hypoglycaemic metabolites. *Phytochemistry*, *199*, p.113189.
- Serino, E., Rigano, D., Bruno, M., Pastore, A., Stornaiuolo, M., Formisano, C. and Taglialatela-Scafati, O., **2024**. Glucose Uptake-Stimulating Metabolites from Aerial Parts of *Centaurea sicula*. *Journal of Natural Products*, *87*(4), pp.1179-1186.
- Şöhretoğlu, D., Sari, S., Şoral, M., Barut, B., Özel, A. and Liptaj, T., **2018**. Potential of *Potentilla inclinata* and its polyphenolic compounds in α -glucosidase inhibition: Kinetics and interaction mechanism merged with docking simulations. *International journal of biological macromolecules*, *108*, pp.81-87.

CHAPTER 5: NMR-based profiling of *Centaurea patula* sesquiterpenoids and their α -glucosidase inhibitory activity

Centaurea patula is native to the Central Anatolian region of Turkey (Zengin et al., 2016) and has also been reported in Iran (Negaresht, 2018). According to the phylogenetic classification, *C. patula* belongs to the Circum-Mediterranean and Eurosiberian clade (Jacea clade), which is quite distant from the previously studied *C. kotschy* (Serino et al., 2022) and *C. drabifolia* (Formisano et al., 2017), which belong to the Eastern Mediterranean clade (Garcia-Jacaset et al., 2006). This could be reflected in a completely different phytochemical profile, as seen in the previous study for *C. sicula*, which belongs to the section *Seridia* (Serino et al., 2023). Interestingly, *C. patula* shows a unique feature within its genus, as its chromosome number is 7, the lowest found in the genus, reflecting a highly divergent group (Garcia-Jacaset et al., 1996). However, there is little information in the literature on the metabolites of this plant, as only the fatty acid profile (Tekeli et al., 2013) and essential oil composition (Zengin et al., 2016) have been described. More recently, the methanolic extracts of some Turkish *Centaurea* species, including *C. patula* and *C. kotschy* var. *perisca*, were analysed by LC-MS, leading to the annotation of some phenolic compounds that are recurrent among *Centaurea* species (Zengin et al., 2019). Nevertheless, the sesquiterpene profile of this plant has not yet been investigated. *C. patula* extracts have been evaluated for their antioxidant activity (Zengin et al., 2010) and antiproliferative effects (Russo et al., 2016). Furthermore, its aerial organic extracts have displayed inhibitory effect on α -glucosidase enzyme (Fattaheian-Dehkordi et al., 2021). Therefore, given our interest in investigating less studied *Centaurea* species with anti-diabetic potential, we carried out an NMR profiling of the organic aerial parts extract, focusing our attention on the sesquiterpene composition. The activity on α -glucosidase of the pure metabolites was evaluated in order to identify the compounds responsible for the hypoglycaemic activity of the extracts.

5.1 Phytochemical composition of *C. patula*

The dried aerial parts of *Centaurea patula* were defatted with n-hexane and successively extracted with CH₂Cl₂. The CH₂Cl₂ extract was first fractionated by MPLC-DAD and the obtained fractions were preliminarily analysed by ¹H NMR, which revealed a rich abundance of signals attributable to different families of sesquiterpenoids. The sesquiterpene-containing fractions were successively purified by MPLC and HPLC, yielding 20 pure compounds (**Figure 5.1**). *C. patula* has a peculiar sesquiterpenoid profile compared to those previously studied in our laboratory (Formisano et al., 2017; Serino et al., 2022; Serino et al., 2024). A preliminary inspection of these compounds revealed that *C. patula* has a peculiar sesquiterpenoid profile compared to those previously studied in our laboratory (Formisano et al., 2017; Serino et al., 2022; Serino et al., 2024). Indeed, *C. patula* biosynthesizes not only variously decorated guaianolides (**1-13**), but also highly unusual elemanolides and germacranolides (**14-19**). Interestingly, no eudesmanes have been isolated from this species, while these sesquiterpenoids were the main constituents of *C. sicula* (Serino et al., 2024). The known metabolites were identified by comparison of their spectral data with those reported in the literature as 4β-15-dihydro-3-dehydrosolstitialin A (**1**) (Medjroubi et al., 2003), desacylcynaropicrin (**2**) (Choi et al., 2005), cynaropicrin (**3**) (Choi et al., 2005), repdiolidetriol (**4**) (Mohamed et al., 2021), babylin B (**5**) (Sen et al., 2017), cebellin G (**6**) (Singh et al., 1988), chlorojanerin (**7**) (Saklani et al., 2011), cebellin J (**8**) (González et al., 1974), 15-O-deacetyl rhaposerin (**10**) (Bruno et al., 2005), rhaposerine (**11**) (Bruno et al., 2005), centaurepensin (**12**) (Bruno et al., 2005) and hololeucin (**13**), which is a peculiar sesquiterpene lactone exhibiting a carbonate cycle described for the first time in *Centaurea hololeuca* Boiss (Rosselli et al., 2006). Finally, the elemane dehydromelitensin (**14**) (Sen et al., 2017) and the monoterpene daphnauranin E (**20**) (Huang et al., 2018) were identified. The guaianolide cebellin Q (**9**), the elemanolides **15-17** and the germacranolides patulosins A-B (**18-19**) were unprecedented in literature, therefore a whole set of 2D NMR experiments (COSY, HSQC, HMBC and NOESY) and HR-ESIMS analyses were performed in order to obtain a complete characterization of these new metabolites.

PART 2

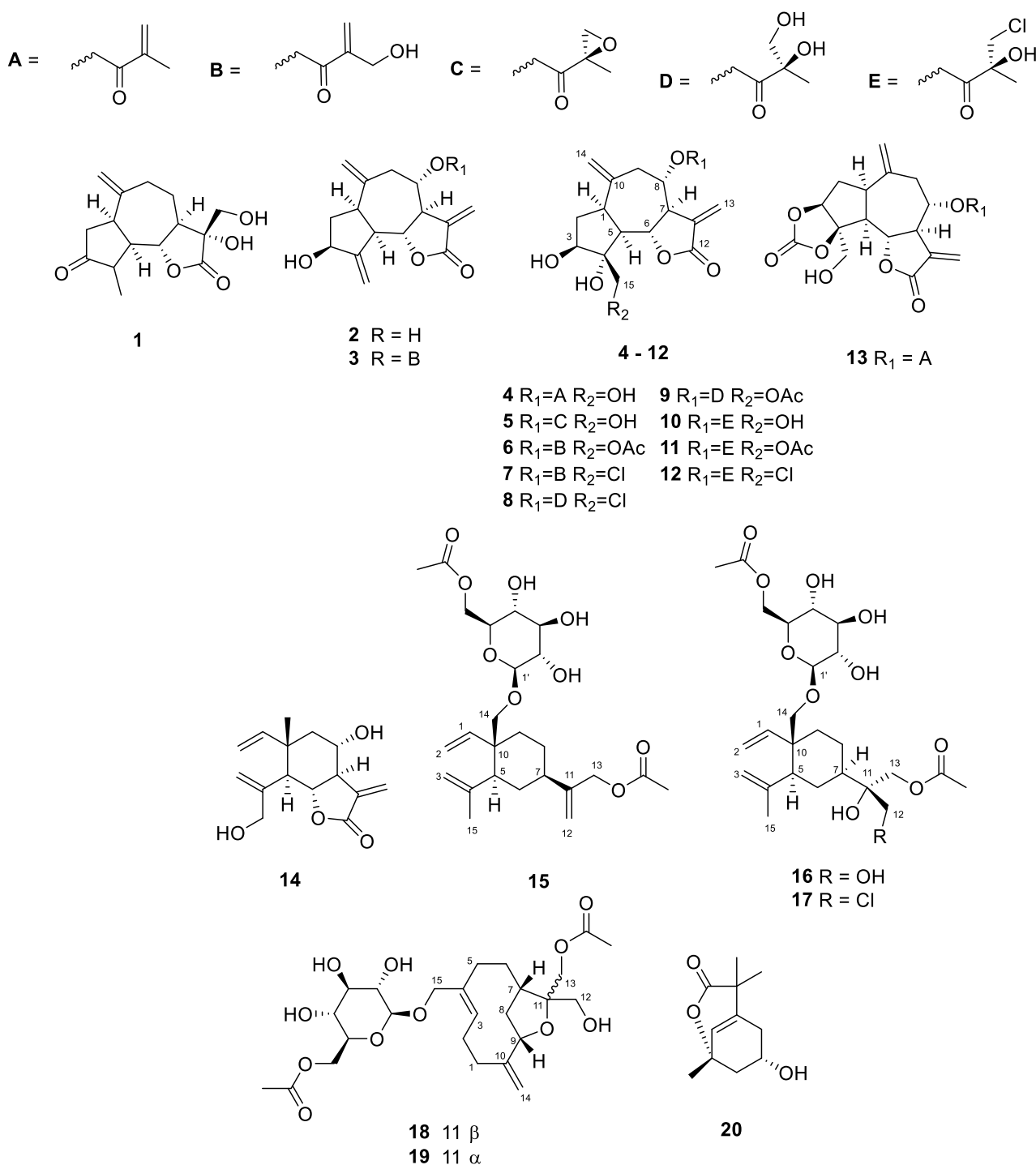


Figure 5. 1 Chemical structures of the terpenoids isolated from *C. patula*

The ¹H NMR spectrum of cebellin Q (**9**) was very similar the one reported in the literature for cebellin J (**8**), co-occurring in this plant, the differences being confined to the presence of a methyl singlet at δ_H 2.17, compatible with the presence of an acetyl group. This hypothesis was supported by HRESIMS, which gave a molecular formula of C₂₁H₂₈O₁₀, confirming the presence of an additional acetyl group and the absence of the chlorine atom compared to that of **8**.

Cebellin Q (**9**) was therefore likely to have replaced with an acetyl group the chlorine attached at C-15. The COSY 2D NMR experiment allowed us to build a single spin system, while all the protons were assigned to their directly linked carbon via HSQC. The HMBC spectrum guided us in the assembly of the structure; in particular the diagnostic correlations of H₂-15 (with C-3, C-4 and C-5) of H₂-14 (with C-9, C-10 and C-1) and of H₂-13 (with C-7, C-11 and C-12) allowed us to build up the guaiane skeleton. The diagnostic correlations between the H₂-15 to the acetyl carbon at δ_c 172.8, which unambiguously placed the acetyl group, while the H-8 (δ_H 5.14)/C-1' cross-peak confirmed the position of the lateral side chain. The NOESY correlations H-5/H-7, H-6/H-8 paralleled those of **8**, suggesting that **9** shared the same stereochemical features. Given the evident relationship between the new cebellin Q (**9**) and cebellin J, we assume that the two compounds share the same absolute configuration at the common stereogenic centers.

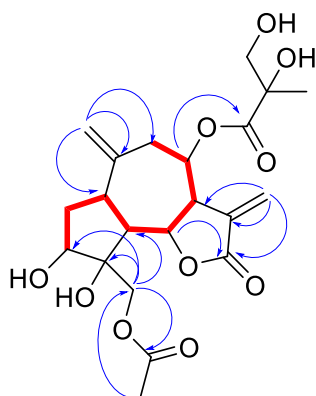


Figure 5. 2 COSY (red bold) and diagnostic HMBC correlation (blue arrow) of compound **9** (planar structure).

Compound **15** showed a HRESIMS [M+Na]⁺ peak at m/z 505.2473, indicating C₂₅H₃₈O₉ as molecular formula. The ¹H NMR of **15** displayed the typical signals of an elemene skeleton in the low-field region (δ_H 5.79, dd; δ_H 5.11, d; δ_H 4.97, d; δ_H 4.85, s and δ_H 4.66, s). In the mid-field region, two 2H-integrating singlets resonating at δ_H 5.03 and 4.60 were present, along with two coupled doublets at δ_H 4.00 and 3.57, and a series of multiplets between δ_H 4.63 and δ_H 3.17. The high-field region showed only three methyl singlets at δ_H 1.73, 2.04 and 2.07, suggesting the presence of two acetyl groups, and a series of multiplets between δ_H 2.18 and 1.33. A combined analysis of the 2D NMR experiments (COSY, HSQC and HMBC) enabled to build-up the planar structure of **15**, which is made of a peculiar elemanolide scaffold decorated with a sugar unit and two acetyl moieties. The 2D NMR COSY spectrum led us to arrange the

protons signals in three different spin systems, two belonging to the elemene unit and another to a sugar moiety. The presence of a sugar unit was supported by the HSQC spectrum, that showed the correlation of the doublet at δ_{H} 4.22 to the anomeric C-1' (δ_{C} 103.5). The HSQC spectrum also allowed us to classify the three uncoupled methylene groups as one sp^2 methylene (δ_{H} 5.03/ δ_{C} 110.6) and two sp^3 oxymethylenes. The HMBC cross-peaks $\text{H}_3\text{-15/C-3}$ and C-5 confirmed the position of the methyl group on the unprotonated carbon C-4, while the HMBC correlations of $\text{H}_2\text{-14}$ with C-1, C-9 and C-10 attached the oxymethylene at C-10. The nature of the substituent on the methine C-7 was deduced from the cross-peaks of $\text{H}_2\text{-12}$ with C-7 and C-13, along with those of $\text{H}_2\text{-13}$ with C-7, C-11 and C-12. Finally, the sugar moiety was placed in C-14 basing on the key HMBC correlation $\text{H}_2\text{-14/C-1'}$, while the two acetyl groups were placed at C-13 and C-6' basing on the correlations of $\text{H}_2\text{-13}$ and $\text{H}_2\text{-6'}$ with the corresponding ester carbonyls (δ_{C} 171.0). 13,14-Dihydroxyelemanolides are extremely rare in *Centaurea* plants and, to the best of our knowledge, only a single example has been reported before, in *C. athoa* DC (Demir et al., 2016). The stereochemical features of **15** were investigated through a detailed analysis of the NOESY spectrum showing that $\text{H}_3\text{-15}$ and H-14a and $\text{H}_3\text{-7}$ and H-5 are co-facial, in line with the configuration assignment reported in the literature for its analogues isolated in other *Centaurea* species. Considering the biogenetic relationship between **15** and the co-present dehydromelitensin (**14**), whose absolute configuration has been reported in the literature (Sen et al., 2017), we can postulate that **15** shares the same absolute configuration of **14**. Moreover, given the proton-proton coupling constants and the ^{13}C NMR values deduced by the HSQC spectrum, we assigned the sugar as a β -glucopyranosyl unit.

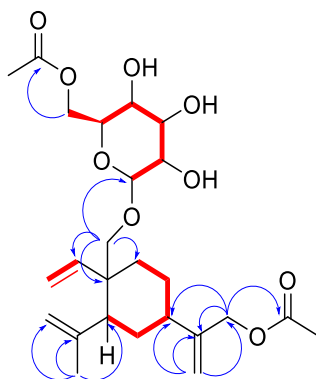


Figure 5.3 COSY (red bold) and HMBC diagnostic correlations (blue arrow) of compound **15** (planar structure).

The ^1H NMR of compound **16** closely resembled the one of **15**. The most remarkable difference was the absence of signal at δ_{H} 5.03, while two mutually coupled doublets were

detected at δ_{H} 3.60 and 3.57, suggesting the oxidation of the exomethylene moiety. This hypothesis was in agreement with the HRESIMS data which showed $\text{C}_{25}\text{H}_{40}\text{O}_{11}$ as molecular formula, indicating the presence of two additional oxygen atoms compared to **15** and of six degrees of unsaturation, one less than the previously reported **15**. The 2D NMR spectra allowed the structural assignment of **16** confirming that the single difference with **15** was confined in the moiety attached at C-7. In particular, the HSQC cross-peaks revealed that the proton signals at δ_{H} 3.60, δ_{H} 3.57 are linked to a carbon atom resonating at δ_{C} 62.6, thus indicating its hydroxymethylene nature, that was unambiguously attached at C-11 basing on the HMBC cross peaks of H_2 -12 with C-7, the oxygenated C-11 (δ_{C} 74.4) and C-13. The position and nature of the sugar moiety and of the acetyl groups were as described for compound **15**. Similarly, the NOESY spectrum of **16** paralleled that of **15**, but the relative configuration at the additional stereogenic center C-11 of **15** remained to be determined. This stereochemical problem was faced by calculating the ^{13}C NMR spectra for the 11*S* and 11*R* diastereomers. A preliminary conformational search by Simulated Annealing (INSIGHT II package), followed by DFT calculations using the Gaussian09 package, revealed, for both diastereomers, only two conformers accounted for 91.5% (11*R*) and 82.0% (11*S*) of the total population (**Figure 5.4**).

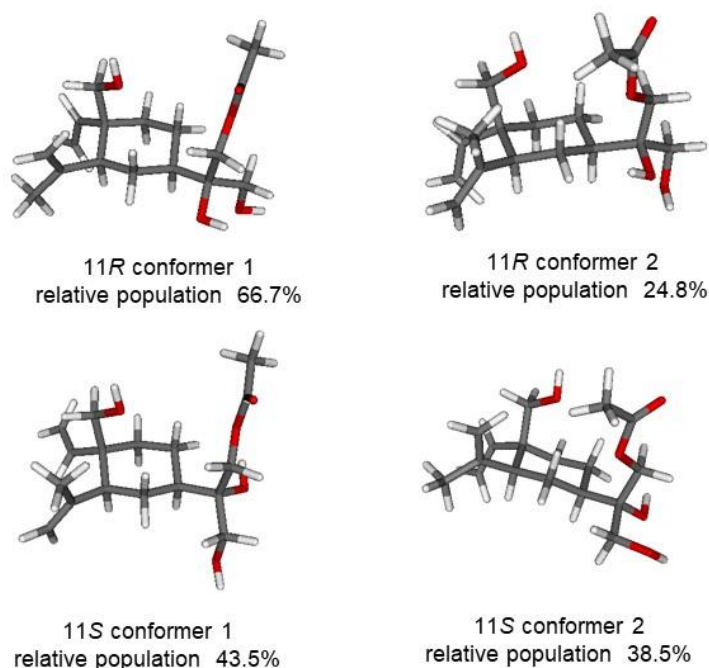


Figure 5.4 Most populated conformers for 11*R* and 11*S* diastereomers of **16**.

GIAO ^{13}C calculations were performed using the mPW1PW91 functional and 6-31G(d,p) basis set using the geometry previously optimized at the mPW1PW91/6-311G(d) level as input. After Boltzmann average, DP4+ probabilities were computed and clearly revealed the 11*S* as the most likely (DP4+ = 100%) configuration at C-11. Details about the calculations are given in the Appendix C.

Compound **17** had $\text{C}_{25}\text{H}_{39}\text{ClO}_{10}$ as molecular formula, as deduced from the HR-ESIMS peak at m/z 557.2189 ($\text{C}_{25}\text{H}_{39}^{35}\text{ClO}_{10}$). The presence of the chlorine atom was supported by the isotopic pattern showing an $[\text{M}+2]^+$ peak in 3:1 ratio. The close resemblance between the ^1H NMR spectral data of **17** and those of **16**, and their molecular formula comparison, strongly suggested that these two compounds only differ for the presence of a chlorine in place of an -OH group. A detailed analysis of the 2D NMR data of **17** confirmed this hypothesis, and COSY, HSQC and HMBC spectra closely paralleled those of **16** with the single exception of the marked upfield-shift of C-13, from δ_{C} 62.6 of **16** to δ_{C} 46.5 of **17**, fully supporting the linkage with a chlorine atom. Given the parallel NOESY correlations, we can assume that these **16** and **17** share also the same stereochemical features. Similar computational calculations established the 11*R* configuration. Computational details of the calculations are reported in the Appendix

C. Compounds **16** and **17** can be viewed as the products of an alternative 11,12-epoxide opening with H₂O and Cl⁻ nucleophilic attack.

Patulosin A (**18**) exhibited an HR-ESIMS [M+Na]⁺ peak at *m/z* 537.2348, implying the molecular formula C₂₅H₃₈O₁₁ in agreement with seven unsaturation degrees. The ¹H NMR spectrum of **18** showed a series of multiplets in the high-field region, along with two acetyl signals resonating at δ_H 2.03 and 2.09. The midfield region appeared crowded with multiplets which suggested the presence of a sugar moiety, moreover a broad doublet at δ_H 5.64 and two singlets at δ_H 5.42 and 5.10 were detected. A combined analysis of the COSY and HSQC spectra led us to the identification of four uncoupled methylene groups, one of them attributable to an exomethylene moiety (δ_C 115.5), and three different spin systems, including a β-glucopyranosyl unit. The HMBC interpretation led us to assemble the remaining two spin systems in a germacrene-type sesquiterpene skeleton bearing an exomethylene moiety in C-10 (δ_C 152.1) (correlations of H₂-14 with C-1, C-9 and C-10) and a trisubstituted double bond (correlations of the oxymethylene H₂-15 with C-5 and the *sp*² C-3 and C-4). The remaining uncoupled methylene groups were placed at C-13 and C-12 through their HMBC correlations with both C-7 and the unprotonated C-11 (δ_C 82.1), as shown in **Figure 5.5**. Finally, the position of the sugar unit was determined on the basis of the H₂-15/C-1' HMBC cross-peak and the two acetyl groups were attached at positions 13 and 6' on the basis of the HMBC correlations of the corresponding protons with the acetyl ester carbonyls. The molecular formula of **18** implied the existence of a further degree of unsaturation that was rationalized with the presence of an oxygen bridge connecting C-9 and C-11, inferred by the diagnostic HMBC cross-peak H-9/C-11. To our knowledge, this unusual germacrene scaffold that includes the presence of an oxygenated five-membered ring is unprecedented. The diagnostic NOESY correlation between H-7/H-9 and H-9/H₂-13 revealed the spatial arrangements of the substituents on the germacrene scaffold. The geometry of the Δ^{3,4} was deduced as Z on the basis of the NOESY cross-peak H-3/H₂-5. The absolute configuration of **18** is under determination by comparison of computationally calculated and experimental ECD spectra.

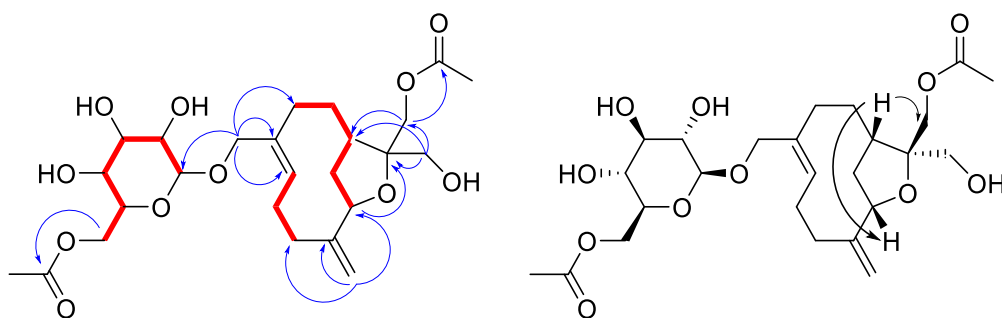


Figure 5.5 Left: COSY (red bold) and HMBC correlations (blue arrows) of compound **18**. Right: diagnostic NOESY correlations (black arrows) of compound **18** (relative stereochemistry).

The spectral data, HRESIMS, 1D and 2D NMR, of patulosin B (**19**) were very similar to those of **18**, except for some slight changes in the chemical shifts around the stereogenic center C-11. These data suggested that compound **19** could be the C-11 diastereoisomer of **18**. This hypothesis was supported by the NOESY cross-peak of H-9 with the hydroxymethylene (H_2 -12) in place of the acetoxymethylene (H_2 -13).

The presence of the two diastereomers patulosins A and B could be either due to a ring closure on an achiral carbocation deriving from an acetylated precursor (pathway A, **Figure 5.6**), or to a late-stage alternative acetylation on a prochiral diol (pathway B, **Figure 5.6**). Unfortunately, we did not isolate any of the postulated precursors, thus leaving both hypotheses open.

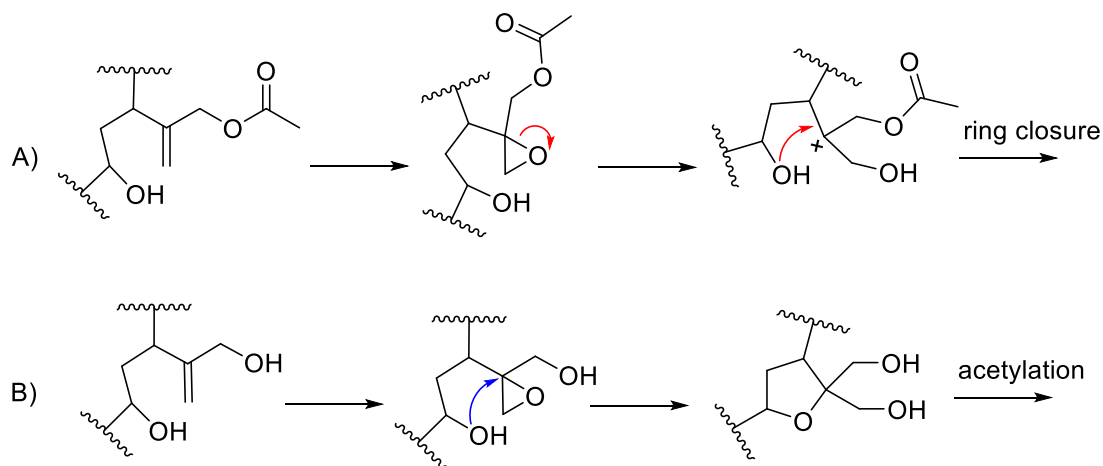


Figure 5.6 Possible formation of the two diastereomers **18** and **19**.

4.2 Alpha-glucosidase inhibitory activity

The inhibitory effects of the sesquiterpene lactones **1-19** isolated from *Centaurea patula* against α -glucosidase were evaluated in comparison with the antidiabetic acarbose used as reference compound. The α -glucosidase enzyme is involved in the carbohydrate metabolism and the release of glucose. Therefore, its inhibition is an effective strategy of delaying glucose absorption and lowering the postprandial blood glucose level, which can potentially suppress the progression of diabetes mellitus (Kilinc et al., 2020; Masullo et al., 2022). Tested compounds exhibited IC_{50} values in the range 27.8 to 232.4 μ M. Among the guaianolides, chlorojanerin (**7**) showed the lowest IC_{50} , showing that the substitution at C-15 and the acyl group at C-8 are able to significantly modulate the activity, especially if we compare the activity of **7** to that of its analogues **6** and **8**. However, the highest activity was detected among glycosylated sesquiterpenes, with the newly described elemanolide **16** being the most active, with an IC_{50} of 27.8 μ M, comparable to that observed for acarbose (IC_{50} = 20.8 μ M), the antidiabetic compound used as a reference. Interestingly, in this case the chlorine substitution present in compound **17** causes a decrease in activity.

Table 5. 1 α -Glucosidase inhibitory activity of **1-19** isolated by *Centaurea patula*.

Compounds	$IC_{50} \pm SD$ (μ M/mL)
1	145.43 $\pm$ 4.43
2	100.80 $\pm$ 4.31
3	79.88 $\pm$ 3.27
4	110.18 $\pm$ 2.94
5	86.74 $\pm$ 3.71
6	162.70 $\pm$ 2.93
7	42.45 $\pm$ 2.53
8	157.59 $\pm$ 5.43
9	107.98 $\pm$ 3.45
10	57.05 $\pm$ 2.54
11	59.56 $\pm$ 4.10
12	232.42 $\pm$ 6.15
13	75.00 $\pm$ 1.85
14	130.34 $\pm$ 3.05
15	36.70 $\pm$ 1.32
16	27.80 $\pm$ 0.55
17	37.66 $\pm$ 1.57
18	49.20 $\pm$ 2.06
19	59.23 $\pm$ 2.33
Acarbose	20.81 $\pm$ 0.66

Data are expressed as the mean values $\pm$ SD of three experiments.

PART 2

To further investigate the α -glucosidase inhibition of the most active compounds, an inhibition kinetic study was performed for **7** and **15-17** using Lineweaver-Burk plots. The results are shown in **Figure 5.7** and **Table 5.2**. In this study, the maximum velocity ($V_{\max}$) value was determined as 40.82 $\mu\text{M}/\text{min}$ and Michaelis-Menten constant (K_m) as 0.15 mM for 4- *p*-nitrophenyl- β -glucopyranoside (pNPG). The kinetic results showed that the inhibition mode of compounds **7**, **15-17** is of the mixed-typed because the regression lines did not intersect both the x- and y-axes in the Lineweaver-Burk plot (Şöhretoğlu et al., 2018).

Table 5.2 Kinetic parameters of α -glucosidase in the presence of **7** and **15-17**

Inhibitors	Inhibition type	$V_{\max}$ ($\mu\text{M}/\text{min}$); K_m (mM)
7	Mix	16.13, 0.25
15	Mix	23.09, 0.19
16	Mix	20.79, 0.51
17	Mix	26.88, 0.52

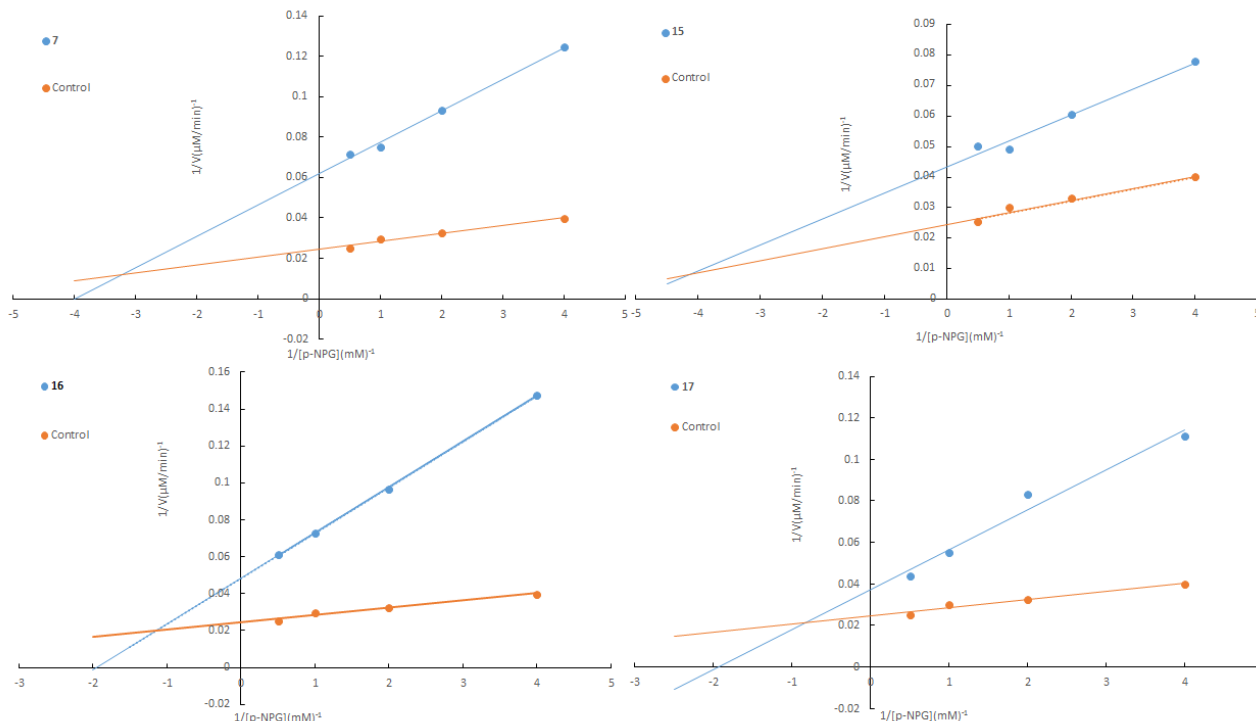


Figure 5.7 Lineweaver-Burk plots for the inhibition of compounds **7**, **15-17** at different concentrations of *p*-NPG (0.25, 0.50, 1.00 and 2.00 mM) on α -glucosidase enzyme.

5.3 Conclusions

An NMR-guided investigation of *C. patula* aerial parts led to the isolation of 19 sesquiterpenoids belonging to three different structural families, five of which were new compounds, including patulosins A-B (**18-19**) that show an unprecedented C-9/C-11 oxygen bridge. The sesquiterpene profile of this plant is quite peculiar among the previously described *Centaurea* species, not only because of the combination of different scaffolds and their structural rarity within the genus, but also for the unusual abundant presence of acetylated metabolites. The plant is traditionally used for its antidiabetic potential and preliminary investigations on the extract had revealed a significant inhibitory activity on α -glucosidase. Among the compounds tested, all glycosylated sesquiterpenes showed good inhibitory activity, suggesting that the presence of the sugar moiety may influence the ability to interact with the enzyme. In particular, the guaianolide chlorojanerin (**7**) and the newly described glycosylated elemanolide **16** were found to be the most active compound. Furthermore, the activity seems to be influenced by minor structural details such as the presence of chlorine in place of an -OH group or the orientation of substituents, thus pointing to a specific interaction with the target enzyme that we are currently looking forward to evaluating.

5.4 Experimental section

5.4.1 Plant Material

C. patula DC. aerial parts were collected in May and June 2009 from Konya and Golbasi (Ankara), Turkey. The plants have been identified by Dr. Tuna UYSAL from Section of Botany, Department of Biology, Faculty of Science, Selcuk University. The voucher specimens have been deposited in KNYA herbarium at Department of Biology, Selcuk University.

5.4.2 Extraction and Isolation

C. patula L. aerial parts (500 g) were dried in an oven at 40°C and extracted at room temperature using *n*-hexane (1.3 L for 3 days, three times) and CH₂Cl₂ (1.3 L for 3 days, three times), successively. After filtration and evaporation of the solvent to dryness in vacuo, 2.4 g of crude CH₂Cl₂ extract was obtained. This phase was subjected to MPLC-DAD on a Cartridge C18 60Å

50 μ m - size 40 (55 g) – column volume (CV) 50 mL. The mobile phase was a mixture of A: H₂O with a 0.1% formic acid, B: MeCN with the following gradient method: 90%A – 10%B for CV 0-2; 80%A – 20%B for CV 4-7; 70%A – 30%B for CV 9-12; 60%A – 40%B for CV 14-17; 50%A – 50%B for CV 19-22; 40%A – 60%B for CV 24-27; 30%A – 70%B for CV 29-32; 20%A – 80%B for CV 34-37; 10%A – 90%B for CV 39-42, setting the flow rate at 20.0 mL/min. The UV detection wavelength was set at 227 nm. This separation afforded 11 fractions (labeled **A-L**). All the fractions were purified on RP-HPLC or MPLC, resulting in the isolation of compounds **1–19**. Fraction E (49.1 mg) was subjected to an MPLC purification with the following gradient method: 90%A – 10%B for CV 0-5; 70%A – 30%B for CV 45-50, on a Silica Cartridge C18 60 $\text{\AA}$ 50 μ m - size 4 (5.9 g)- column volume (CV) 5 mL, UV detection wavelength at 227 nm, to afford four subfractions (E0–E4). Subfraction E0 (9.9 mg) was further separated by RP-HPLC on a Luna C18 5 μ m 100 $\text{\AA}$ 250 x 4.6mm using the isocratic elution: H₂O-MeOH 55:45, flowrate 1.0 mL/min, to give the unreported compound **9** (1.1 mg, t_R = 5.8 min) and cebellin J (**8**) (1.6 mg, t_R = 7.4 min). Subfraction E1 (4.5 mg) was further separated by RP-HPLC on a Kinetex C18 5 μ m 100 $\text{\AA}$ 250 x 4.6mm using a mixture of H₂O-ACN 85:15, flowrate 1.0 mL/min, yielding to compound **2** (0.6 mg, t_R = 11.9 min) and **1** (0.5 mg, t_R = 23.2 min). Fraction F (159.3 mg) was purified on an MPLC equipped with a Silica Cartridge C18 100 $\text{\AA}$ 25 μ m - size 12 (20g) - column volume (CV) 19 mL, using the following gradient method: 90%A – 10%B for CV 0-3; 80%A – 20%B for CV 6-9, 70%A – 30%B for CV 11-14, 60%A – 40%B for CV 17-21. The UV detection wavelength was set at 227 nm and the flowrate was set at 15 mL/min. From this purification compound **20** (4.9 mg, CV 9-10) was obtained in pure form, while subfractions F0-F3 were purified on HPLC. F0 (12.2 mg) was subjected to HPLC, using H₂O:MeOH 60:40 as mobile phase, on a Luna C18 5 μ m 100 $\text{\AA}$ 250 x 4.6mm to afford compounds **6** (2.6 mg, t_R = 9.7 min), **14** (0.6 mg t_R = 10.7 min), the **10** (1.3 mg t_R = 12.2 min) and compound **19** (1.4 mg t_R = 15.3 min). Subfraction F1 (19.8 mg) was purified as subfraction F0, yielding to compound **13** (0.9 mg t_R = 9.6 min), **5** (0.7 mg t_R = 14.8 min), **7** (1.9 mg t_R = 15.7 min) and the unreported **18** (1.0 mg t_R = 35.1 min). Subfraction F2 (21.2 mg) was subjected to HPLC (H₂O-MeOH 50:50) as mobile phase on a Luna C18 5 μ m 100 $\text{\AA}$ 250 x 4.6mm to afford compounds **11** (1.6 mg t_R = 9.6 min), **3** (1.7 mg t_R = 13.0 min) and the unprecedented **16** (5.4 mg t_R = 25.6 min). Fraction G (50.6 mg) was subjected to an MPLC equipped with a Silica Cartridge C18 60 $\text{\AA}$ 50 μ m - size 4 (5.9 g) - column volume (CV) 5 mL, with the following gradient method: 80%A – 20%B for CV 0-5; 60%A – 40%B for CV 40-45, the flowrate was set at 10 mL/min and the UV detection wavelength at 227 nm, to afford four subfractions (G0–G5). From this

purification of **12** (9.0 mg, CV 24-25) was obtained in pure form, while subfractions G0 was purified on HPLC as subfraction F2 to afford compounds **4** (3.8 mg t_R = 23.6 min). Fraction H (136.3 mg) was purified on a MPLC using a Silica Cartridge C18 100Å 25µm - size 4 (5.4 g) – column volume (CV) 5 mL, at 10 mL/min as flowrate, setting UV wavelength at 270 nm, with the following gradient method: 60%A – 40%B for CV 0-5; 55%A – 60%B for CV 30-70, leading to the isolation of the previously undescribed **17** (17.6 mg CV 56-57). Fraction I (85.3 mg) was subjected to an MPLC purification using a Silica Cartridge C18 60Å 50µm - size 4 (5.9 g) - column volume (CV) 5 mL, at 10 mL/min flowrate, applying the following gradient method: 50%A – 50%B for CV 0-5; 40%A – 60%B for CV 20-25, 30%A – 70%B for CV 40-45, setting the UV wavelength at 280 nm, to give the unprecedented **15** (17.6 mg CV 38).

5.4.3 Spectral Data

Compound 9: Colorless amorphous solid, $[\alpha]_D^{25} + 21.4$ (c 0.08, CHCl₃); HRESIMS $[M+Na]^+$ m/z found 463.1642 (C₂₁H₂₈O₁₀Na⁺ requires 463.1575). ¹H NMR (CDCl₃, 600 MHz) 6.22 (1H, d, J 3.10 Hz, H-13a), 5.60 (1H, d, J 3.10 Hz, H-13b), 5.14 (2H, overlapped, H-8, H-14a), 5.05 (1H, d, J 12.8 Hz, H-15b), 5.02 (1H, s, H-14b), 4.88 (1H, t, J 10.5 Hz, H-6), 4.20 (1H, d, J 12.8 Hz, H-15a), 3.87 (1H, d, J 11.2 Hz, H-4'a), 3.78 (1H, m, H-3), 3.63 (1H, d, J 11.2 Hz, H-4'b), 3.57 (1H, m, H-1), 3.14 (1H, m, H-7), 2.66 (1H, dd, J 14.9, 4.8 Hz, H-9a), 2.50 (1H, m, H-2a), 2.44 (1H, t, J 14.9 Hz, H-9b), 2.30 (1H, t, J 10.5 Hz, H-5), 2.17 (3H, s, OAc), 1.60 (1H, overlapped, H-2b), 1.41 (3H, s, H-5'); ¹³C NMR (CDCl₃, 125 MHz) 175.5 (C-2'), 172.8 (OAc), 168.6 (C-12), 142.0 (C-10), 136.3 (C-11), 122.4 (C-13), 118.4 (C-14), 84.4 (C-4), 76.4 (C-6), 76.3 (C-3), 75.9 (C-8), 75.8 (C-3'), 68.4 (C-4'), 66.2 (C-15), 57.1 (C-5), 46.3 (C-7), 46.1 (C-1), 37.1 (C-2), 34.8 (C-9), 21.8 (C-5'), 20.8 (OAc).

Compound 15: Colorless amorphous solid, $[\alpha]_D^{25} - 6.27$ (c 0.1, CHCl₃); HRESIMS $[M+Na]^+$ m/z found 505.2473 (C₂₅H₃₈O₉Na⁺ requires 505.2408). ¹H NMR (CDCl₃, 600 MHz) 5.79 (1H, dd, J 17.8, 10.9 Hz, H-1), 5.11 (1H, d, J 17.8 Hz, H-2a), 5.03 (2H, s, H-12), 4.97 (1H, d, J 10.9 Hz, H-2b), 4.85 (1H, overlapped, H-3a), 4.66 (1H, s, H-3b), 4.60 (2H, s, H-13), 4.36 (1H, dd, J 11.8, 2.1 Hz, H-6'a), 4.24 (1H, overlapped, H-6'b), 4.22 (1H, d, J 8.1 Hz, H-1'), 4.00 (1H, d, J 9.4 Hz, H-14b), 3.75 (1H, d, J 9.4 Hz, H-14a), 3.45 (1H, m, H-5'), 3.33 (1H, overlapped, H-3'), 3.27 (1H, overlapped, H-4'), 3.17 (1H, J 8.1 Hz, H-2'), 2.18 (1H, dt, J 13.3, 2.9 Hz H-9a), 2.12 (1H, dt, J 13.0, 2.9 Hz, H-5), 2.11 (1H, m, H-7), 2.07 (3H, s, OAc'), 2.04 (3H, s, OAc, s), 1.75 (1H, q, J 13.0 Hz, H-6a), 1.73

(3H, s, H-15), 1.65 (1H, m, H-8b), 1.57 (2H, overlapped H-6b, H-8a), 1.33 (1H, m, H-9b); ^{13}C NMR (CDCl_3 , 125 MHz) 171.0 (OAc, OAc'), 148.6 (C-11), 146.1 (C-1), 145.9 (C-4), 112.0 (C-3), 111.6 (C-2), 110.6 (C-12), 103.5 (C-1'), 76.5 (C-3'), 73.8 (C-5'), 73.6 (C-2'), 70.3 (C-4'), 68.8 (C-14), 65.8 (C-13), 63.2 (C-6'), 52.5 (C-5), 42.9 (C-10), 41.8 (C-7), 33.4 (C-6), 32.3 (C-9), 26.6 (C-8), 24.1 (C-15), 19.2 (OAc, OAc').

Compound 16: Colorless amorphous solid, $[\alpha]_D^{25} - 8.3$ (c 0.28, CHCl_3); HRESIMS $[\text{M}+\text{Na}]^+ m/z$ found 539.2519 ($\text{C}_{25}\text{H}_{40}\text{O}_{11}\text{Na}^+$ requires 539.2463); ^1H NMR (CDCl_3 , 700 MHz) 5.75 (1H, dd, J 17.8, 11.0 Hz, H-1), 5.09 (1H, d, J 17.8 Hz, H-2a), 4.96 (2H, d, J 11.0 Hz, H-2b), 4.85 (1H, overlapped, H-3a), 4.66 (1H, s, H-3b), 4.36 (1H, dd, J 11.9, 1.97 Hz, H-6'a), 4.25 (1H, m, H-6'b), 4.23 (1H, d, J 7.4 Hz, H-1'), 4.17 (1H, d, J 11.2 Hz, H-13a), 4.10 (1H, d, J 11.2 Hz, H-13b), 3.60 (1H, d, J 11.3 Hz, H-12a), 3.57 (1H, d, J 11.3 Hz, H-12b), 3.98 (1H, d, J 9.4 Hz, H-14b), 3.73 (1H, d, J 9.4 Hz, H-14a), 3.44 (1H, m, H-5'), 3.34 (1H, t, J 8.3 Hz, H-3'), 3.27 (1H, t, J 8.3 Hz, H-4'), 3.16 (1H, t, J 8.3 Hz, H-2'), 2.17 (1H, d, J 13.4 Hz, H-9a), 2.06 (1H, overlapped, H-5), 1.72 (1H, overlapped, H-7), 2.08 (3H, s, OAc), 2.05 (3H, s, OAc'), 1.79 (1H, m, H-6b), 1.74 (3H, s, H-15), 1.60 (1H, m, H-8a), 1.57 (1H, m, H-8b), 1.52 (1H, m, H-6a), 1.28 (1H, m, H-9b); ^{13}C NMR (CDCl_3 , 150 MHz) 171.2 (OAc), 171.1 (OAc'), 146.9 (C-1), 146.0 (C-4), 112.0 (C-3), 111.3 (C-2), 103.5 (C-1'), 76.7 (C-3'), 74.4 (C-11), 74.0 (C-5'), 73.8 (C-2'), 70.2 (C-4'), 69.3 (C-14), 65.1 (C-13), 63.5 (C-6'), 62.6 (C-12), 52.6 (C-5), 42.7 (C-10), 41.2 (C-7), 27.6 (C-6), 32.6 (C-9), 20.9 (C-8), 24.1 (C-15), 19.4 (OAc'), 19.2 (OAc').

Compound 17: Colorless amorphous solid, $[\alpha]_D^{25} - 4.1$ (c 0.10, CHCl_3); HRESIMS $[\text{M}+\text{Na}]^+ m/z$ found 557.2189 ($\text{C}_{25}\text{H}_{39}^{35}\text{ClO}_{10}\text{Na}^+$ requires 557.2124); ^1H NMR (CDCl_3 , 700 MHz) 5.75 (1H, dd, J 17.7, 10.8 Hz, H-1), 5.09 (1H, dd, J 17.7, 1.0 Hz, H-2a), 4.96 (2H, dd, J 10.8, 1.0 Hz, H-2b), 4.85 (1H, s, H-3a), 4.67 (1H, s, H-3b), 4.36 (1H, dd, J 12.0, 2.1 Hz, H-6'a), 4.24 (1H, overlapped H-6'b), 4.23 (1H, d, J 11.4 Hz, H-13a), 4.22 (1H, d, overlapped, H-1'), 4.16 (1H, d, J 11.4 Hz, H-13b), 3.98 (1H, d, J 11.2 Hz, H-14b), 3.72 (1H, d, J 11.3 Hz, H-12b), 3.71 (1H, d, J 11.2 Hz, H-14a), 3.65 (1H, d, J 11.3 Hz, H-12a), 3.45 (1H, m, H-5'), 3.34 (1H, t, J 9.2 Hz, H-3'), 3.27 (1H, t, J 9.2 Hz, H-4'), 3.17 (1H, t, J 9.2 Hz, H-2'), 2.17 (1H, dt, J 13.4, 3.8 Hz, H-9a), 2.09 (3H, s, OAc'), 2.06 (1H, overlapped, H-5), 2.04 (3H, s, OAc), 1.84 (1H, m, H-7), 1.81 (1H, t, J 12.6 Hz, H-6a), 1.74 (3H, s, H-15), 1.59 (2H, overlapped, H-8), 1.56 (1H, overlapped, H-6b), 1.29 (1H, overlapped H-9b); ^{13}C NMR (CDCl_3 , 150 MHz) 171.3 (OAc, OAc'), 146.2 (C-1), 146.6 (C-4), 112.1 (C-3), 111.4 (C-2),

PART 2

103.6 (C-1'), 76.7 (C-3'), 74.0 (C-11), 73.4 (C-5'), 73.6 (C-2'), 70.1 (C-4'), 69.3 (C-14), 65.1 (C-13), 63.4 (C-6'), 46.5 (C-12), 52.5 (C-5), 42.6 (C-10), 41.5 (C-7), 27.0 (C-6), 32.6 (C-9), 24.7 (C-8), 24.1 (C-15), 19.4 (OAc'), 19.2 (OAc).

Compound 18: Colorless amorphous solid, $[\alpha]_D^{25} +9.3$ (c 0.24, MeOH); HRESIMS $[M+Na]^+$ m/z found 537.2348 ($C_{25}H_{38}O_{11}Na^+$ requires 537.2306). ^1H NMR ($CDCl_3$, 700 MHz) 5.64 (1H, dd, J 12.2, 2.8 Hz, H-3), 5.42 (1H, s, H-14a), 5.10 (1H, s, H-14b), 4.45 (1H, d, J 12.0 Hz, H-6'a), 4.30 (1H, d, J 11.8 Hz, H-15a), 4.29 (1H, overlapped, H-9), 4.25 (1H, d, J 7.9 Hz, H-1'), 4.16 (1H, d, J 12.0 Hz, H-6'b), 4.10 (1H, s, H-13a), 4.08 (1H, s, H-13b), 4.05 (1H, d, J 11.8 Hz, H-15b), 3.70 (1H, d, J 11.3 Hz, H-12a), 3.57 (1H, d, J 11.3 Hz, H-12b), 3.39 (1H, m, H-5'), 3.32 (1H, overlapped, H-3'), 3.28 (1H, overlapped, H-4'), 3.17 (1H, t, J 8.5 Hz, H-2'), 2.65 (1H, m, H-5a), 2.43 (1H, overlapped, H-2a), 2.44 (1H, overlapped, H-8a), 2.42 (1H, overlapped, H-1a), 2.36 (1H, m, H-2b), 2.09 (3H, s, OAc'), 2.03 (3H, OAc, s), 2.00 (1H, overlapped, H-5b), 1.81 (1H, t, J 12.4 Hz, H-1b), 1.78 (1H, overlapped, H-6a), 1.65 (1H, m, H-7), 1.54 (1H, m, H-8b), 1.30 (1H, overlapped, H-6b); ^{13}C NMR ($CDCl_3$, 150 MHz): 172.4 (OAc'), 172.4 (OAc), 153.0 (C-10), 135.2 (C-3), 134.1 (C-4), 115.5 (C-14), 101.7 (C-1'), 83.3 (C-11), 79.5 (C-9), 77.6 (C-3'), 75.0 (C-5'), 74.6 (C-2'), 71.4 (C-4'), 66.8 (C-12), 65.2 (C-15), 63.6 (C-13), 64.7 (C-6'), 41.9 (C-7), 33.4 (C-5), 32.8 (C-1), 32.2 (C-2), 28.9 (C-8), 20.4 (OAc'), 20.4 (OAc), 20.2 (C-6).

Compound 19: Colorless amorphous solid, $[\alpha]_D^{25} -7.5$ (c 0.14, MeOH); HRESIMS $[M+Na]^+$ m/z found 537.2365 ($C_{25}H_{38}O_{11}Na^+$ requires 537.2306); ^1H NMR ($CDCl_3$, 700 MHz): 5.64 (1H, dd, J 11.9, 2.9 Hz, H-3), 5.34 (1H, s, H-14a), 5.11 (1H, s, H-14b), 4.43 (1H, d, J 11.8 Hz, H-6'a), 4.34 (1H, d, J 11.2 Hz, H-15a), 4.29 (1H, overlapped, H-9), 4.24 (1H, d, J 8.3 Hz, H-1'), 4.32 (1H, d, J 11.6 Hz, H-13a), 4.05 (1H, d, J 11.6 Hz, H-13b), 4.11 (1H, dd, J 11.8, 6.1 Hz, H-6'b), 3.55 (1H, d, J 11.2 Hz, H-12a), 3.51 (1H, d, J 11.2 Hz, H-12b), 3.40 (1H, m, H-5'), 3.32 (1H, overlapped, H-3'), 3.29 (1H, overlapped, H-4'), 3.19 (1H, t, J 8.3 Hz, H-2'), 2.68 (1H, m, H-5a), 2.44 (2H, overlapped, H-2a, H-8a), 2.41 (1H, overlapped, H-1a), 2.36 (1H, m, H-2b), 2.08 (6H, s, OAc), 2.00 (1H, overlapped, H-5b), 1.81 (1H, t, J 12.4 Hz, H-1b), 1.68 (1H, m, H-7), 1.37 (1H, m, H-8b), 1.28 (2H, overlapped, H-6); ^{13}C NMR ($CDCl_3$, 150 MHz): 173.3 (OAc), 172.3 (OAc'), 152.7 (C-10), 135.1 (C-3), 134.1 (C-4), 115.8 (C-14), 101.8 (C-1'), 83.5 (C-11), 79.5 (C-9), 77.7 (C-3'), 75.0 (C-5'), 74.6 (C-2'), 71.4 (C-4'), 67.6 (C-13), 65.4 (C-15), 62.6 (C-12), 64.6 (C-6'), 41.1 (C-7), 33.4 (C-5), 32.9 (C-1), 32.3 (C-2), 20.0 (C-6), 28.6 (C-8), 20.6 (OAc'), 20.6 (OAc).

5.4.4 Computational Details of compounds 16 and 17

A preliminary conformational search on each stereoisomer was performed by Simulated Annealing in the INSIGHT II package. The MeOH solution phases were mimicked through the value of the corresponding dielectric constant. Using the steepest descent followed by quasi-Newton–Raphson method (VA09A) the conformational energy was minimized. Restrained simulations were carried out for 500 ps using the CVFF force field as implemented in Discover software (Accelrys, San Diego, USA). The simulation started at 1000 K, and then the temperature was decreased stepwise to 300 K. The final step was again the energy minimization, performed in order to refine the structures obtained, using the steepest descent and the quasi-Newton–Raphson (VA09A) algorithms successively. Both dynamic and mechanic calculations were carried out by using 1 (kcal/mol)/Å² flat well distance restraints. One hundred structures were generated. To simulate the solvent chosen for NMR analysis, a distance-dependent dielectric constant set to the value of DMSO (ϵ 32.63) was used during the calculations.

DFT calculations were performed on a Intel(R) Core (TM) i5-4440 processor at 3.0 GHz using the Gaussian09 package (Multiprocessor). GIAO ¹³C calculations were performed using the mPW1PW91 functional and 6-31G(d,p) basis set using the geometry previously optimized at the mPW1PW91/6-311G(d) level as input. DP4+ probabilities were computed using the available DP4+ Toolbox (Excel file). For these calculations, the IEF-PCM solvent continuum model, as implemented in Gaussian (MeOH solvent).

5.4.6 α -Glucosidase Inhibition Assay.

α -Glucosidase inhibitory activity was evaluated as previously reported (Kilinc et al., 2020). Briefly, 0.1 M phosphate buffer at pH 7.0 (150 μ L) was added to each well; successively 10 μ L of samples, dissolved in MeOH at different concentrations (final concentrations 8, 12, 16 μ g/mL) were added, at this point, the reaction was initiated by the addition of 15 μ L of the α -glucosidase enzyme (from *Saccharomyces cerevisiae*) (2 U/ml) in each well, and the plate was incubated for 5 minutes at 37°C; after 5 minutes 75 μ L of the substrate (2.5 mM) 4-p-nitrophenyl α -D-glucopyranoside (4-pNPG) were added, and successively the plate was incubated for 10 minutes at 37°. The absorbance was measured at 405 nm before (T0') and after the incubation with the enzyme (T10'). The positive control was acarbose. Negative control absorbance

PART 2

(phosphate buffer in place of the sample) was also recorded. Inhibition of the enzyme was calculated, and the results were expressed as IC_{50} values. All experiments were carried out in triplicate. IC_{50} values were calculated through non-linear regression using Graphpad Prism 5 Software and expressed as means $\pm$ SD (standard deviation).

References

- Bruno, M., Rosselli, S., Maggio, A., Raccuglia, R.A. and Arnold, N.A., **2005**. Guaianolides from *Centaurea babylonica*. *Biochemical systematics and ecology*, 33(8), pp.817-825.
- Choi, S.Z., Choi, S.U. and Lee, K.R., **2005**. Cytotoxic sesquiterpene lactones from *Saussurea calciccola*. *Archives of pharmacal research*, 28, pp.1142-1146.
- Demir, S., Karaalp, C. and Bedir, E., **2016**. Unusual sesquiterpenes from *Centaurea athoa* DC. *Phytochemistry Letters*, 15, pp.245-250.
- Fattaheian-Dehkordi, S., Hojjatifard, R., Saeedi, M. and Khanavi, M., **2021**. A review on antidiabetic activity of *Centaurea* spp.: A new approach for developing herbal remedies. *Evidence-Based Complementary and Alternative Medicine*, 2021(1), p.5587938.
- Formisano, C., Sirignano, C., Rigano, D., Chianese, G., Zengin, G., Seo, E.J., Efferth, T. and Taglialatela-Scafati, O., **2017**. Antiproliferative activity against leukemia cells of sesquiterpene lactones from the Turkish endemic plant *Centaurea drabifolia* subsp. *detonsa*. *Fitoterapia*, 120, pp.98-102.
- González, A.G., Bermejo, J., Breton, J.L., Massanet, G.M. and Triana, J., **1974**. Chlorohyssopifolin C, D, E and vahlenin, four new sesquiterpene lactones from *Centaurea hyssopifolia*. *Phytochemistry*, 13(7), pp.1193-1197.
- Garcia-Jacas, N., Susanna, A. and İlarıslan, R., **1996**. Aneuploidy in the *Centaureinae* (Compositae): is $n=7$ the end of the series?. *Taxon*, pp.39-42.
- Garcia-Jacas, N., Uysal, T., Romashchenko, K., Suárez-Santiago, V.N., Ertuğrul, K. and Susanna, A., **2006**. *Centaurea* revisited: a molecular survey of the *Jacea* group. *Annals of Botany*, 98(4), pp.741-753.
- Huang, S.Z., Ma, Q.Y., Wang, Q., Dai, H.F., Liu, Y.Q., Zhou, J. and Zhao, Y.X., **2018**. Daphnauranins C–E, three new antifeedants from *Daphne aurantiaca* roots. *Molecules*, 23(10), p.2429.
- Kılınç, H., Masullo, M., D'Urso, G., Karayildirim, T., Alankus, O. and Piacente, S., **2020**. Phytochemical investigation of *Scabiosa sicula* guided by a preliminary HPLC-ESIMS_n profiling. *Phytochemistry*, 174, p.112350.

- Masullo, M., Lauro, G., Cerulli, A., Bifulco, G. and Piacente, S., **2022**. Corylus avellana: A Source of Diarylheptanoids With α -Glucosidase Inhibitory Activity Evaluated by in vitro and in silico Studies. *Frontiers in Plant Science*, 13, p.805660.
- Medjroubi, K., Benayache, F., Leon, F. and Bermejo, J., **2003**. Complete assignment of the ^{13}C and ^1H NMR spectra of two known guaianolides isolated from Centaurea musimomum. *Rev. Colomb. Quim*, 32, pp.17-25.
- Mohamed, T.A., Elshamy, A.I., Abd-ElGawad, A.M., Hussien, T.A., El-Toumy, S.A., Efferth, T. and Hegazy, M.E.F., **2021**. Cytotoxic and chemotaxonomic study of isolated metabolites from Centaurea aegyptiaca. *Journal of the Chinese Chemical Society*, 68(1), pp.159-168.
- Negaresh, K., **2018**. Centaurea patula (Asteraceae), A Forgotten Species for the Flora of Iran. *Taxonomy and Biosystematics*, 37, pp.13-22.
- Rosselli, S., Maggio, A., Bellone, G. and Bruno, M., **2006**. The first example of natural cyclic carbonate in terpenoids. *Tetrahedron letters*, 47(39), pp.7047-7050.
- Russo, A., Cardile, V., Graziano, A.C., Rigano, D., Aktumsek, A., Zengin, G. and Senatore, F., **2016**. Effect of three Centaurea species collected from central Anatolia region of Turkey on human melanoma cells. *Natural Product Communications*, 11(3), p.1934578X1601100302.
- Saklani, A., Sahoo, M.R., Mishra, P.D. and Vishwakarma, R., **2011**. Saussurea heteromalla (D. Don) Hand.-Mazz.: A new source of arctiin, arctigenin and chlorojanerin.
- Sen, A., Ozbas Turan, S. and Bitis, L., **2017**. Bioactivity-guided isolation of anti-proliferative compounds from endemic Centaurea kilaea. *Pharmaceutical biology*, 55(1), pp.541-546.
- Serino, E., Chianese, G., Musto, G., Zengin, G., Rigano, D., Stornaiuolo, M., Formisano, C. and Taglialatela-Scafati, O., **2022**. Guaiane-rich phytochemical profile of Centaurea kotschyi subsp. persica (Boiss.) Wagenitz and identification of hypoglycaemic metabolites. *Phytochemistry*, 199, p.113189.
- Serino, E., Rigano, D., Bruno, M., Pastore, A., Stornaiuolo, M., Formisano, C. and Taglialatela-Scafati, O., **2024**. Glucose Uptake-Stimulating Metabolites from Aerial Parts of Centaurea sicula. *Journal of Natural Products*, 87(4), pp.1179-1186.
- Singh, P. and Bhala, M., **1988**. Guaianolides from Saussurea candicans. *Phytochemistry*, 27(4), pp.1203-1205.

- Şöhretoğlu, D., Sari, S., Şoral, M., Barut, B., Özel, A. and Liptaj, T., **2018**. Potential of *Potentilla inclinata* and its polyphenolic compounds in α -glucosidase inhibition: Kinetics and interaction mechanism merged with docking simulations. *International journal of biological macromolecules*, 108, pp.81-87.
- Tekeli, Y., Zengin, G., Aktumsek, A. and Sezgin, M., **2013**. Comparison of the fatty acid compositions of Six *Centaurea* species. *Chemistry of Natural Compounds*, 49, pp.496-498.
- Zengin, G., Cakmak, Y.S., Guler, G.O. and Aktumsek, A., **2010**. In vitro antioxidant capacities and fatty acid compositions of three *Centaurea* species collected from Central Anatolia region of Turkey. *Food and Chemical Toxicology*, 48(10), pp.2638-2641.
- Zengin, G., Aktumsek, A., Boga, M., Ceylan, R. and Uysal, S., **2016**. Essential Oil Composition of an Uninvestigated *Centaurea* Species from Turkey: *Centaurea patula* DC. *Journal of Essential Oil Bearing Plants*, 19(2), pp.485-491.

APPENDIX C

Here are reported the spectra of the previously undescribed compounds isolated from *Centaurea sicula* (described in PART 2 – Chapter 5 of this manuscript).

APPENDIX C	132
Figures S1-S5: NMR spectra of compound 9 (CDCl ₃ , 700MHz)	133
Figures S6-S10: NMR spectra of compound 14 (MeOD, 600MHz)	135
Figures S11-S15: NMR spectra of compound 15 (MeOD, 700MHz)	138
Figures S16-20: NMR spectra of compound 16 (MeOD, 700MHz)	140
Figures S21-S25: ¹ H NMR spectrum of compound 17 (MeOD, 700MHz)	143
Figures S26-S30: NMR spectra of compound 18 (MeOD, 600MHz)	145
Table S1: Conformational Analysis of the stereoisomers in the MeOH solution of compound 16	148
Table S2: The calculation results of compound 16 , with mean absolute errors (MAE) values and DP4+ probabilities	148
Table S3: Comparisons of ¹³ C values for compound 16	149
Table S4: Conformational Analysis of the stereoisomers in the MeOH solution of compound 17	150
Table S5: The calculation results of compound 17 , with mean absolute errors (MAE) values and DP4+ probabilities	150
Table S6: Comparisons of ¹³ C values for compound 17	151

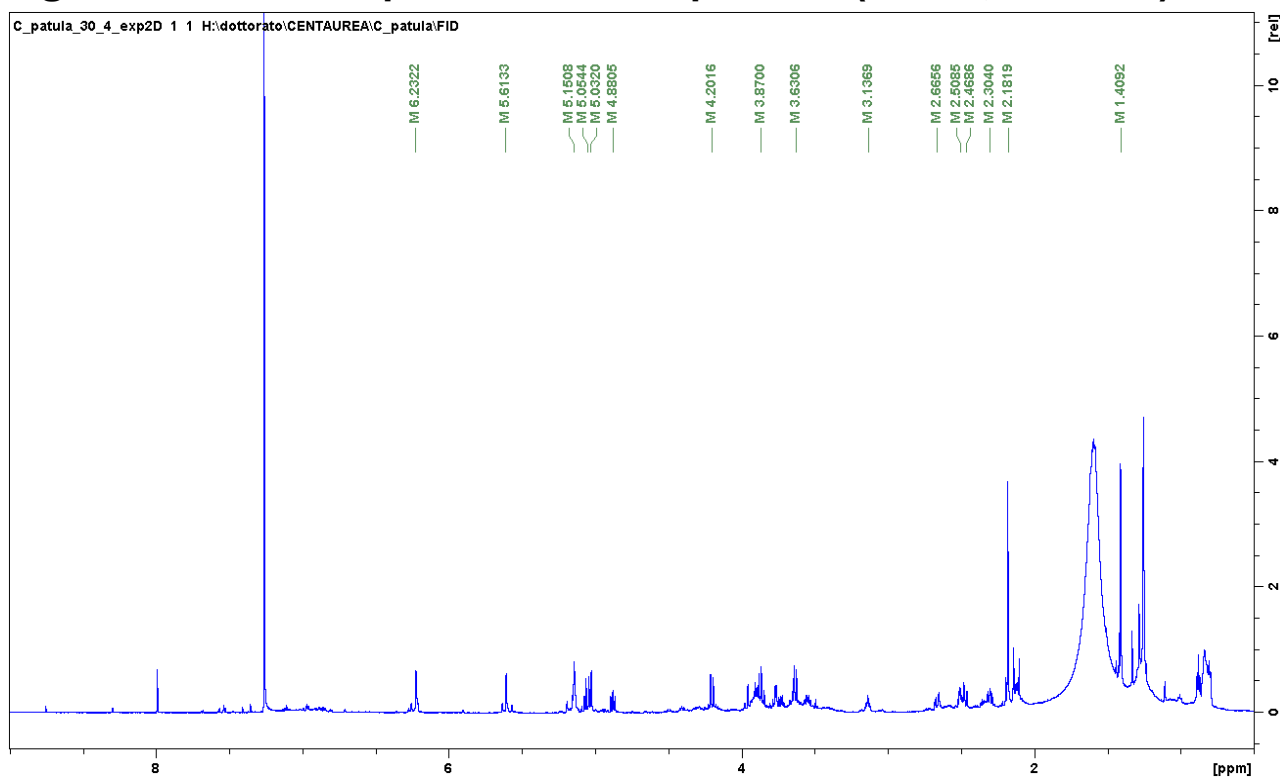
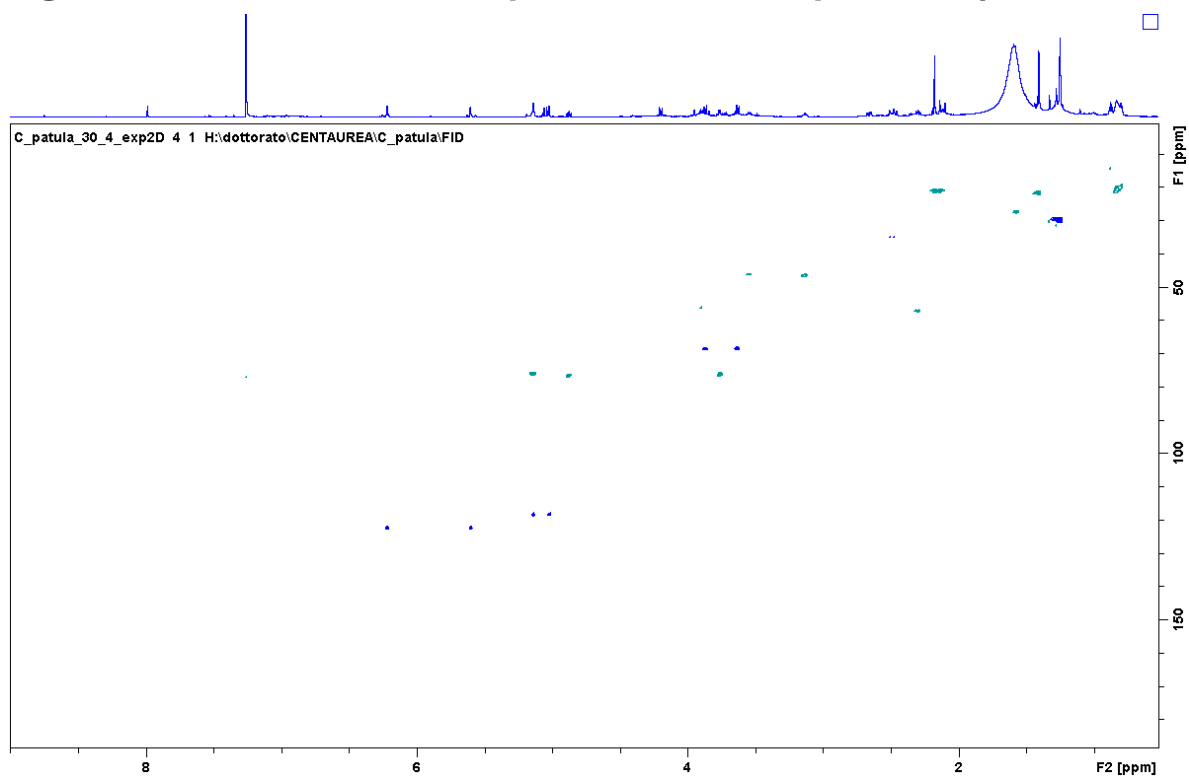
Figure S1: ^1H NMR spectrum of compound 9 (CDCl_3 , 700MHz)**Figure S2: HSQC 2D NMR spectrum of compound 9 (CDCl_3 , 700MHz)**

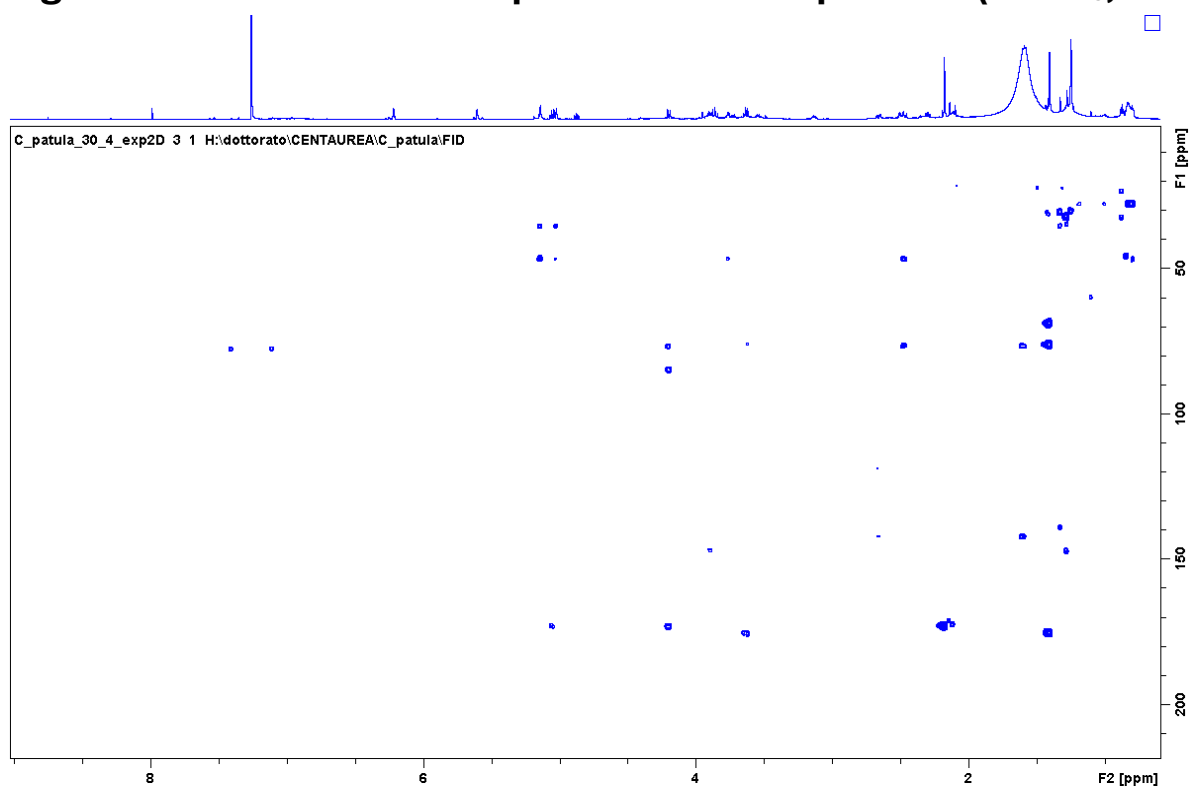
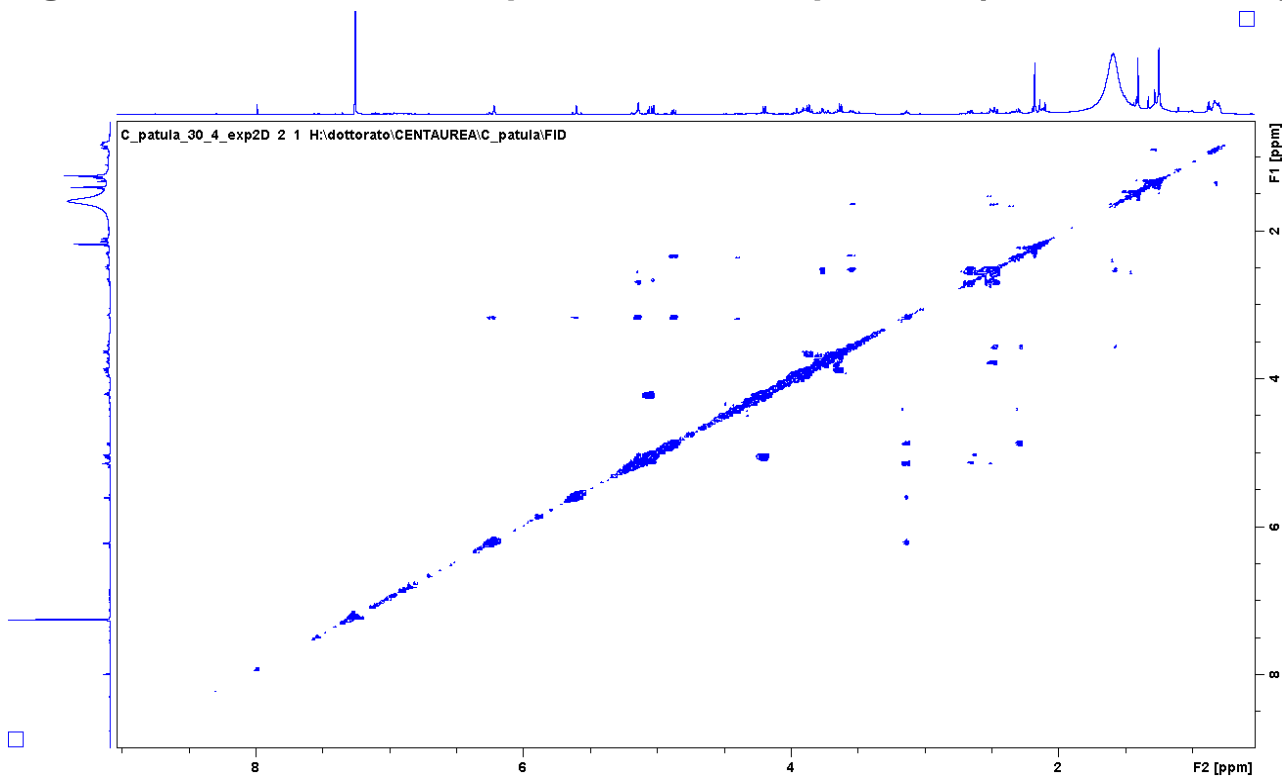
Figure S3: HMBC 2D NMR spectrum of compound 9 (CDCl₃, 700MHz)**Figure S4: COSY 2D NMR spectrum of compound 9 (CDCl₃, 700MHz)**

Figure S5: NOESY 2D NMR spectrum of compound 9 (CDCl₃, 700MHz)

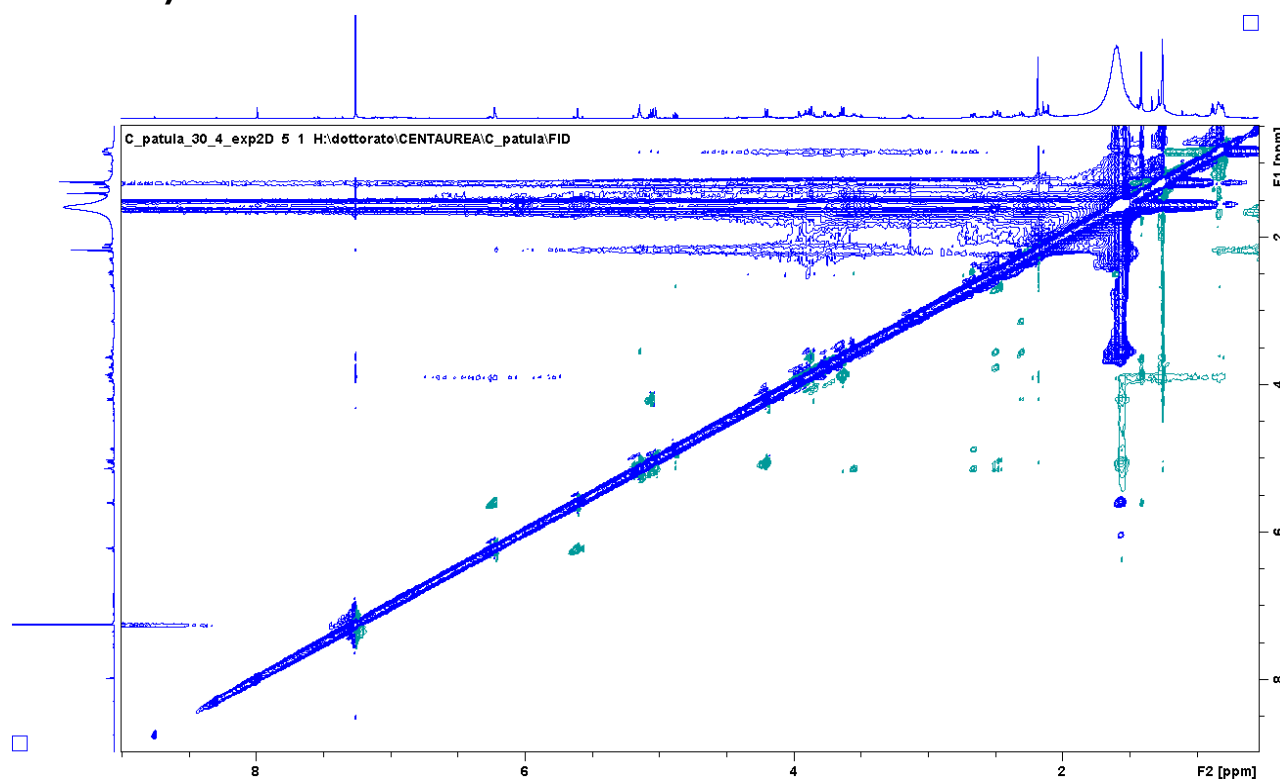


Figure S6: ¹H NMR spectrum of compound 14 (MeOD, 600MHz)

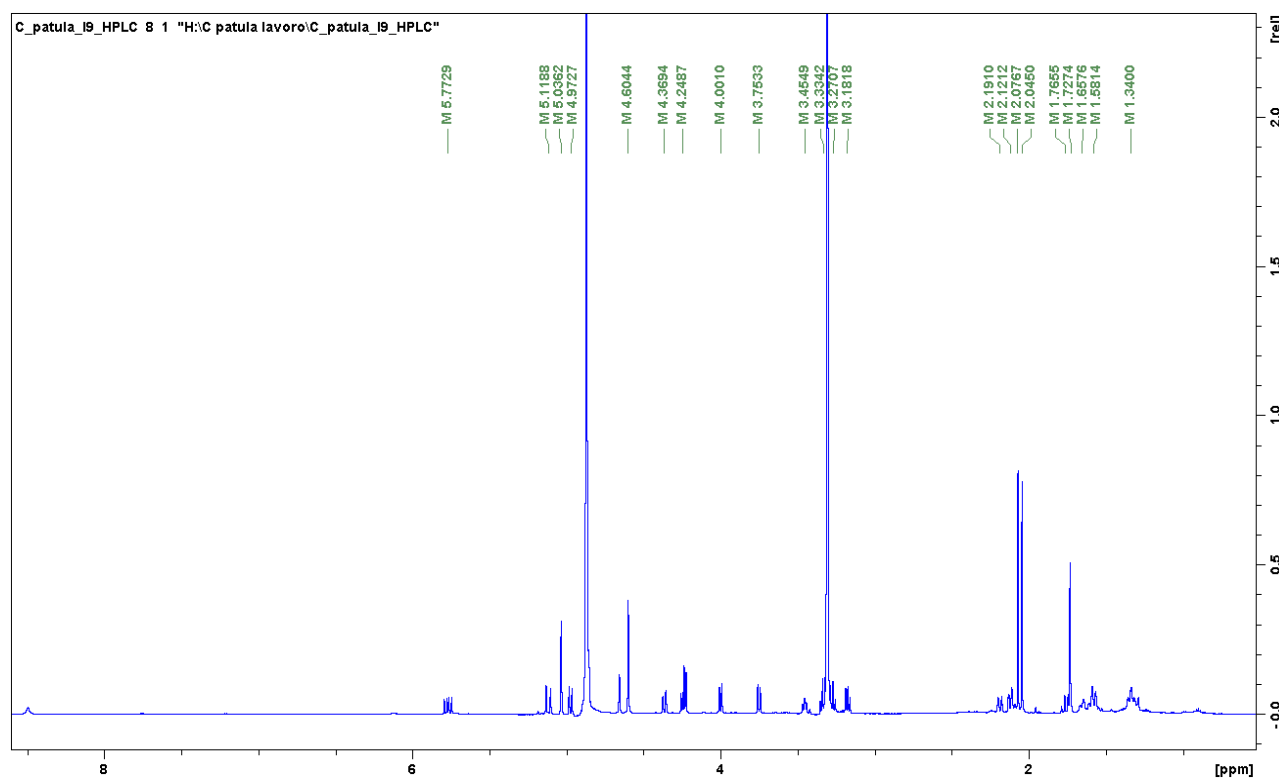


Figure S7: HSQC 2D NMR spectrum of compound 14 (MeOD, 600MHz)

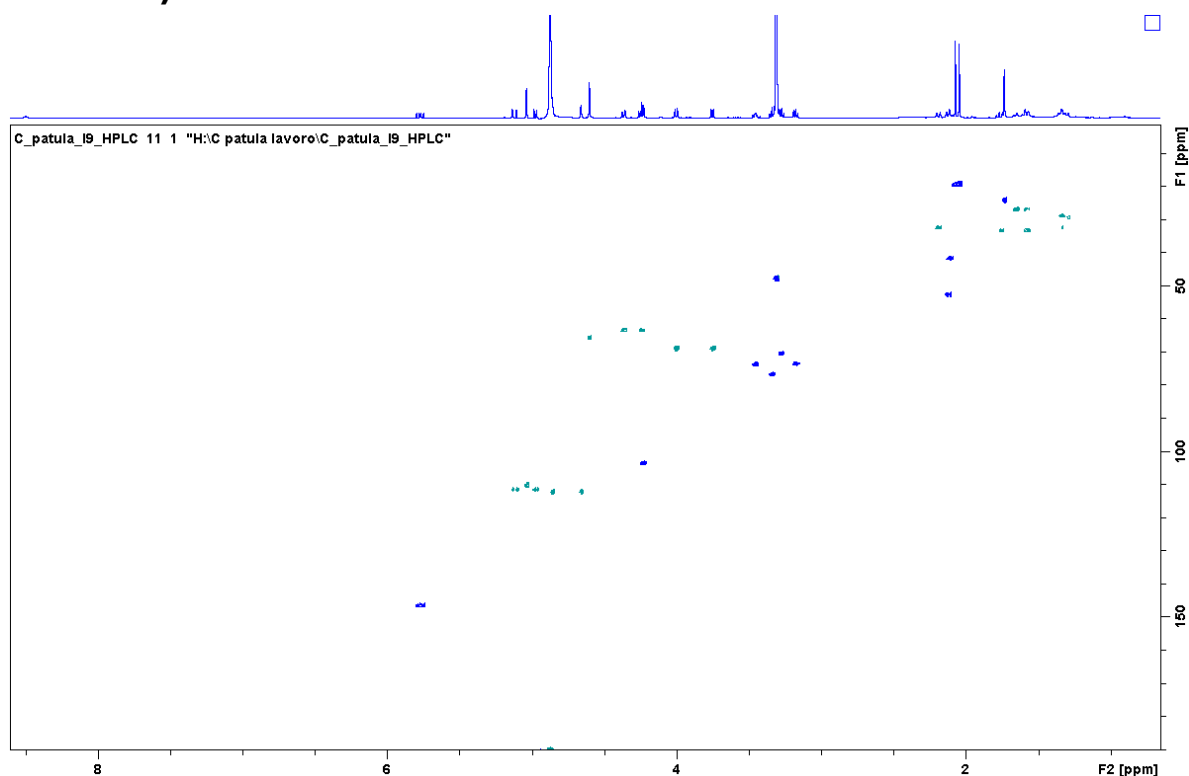


Figure S8: HMBC 2D NMR spectrum of compound 14 (MeOD, 600MHz)

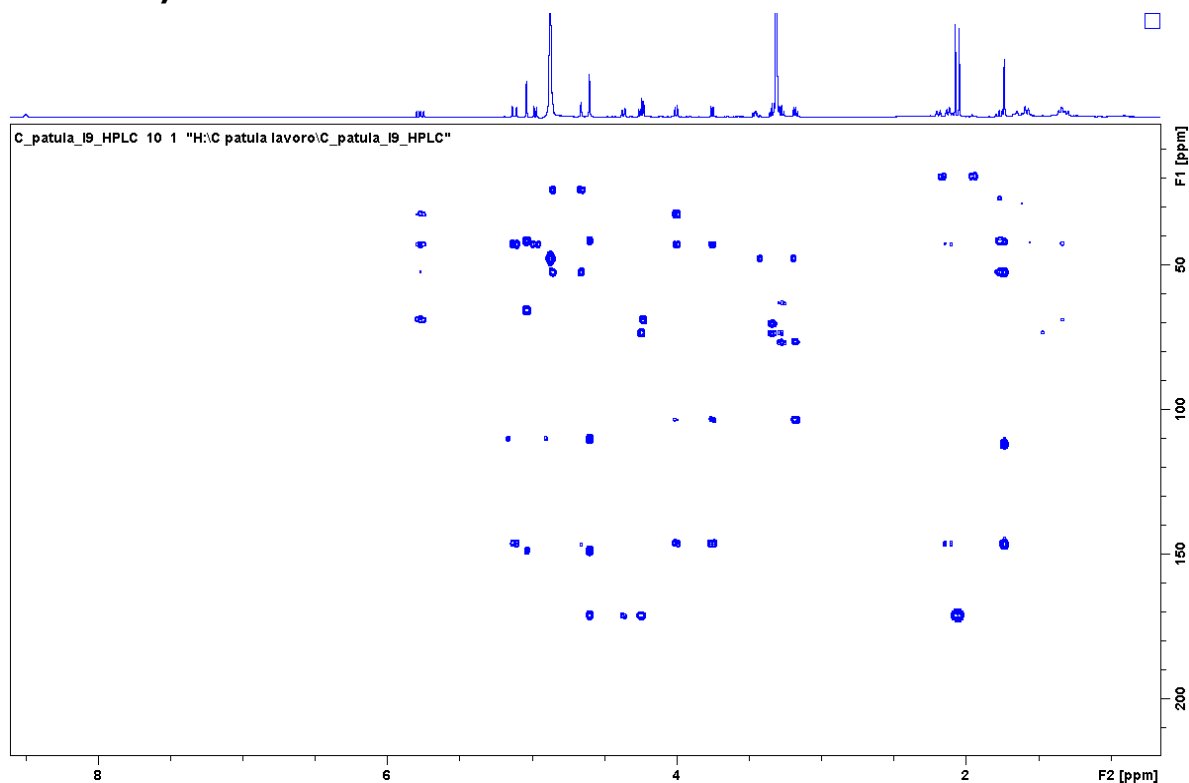


Figure S9: COSY 2D NMR spectrum of compound 14 (MeOD, 600MHz)

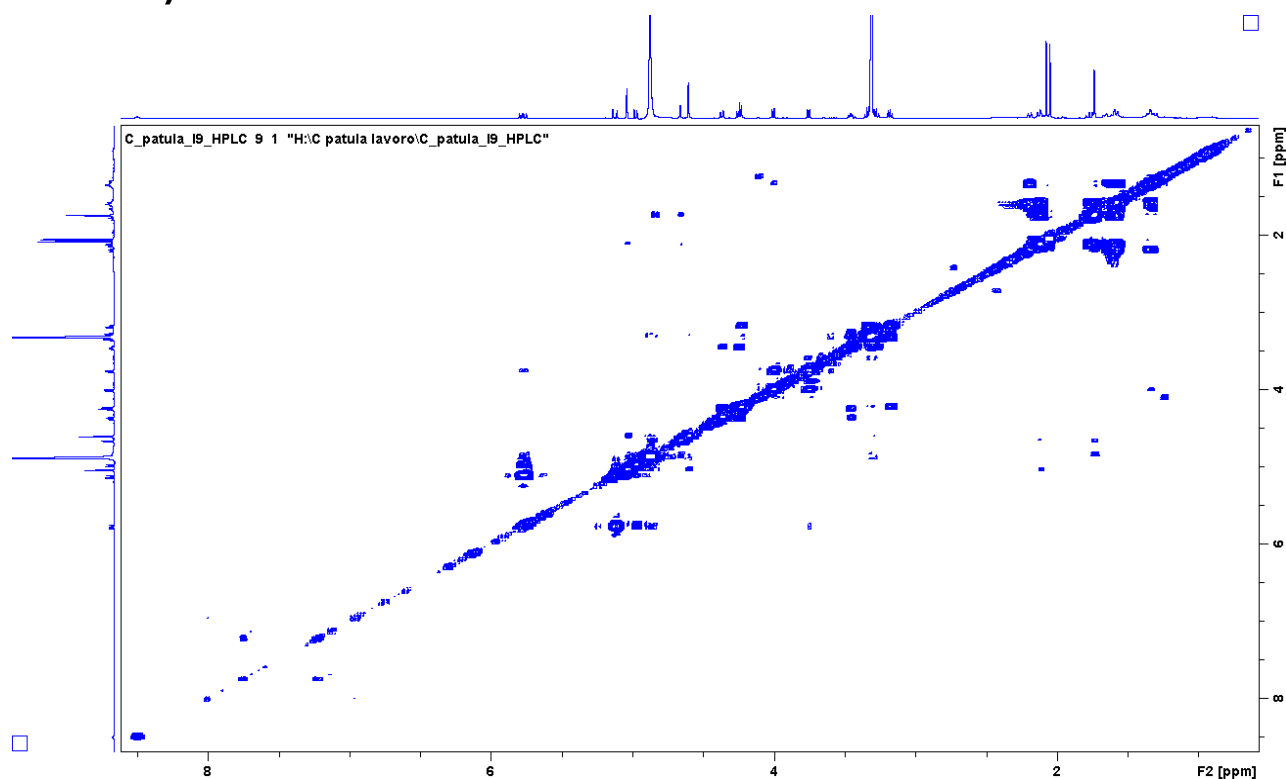


Figure S10: NOESY 2D NMR spectrum of compound 14 (MeOD, 600MHz)

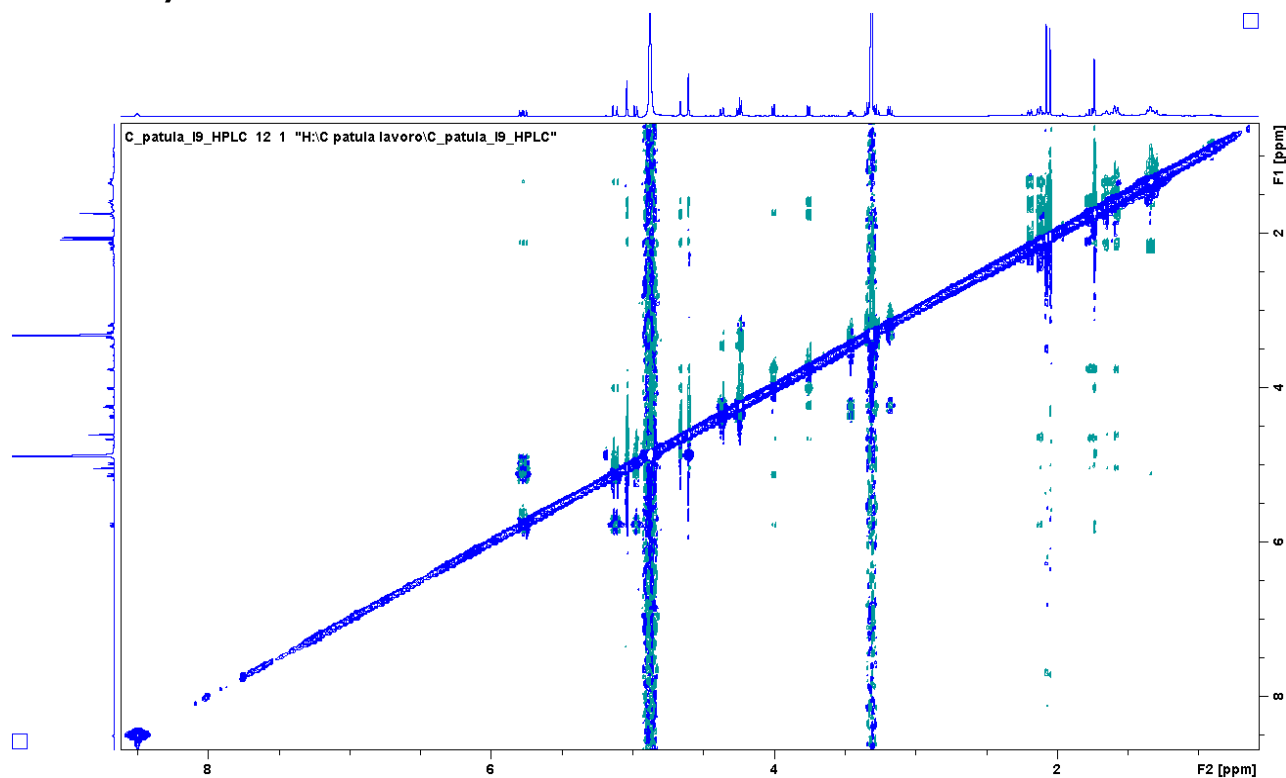


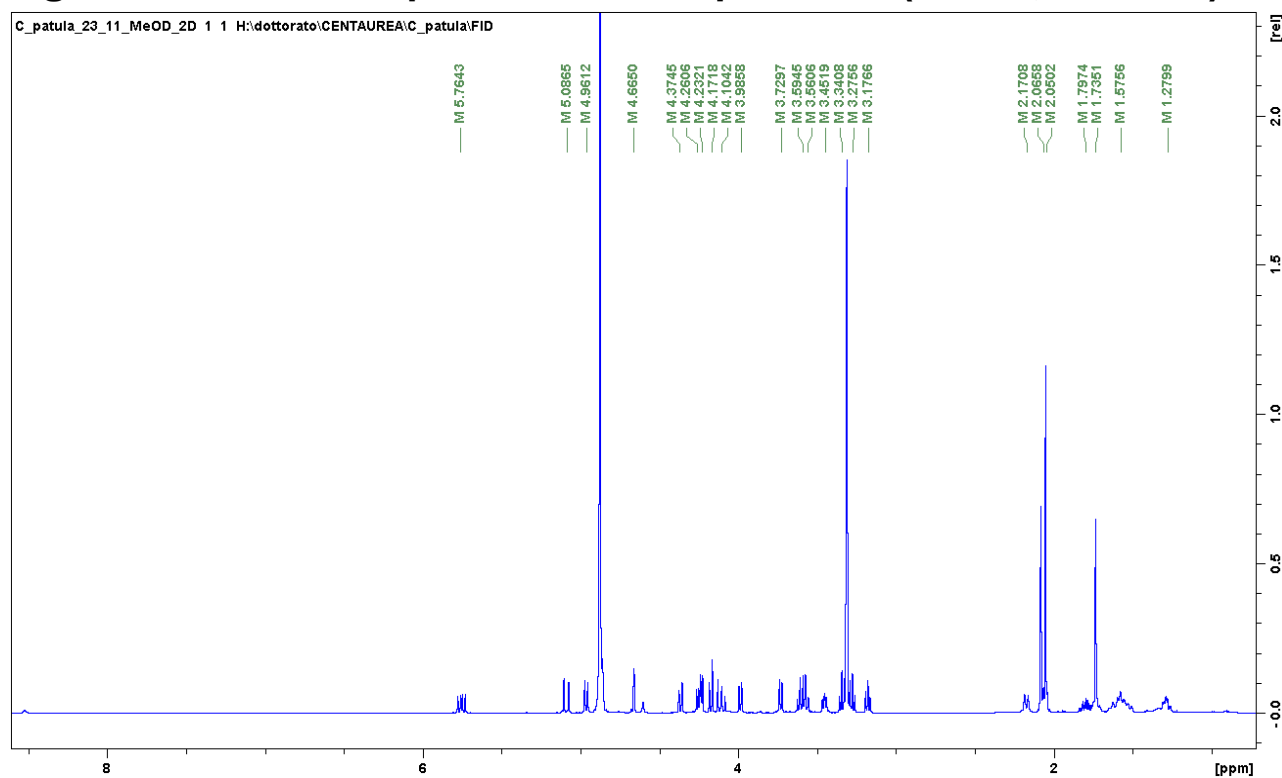
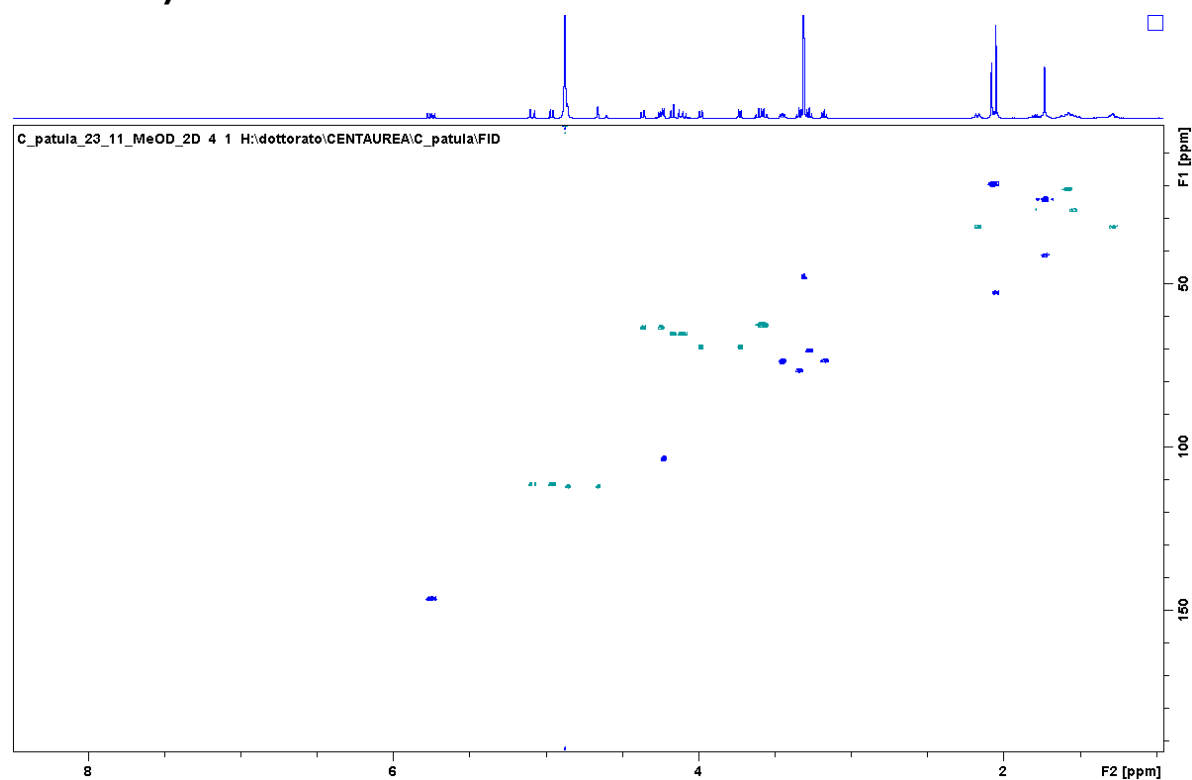
Figure S11: ^1H NMR spectrum of compound 15 (MeOD, 700MHz)**Figure S12: HSQC 2D NMR spectrum of compound 15 (MeOD, 700MHz)**

Figure S13: HMBC 2D NMR spectrum of compound 15 (MeOD, 700MHz)

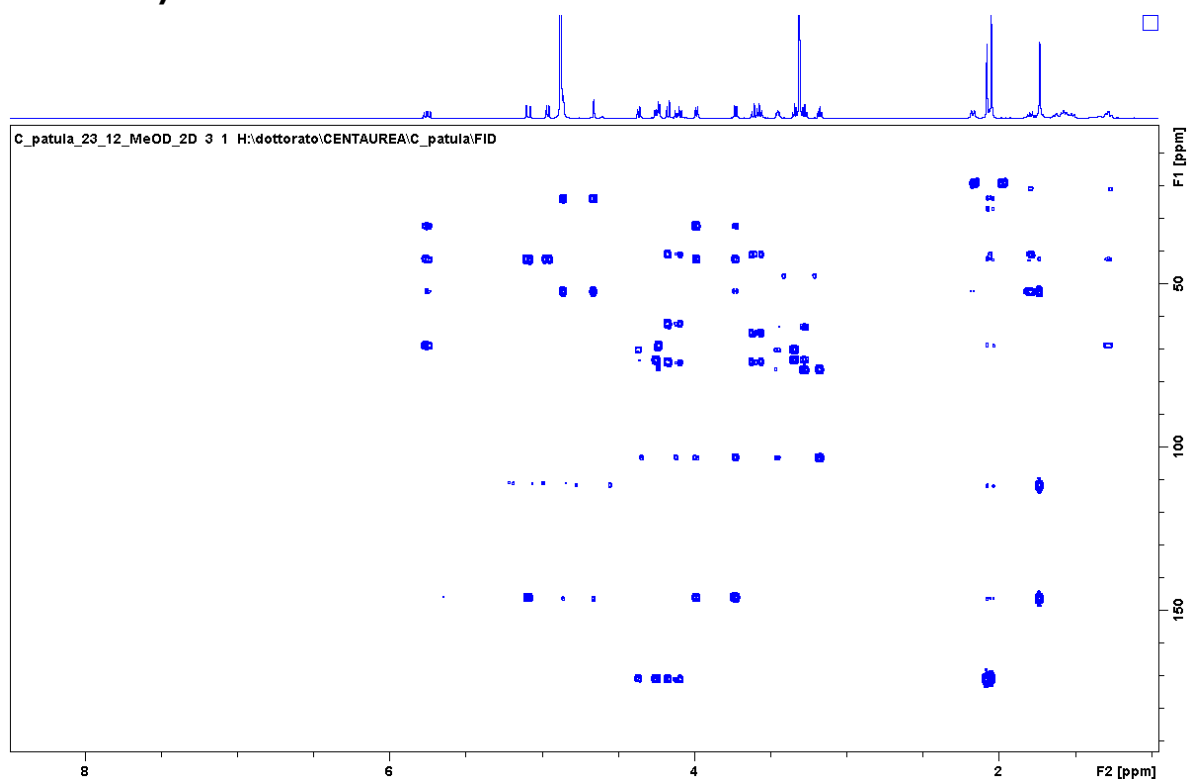


Figure S14: COSY 2D NMR spectrum of compound 15 (MeOD, 700MHz)

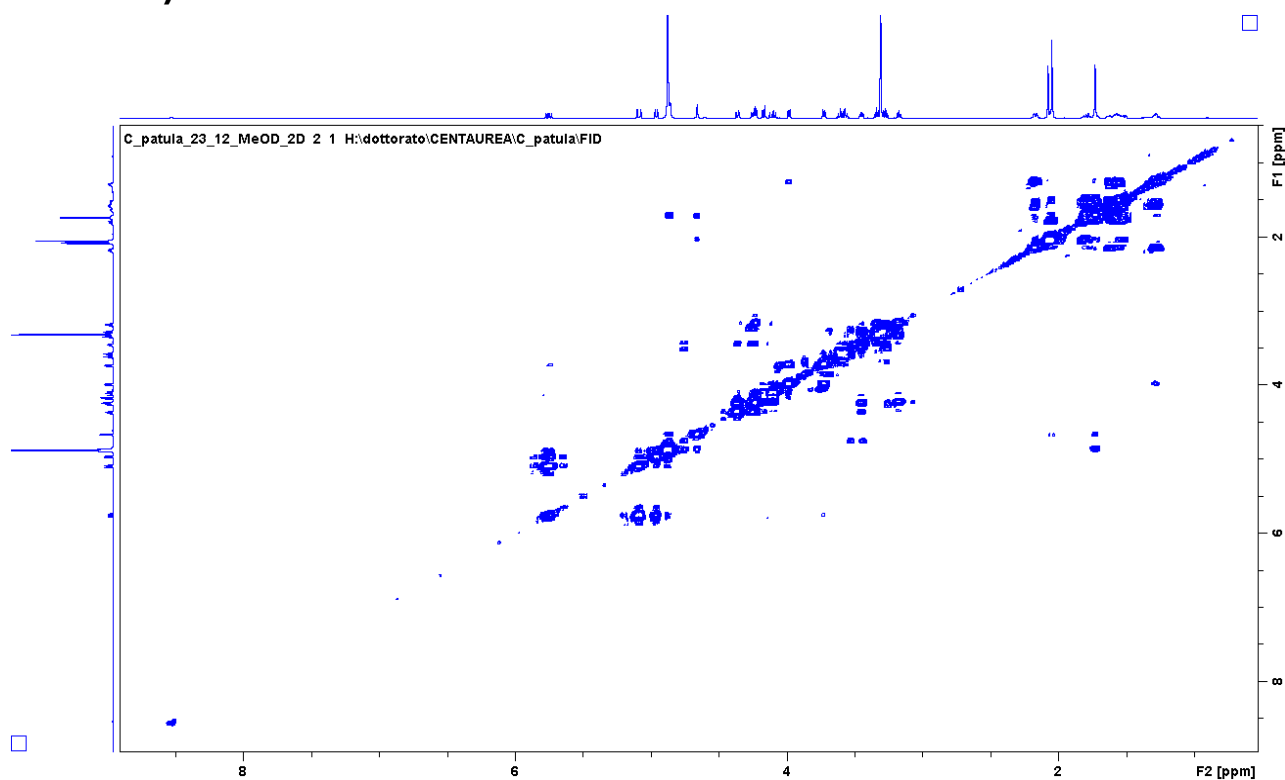


Figure S15: NOESY 2D NMR spectrum of compound 15 (MeOD, 700MHz)

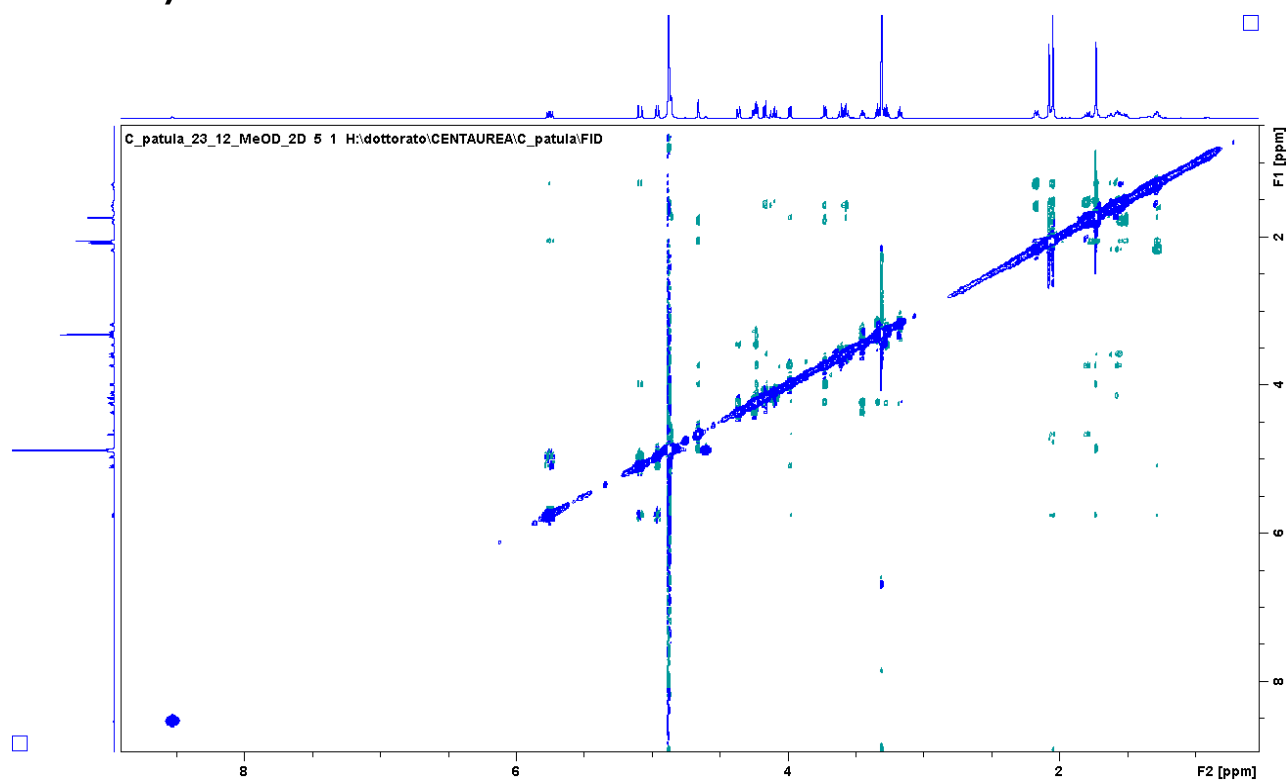


Figure S16: ^1H NMR spectrum of compound 16 (MeOD, 700MHz)

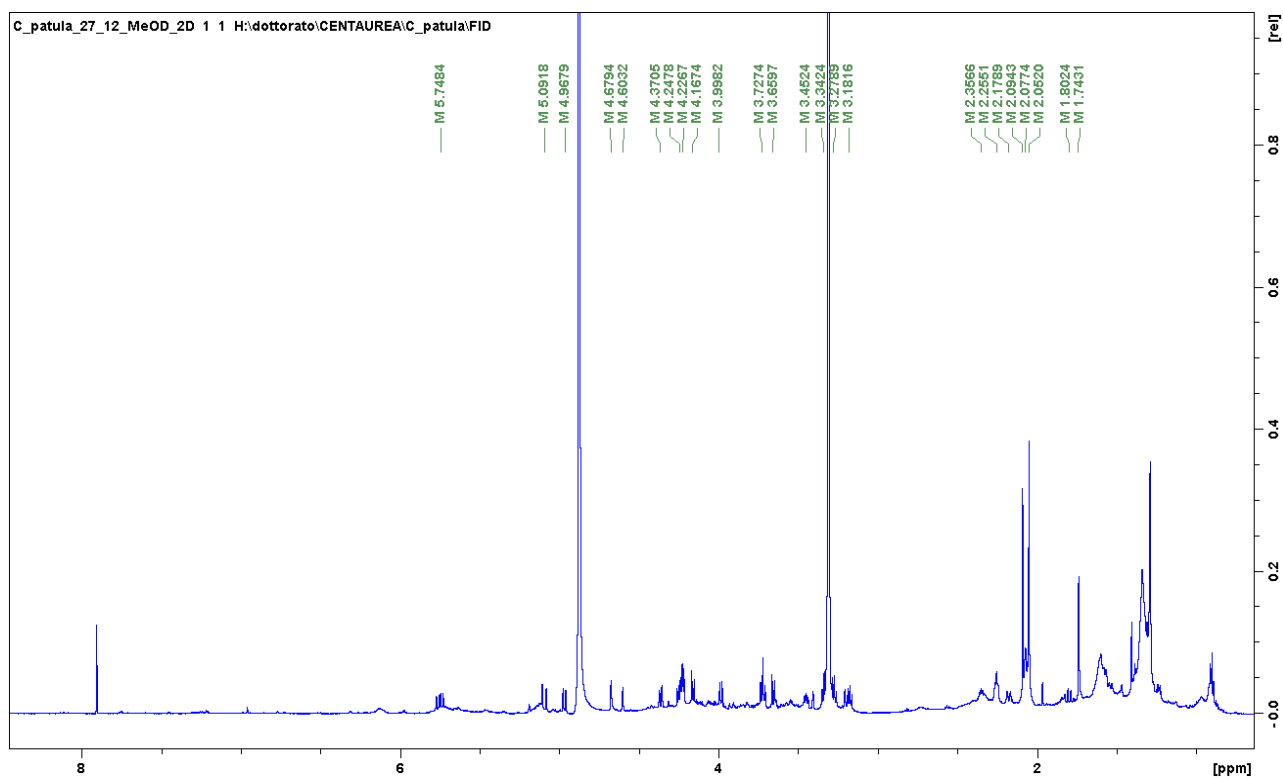


Figure S17: HSQC 2D NMR spectrum of compound 16 (MeOD, 700MHz)

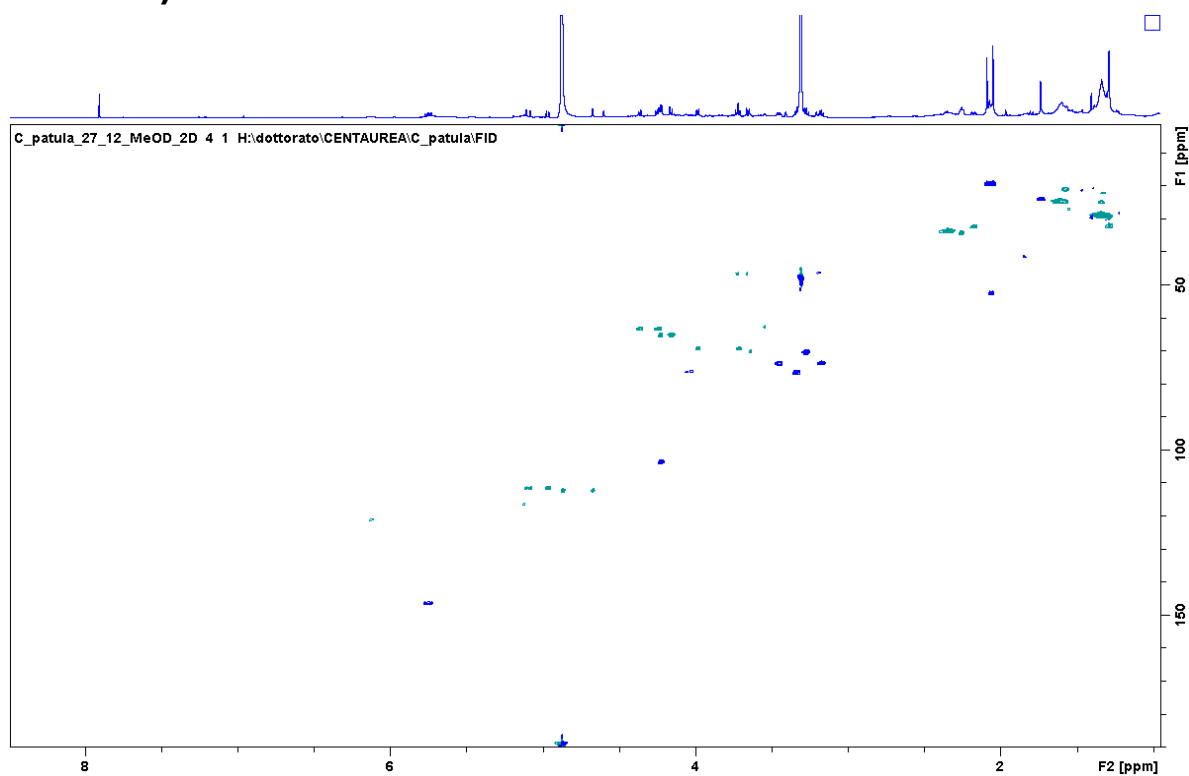


Figure S18: HMBC 2D NMR spectrum of compound 16 (MeOD, 700MHz)

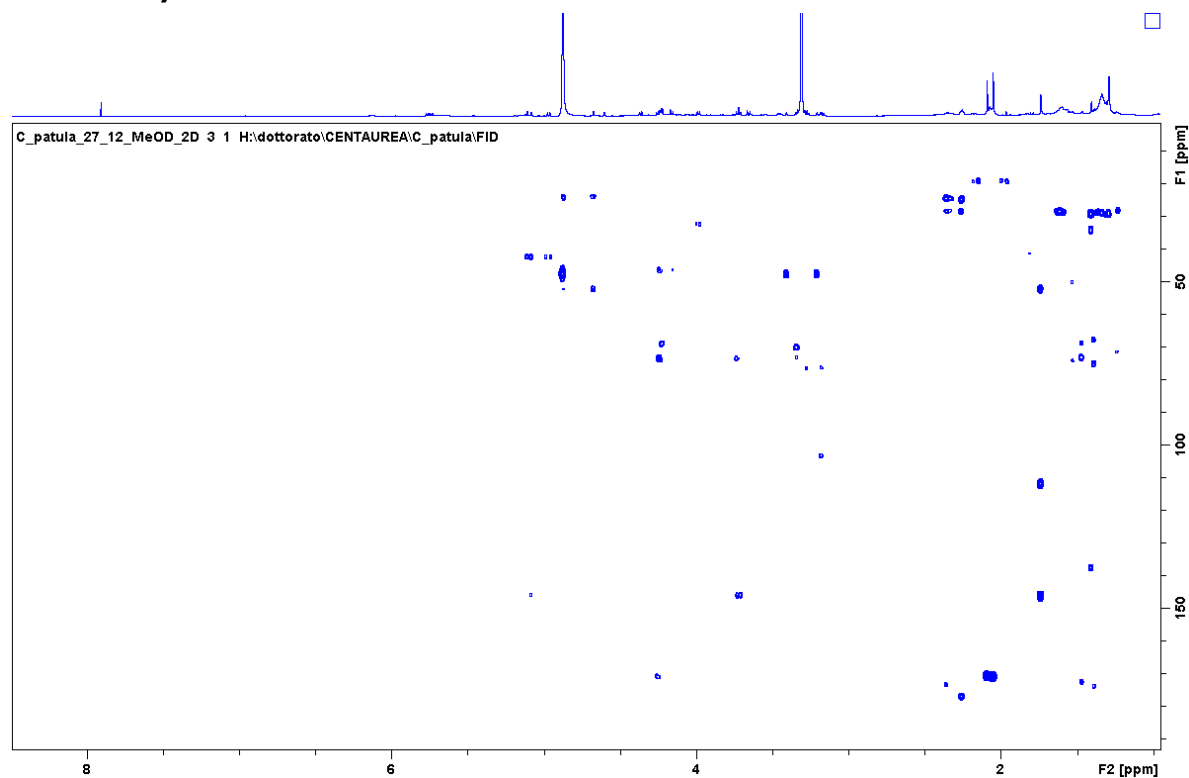


Figure S19: COSY 2D NMR spectrum of compound 16 (MeOD, 700MHz)

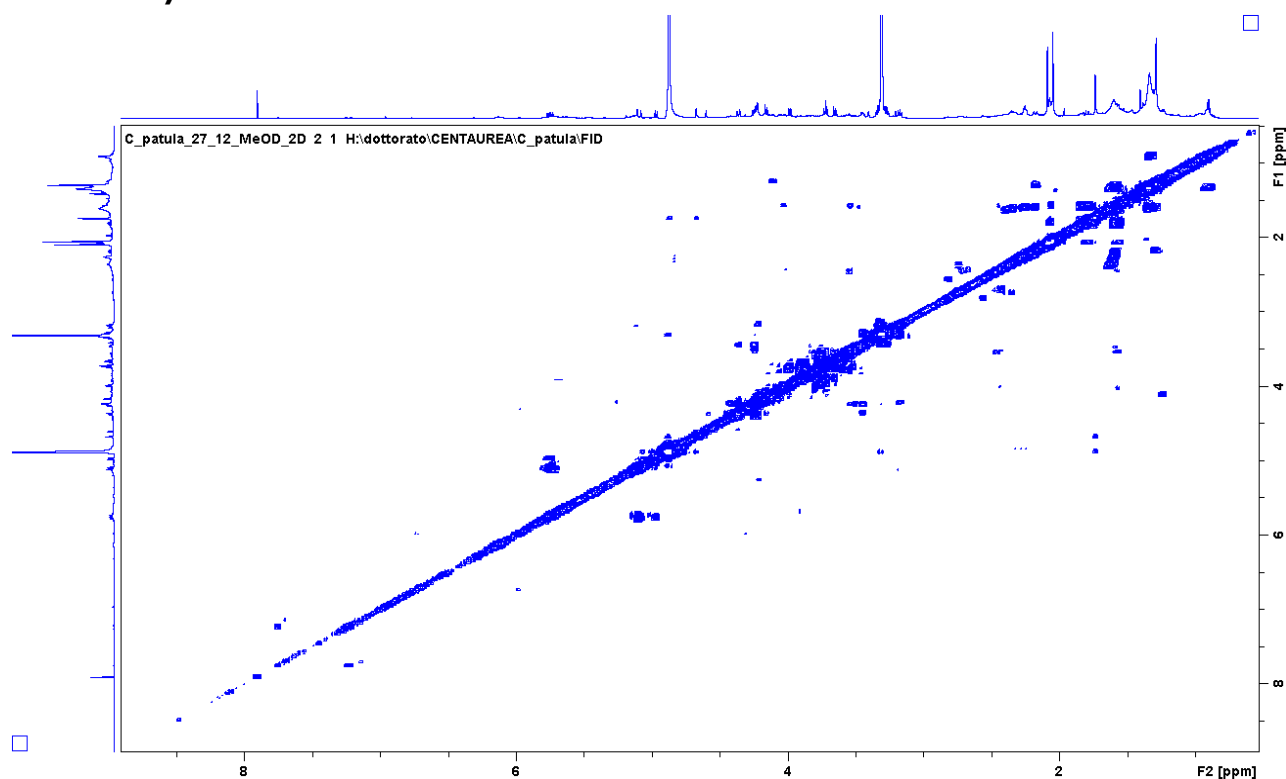


Figure S20: NOESY 2D NMR spectrum of compound 16 (MeOD, 700MHz)

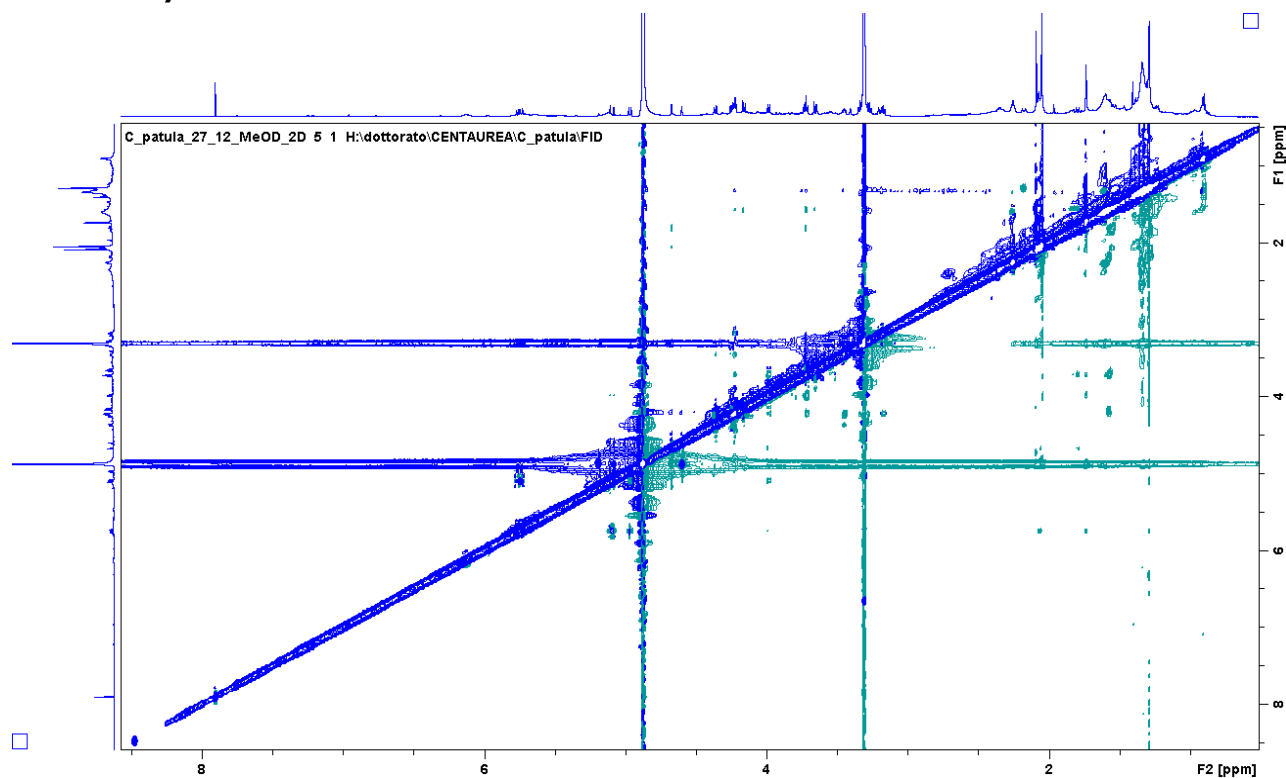


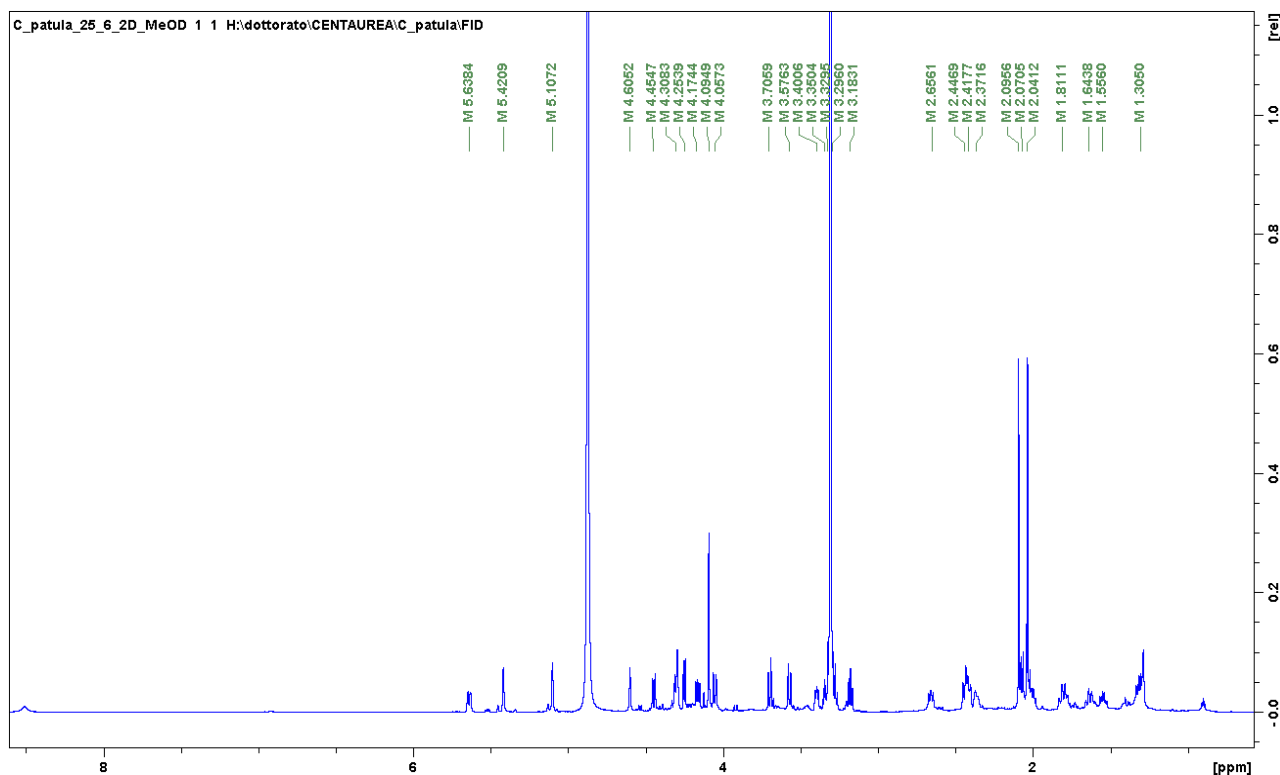
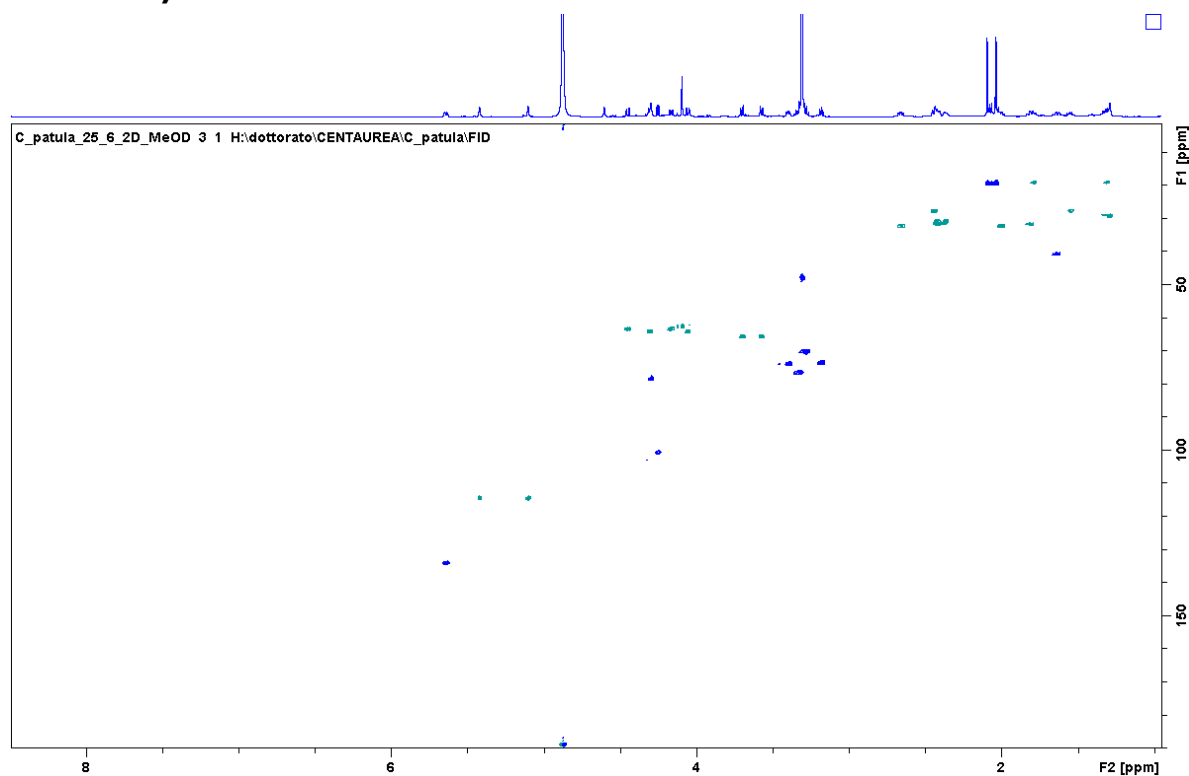
Figure S21: ^1H NMR spectrum of compound 17 (MeOD, 700MHz)**Figure S22: HSQC 2D NMR spectrum of compound 17 (MeOD, 700MHz)**

Figure S23: HMBC 2D NMR spectrum of compound 17 (MeOD, 700MHz)

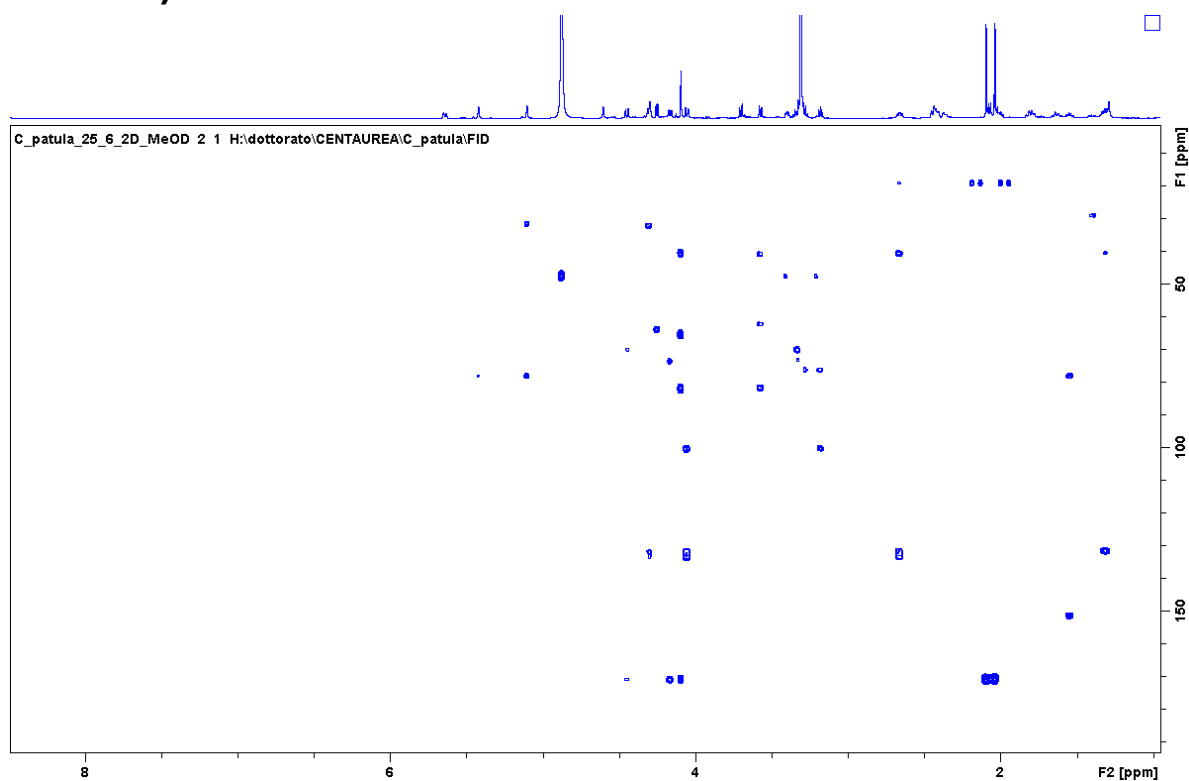


Figure S24: COSY 2D NMR spectrum of compound 17 (MeOD, 700MHz)

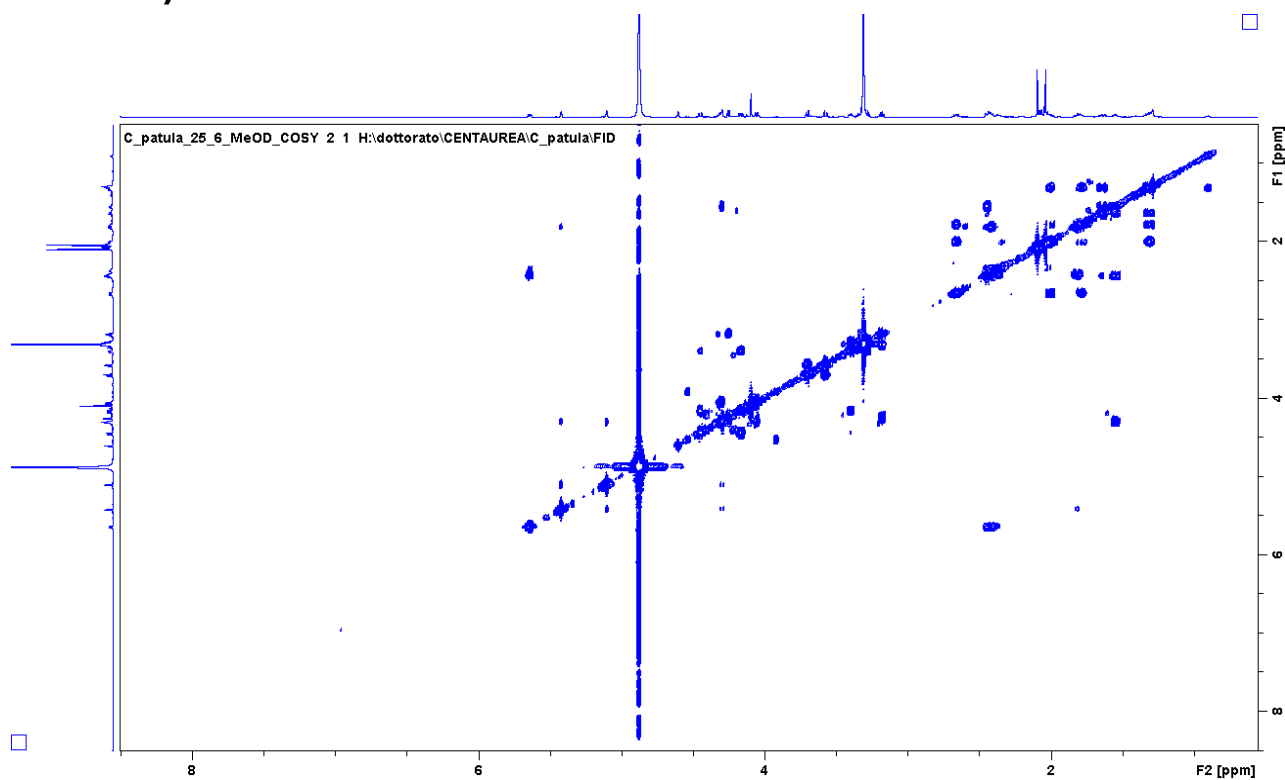


Figure S25: NOESY 2D NMR spectrum of compound 17 (MeOD, 700MHz)

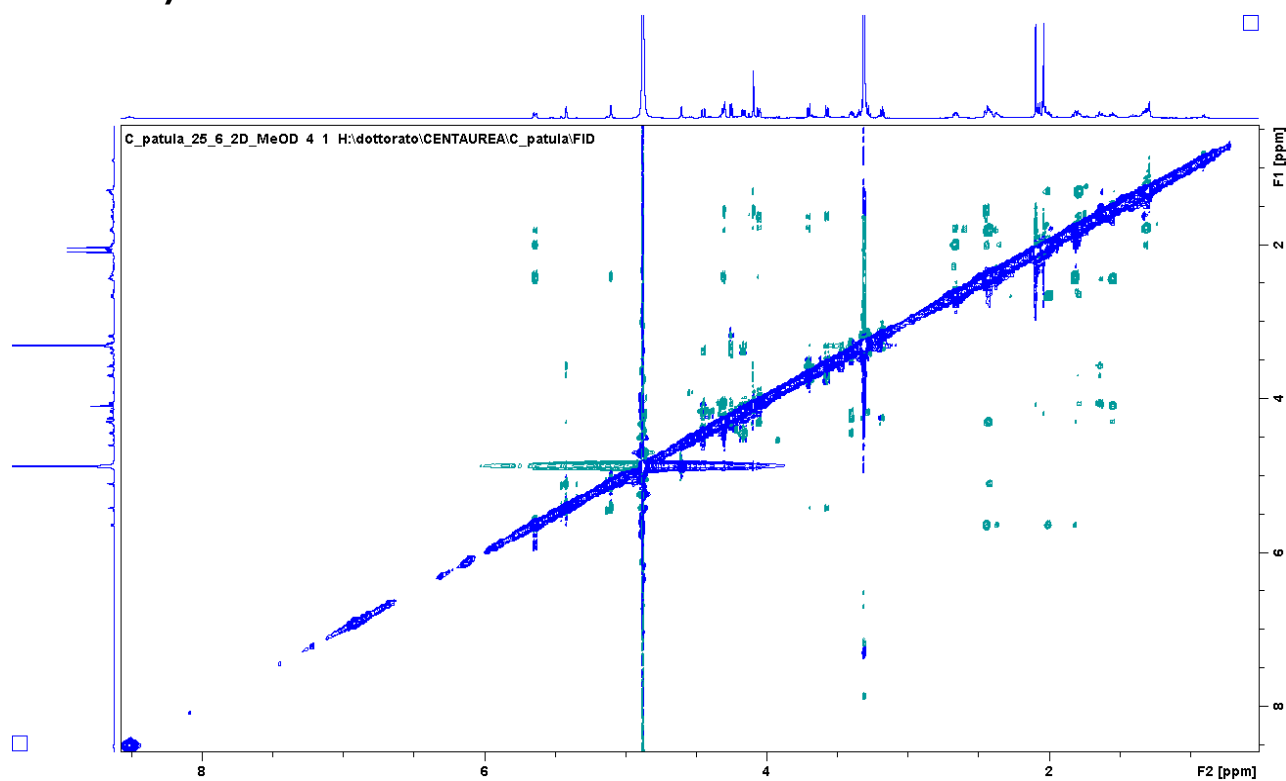


Figure S26: ^1H NMR spectrum of compound 18 (MeOD, 600MHz)

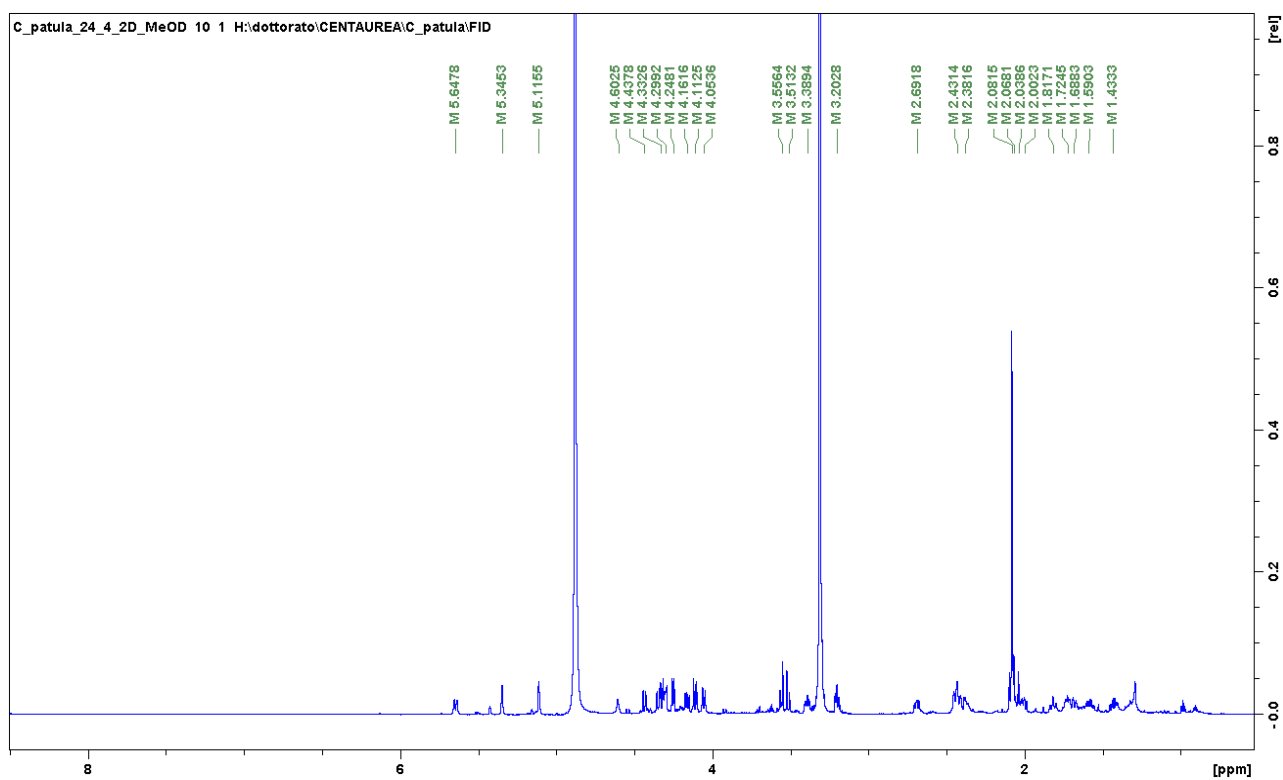


Figure S27: HSQC 2D NMR spectrum of compound 18 (MeOD, 600MHz)

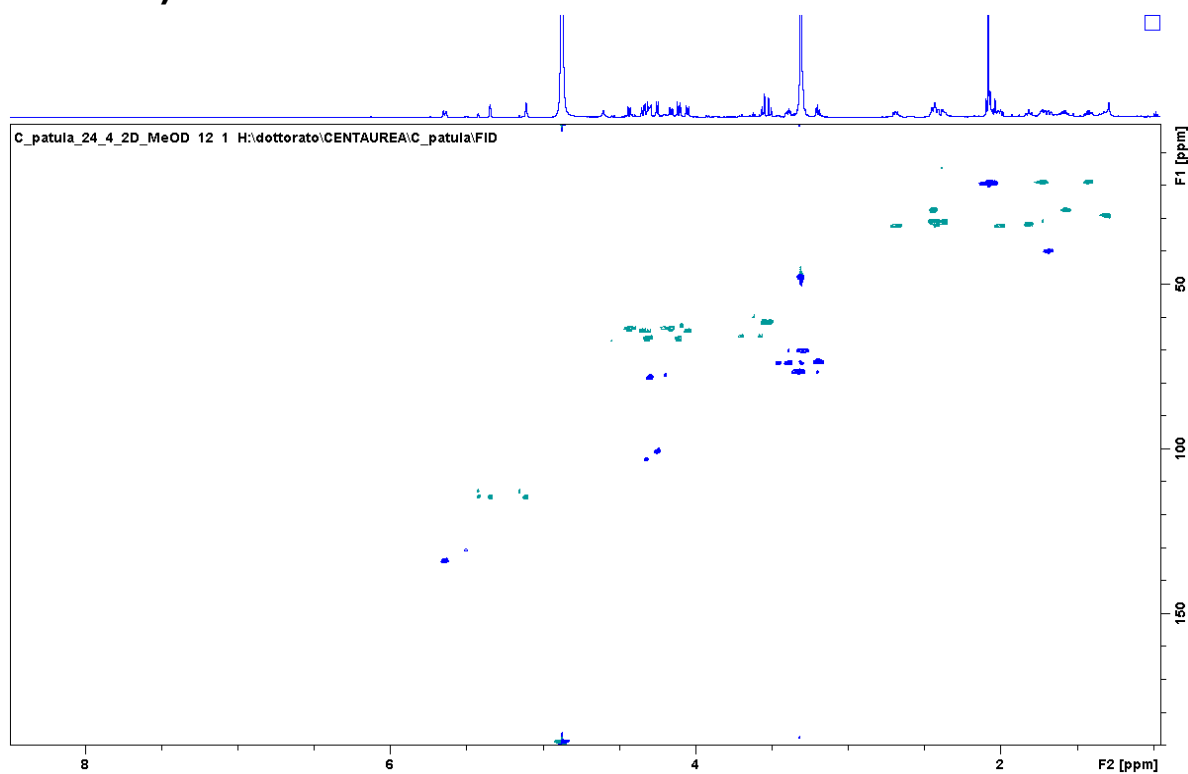


Figure S28: HMBC 2D NMR spectrum of compound 18 (MeOD, 600MHz)

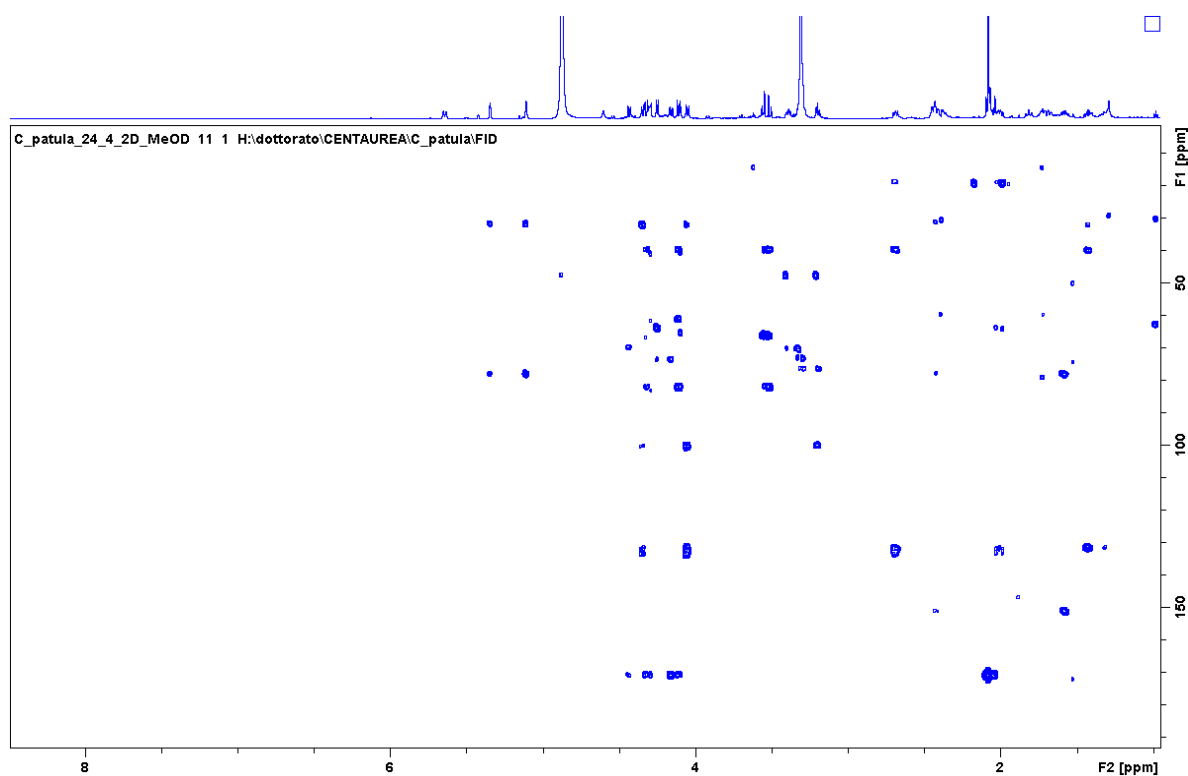


Figure S29: COSY 2D NMR spectrum of compound 18 (MeOD, 600MHz)

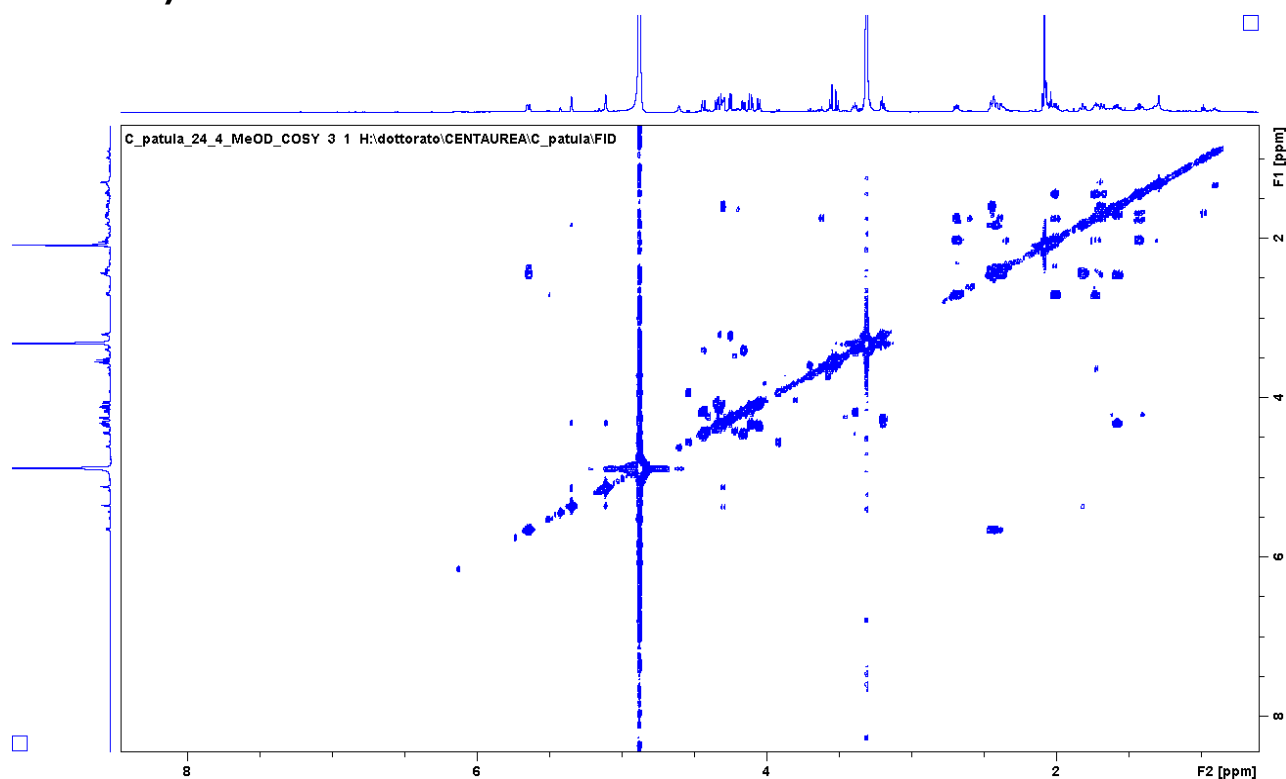


Figure S30: NOESY 2D NMR spectrum of compound 18 (MeOD, 600MHz)

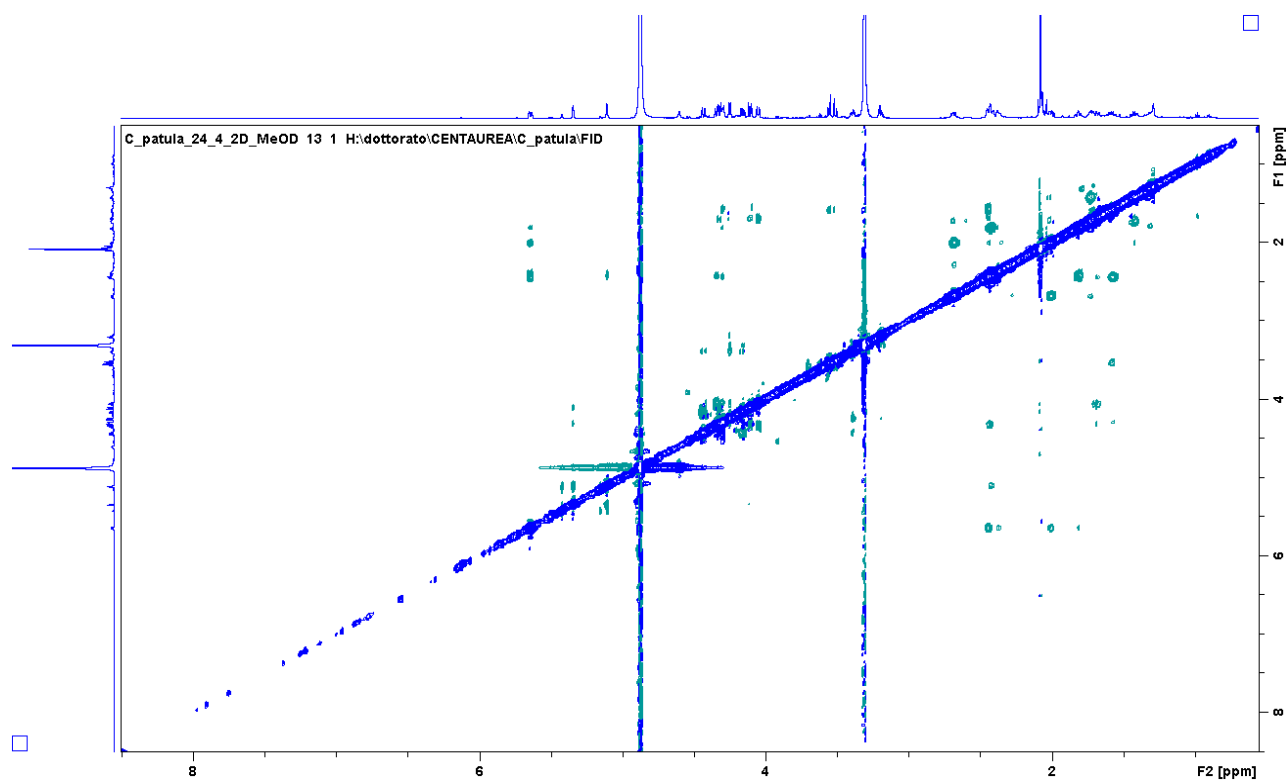


Table S1: Conformational Analysis of the stereoisomers in the MeOH solution of compound 16.

16 S		ΔE^a	ΔG^b	P% ^c
	Conformer 1-1	0.00	0.00	43.54
	Conformer 1-2	0.07	0.31	38.48
	Conformer 1-3	0.88	3.69	9.83
	Conformer 1-4	1.13	4.71	6.51
	Conformer 1-5	2.19	9.16	1.08
	Conformer 1-6	2.59	10.84	0.55
16 R				
	Conformer 2-1	0.00	0.00	66.7
	Conformer 2-2	0.59	2.45	24.8
	Conformer 2-3	1.71	7.16	3.7
	Conformer 2-4	1.72	7.20	3.7
	Conformer 2-5	2.45	10.26	1.1

^aRelative energy (kcal/mol). ^bRelative Gibbs free energy (kcal/mol).

^cConformational distribution calculated at the at the mpw1pw91/6-31G(d) level.

Table S2: The calculation results of compound 16, with mean absolute errors (MAE) values and DP4+ probabilities.

	MAE Value (ppm)		¹³ C Data DP4+ Probability		
	¹³ C MAE	¹³ C CMAE	sDP4+	uDP4+	DP4+
16 S	0.92	1.57	95.85%	99.76%	100.00%
16 R	1.32	3.54	4.15%	0.24%	0.00%

Table S3: Comparisons of ^{13}C values for compound 16.

Carbon	16 S	16 R	Experimental
1	145.6	145.8	146.9
2	109.8	109.8	111.3
3	112.7	112.0	112.0
4	145.0	145.9	146.0
5	52.9	52.6	52.6
6	27.7	27.2	27.6
7	44.3	42.4	41.2
8	21.6	24.2	20.9
9	33.4	33.8	32.6
10	43.5	43.7	42.7
11	74.4	74.7	74.4
12	64.7	66.8	62.6
13	64.4	67.5	65.1
14	69.2	69.8	69.3
15	24.4	26.7	24.1
16	20.7	20.8	19.2
17	170.5	170.4	171.2

Table S4: Conformational Analysis of the stereoisomers in the MeOH solution of compound 17.

17 S		ΔE^a	ΔG^b	P% ^c
	Conformer 1-1	0.00	0.00	28.49
	Conformer 1-2	0.04	0.16	26.75
	Conformer 1-3	0.41	1.71	14.32
	Conformer 1-4	0.58	2.43	10.71
	Conformer 1-5	0.83	3.46	7.05
	Conformer 1-6	0.95	3.96	5.78
	Conformer 1-7	1.06	4.44	4.75
	Conformer 1-8	1.70	7.11	1.62
	Conformer 1-9	2.36	9.86	0.53
17 R				
	Conformer 2-1	0.00	0.00	35.18
	Conformer 2-2	0.36	1.51	19.14
	Conformer 2-3	0.51	2.11	14.99
	Conformer 2-4	0.51	2.12	14.94
	Conformer 2-5	0.72	3.02	10.40
	Conformer 2-6	1.49	6.24	2.84
	Conformer 2-7	1.83	7.67	1.60
	Conformer 2-8	2.16	9.05	0.91
^a Relative energy (kcal/mol). ^b Relative Gibbs free energy (kJoule/mol). ^c Conformational distribution calculated at the at the mpw1pw91/6-31G(d) level.				

Table S5: The calculation results of compound 17, with mean absolute errors (MAE) values and DP4+ probabilities.

	MAE Value (ppm)		¹³ C Data DP4+ Probability		
	¹³ C MAE	¹³ C CMAE	sDP4+	uDP4+	DP4+
17 S	3.33	2.64	0.01%	0.32%	0.00%
17 R	2.12	2.02	99.99%	99.68%	100.00%

Table S6: Comparisons of ^{13}C values for compound 17.

Carbon	17 S	17 R	Experimental
1	144.2	144.6	146.2
2	113.6	111.2	111.4
3	116.1	114.5	112.1
4	143.0	144.1	146.6
5	54.7	52.4	52.5
6	29.1	28.6	27.0
7	45.0	42.8	41.5
8	24.2	24.3	24.7
9	34.4	31.9	32.6
10	50.2	45.7	42.6
11	76.0	73.7	74.0
12	53.4	55.5	46.5
13	69.3	67.5	65.1
14	62.4	61.4	69.3
15	21.9	23.6	24.1
16	23.1	20.6	19.2
17	170.0	170.5	171.3

PART 3: Dietary molecules from food plants active on Metabolic Syndrome

This part of the manuscript describes the studies of two different food plants and the activity of their secondary metabolites on MetS targets.

In the first chapter, the phytochemical characterisation of *Glycyrrhiza foetida*, a less known species of the genus *Glycyrrhiza*, is reported. This plant is a source of amorfrutins, a class of specialised metabolites highly active as PPAR γ stimulators. The minor members of this family of compounds have not been characterised, leaving their structure-activity relationship undefined. Our phytochemical investigation led to the characterisation of several unprecedented amorfrutins and other polyphenols present in the aerial parts of the plant. The amorfrutins were tested on PPAR α and PPAR γ , as well as on GLUT and as mitochondrial stimulators together with the polyphenols, highlighting the presence of several interesting bioactive compounds. This project was carried out in collaboration with several institutes. Professor Soussen Hamammi of the University of Monastir in Tunisia provided us with the plant extracts, while the pharmacological analyses were carried out at the IBC Institute in Pozzuoli, thanks to the collaboration of Dr Rosa Maria Vitale, and at the laboratory of Professor Mariano Stornaiuolo of the Department of Pharmacy of the University of Naples Federico II. The Supplementary Material is reported in Appendix D.

The second chapter deals with the chemical analysis of Aglianico grape pomace. In this project, in collaboration with Professor Martino Forino of the Department of Agronomy of the University of Naples Federico II, I contributed to the LC-MS/MS and NMR analysis of the various extracts and compounds obtained. The main component of the mixture, oleanolic acid, was tested for its antidiabetic properties thanks to the collaboration of Professor Mariano Stornaiuolo of our department.

CHAPTER 1: *Glycyrrhiza foetida* secondary metabolites and their modulation on MetS

1.1 Genus *Glycyrrhiza*

Glycyrrhiza is a genus of flowering plants belonging to the Fabaceae (Leguminosae) family, that includes around 30 species, with *G. glabra*, commonly known as ‘licorice’, being the most known one. The name comes from the greek words *glykos* (‘sweet’) and *rhiza* (‘root’), that indeed is the part of the plant responsible for the characteristic flavor and sweetness associated with licorice. The plants belonging to this genus are native to the Mediterranean basin and in some regions of Asia, as China, and it is widely cultivated also in America (**Figure 1.1**).

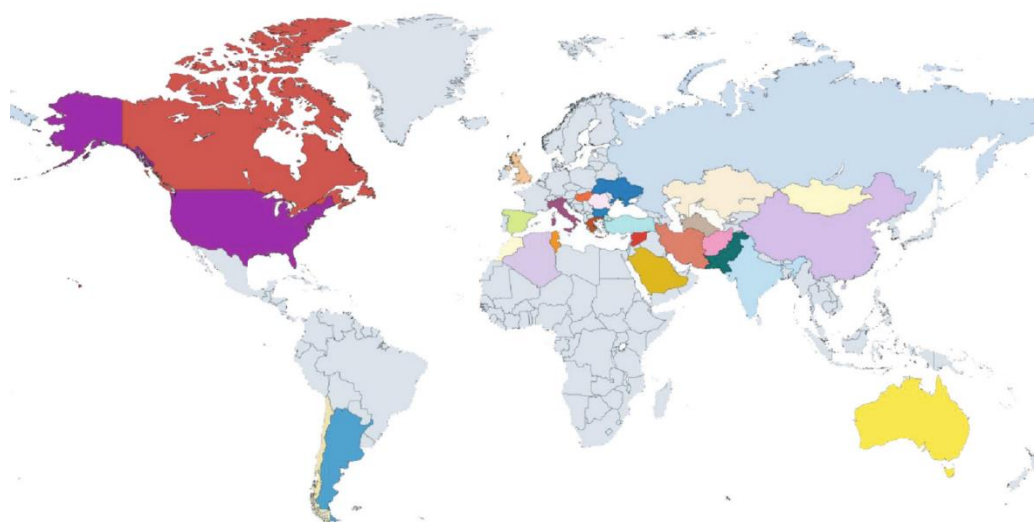


Figure 1.1 *Glycyrrhiza* genus distribution worldwide

Glycyrrhiza plants are perennial subshrub whose size depends on the species and growing conditions, but typically range from 1 to 2.5 m in height. The stems are erect and usually display multiple branches. The leaves are alternate and imparipinnate, with small lanceolate-shaped leaflets. Flowers are arranged in long stalks inflorescence, typical of its family, whose colors range from whitish to pale purple or blue, depending on the species (**Figure 1.2**). The fruits appear as flattened brownish pods that contain several reniform seeds, which are small and smooth, whose colors range to brown, black, or deep grey. *Glycyrrhiza* plants have a highly developed stoloniferous woody roots, thick and branched with yellowish to red color (Sharifi-Rad et al., 2021).

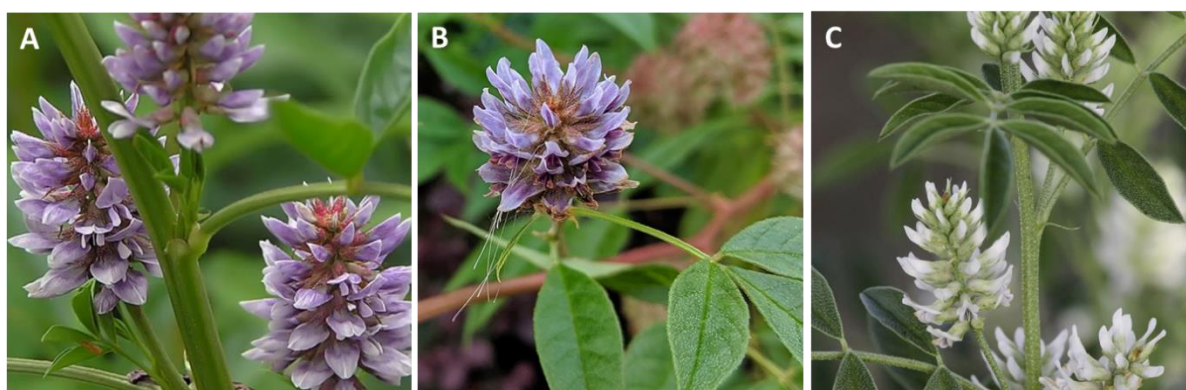


Figure 1. 2 Flowers from different *Glycyrrhiza* species: (A) *G. glabra*; (B) *G. uralensis*; (C) *G. foetida*.

Glycyrrhiza species are not only used in the food industry for their flavour, but also in the pharmaceutical industry and as dietary supplements. Plants of genus *Glycyrrhiza* contain a wealth of bioactive compounds, but only a few species have a long history of use in folk medicine, including *G. glabra*, *G. inflata* and *G. uralensis*. Health benefits and medicinal applications include skin health, antitussive and expectorant properties, digestive and antiulcerative effects, hepatoprotective activity and hormonal balance properties (Pastorino et al., 2018). However, it should be noted that high consumption of liquorice has been associated with side effects due to potential inhibitory activity on different subtypes of cytochrome P450 enzymes, with IC_{50} values depending on the species (Husain et al., 2021).

The phytochemical composition of *Glycyrrhiza* species comprises saponins, as glycyrrhizin and its analogues, flavonoids, especially isoflavones and flavanols, as liquitrin and its derivatives, along with coumarins and chalcones.

In previous years, *G. uralensis* and *G. glabra* were the species examined for their health benefits on MetS. Indeed, the anti-obesity effect of *G. uralensis* extract has been evaluated in vivo, alone and in combination with metformin, showing promising results in preventing obesity complications such as dyslipidaemia, hyperglycaemia, inflammation and fatty liver (Hong et al., 2023). On the other hand, the anti-diabetic properties of *G. glabra* have been evaluated in several studies, for example its alcoholic root extract can improve glucose and lipid blood levels (Qureshi et al., 2020; Sen et al., 2021), together with the anti-obesity property displayed in vivo in combination with Red Ginseng extract (Zheng et al., 2020). The bioactive ingredient is attributed to be glycyrrhizin, one of the main constituents of the plant's root. This triterpenoid is able to improve several MetS parameters, including insulin resistance, oxidative stress, and

blood glucose and lipid levels, probably via PPAR γ modulation (Sil et al., 2013; Behl et al 2022). Glabridin, one of the most abundant flavonoids extracted from the roots of *G. glabra*, also has shown hypoglycaemic effects and can reduce body weight (Pastorino et al., 2018). The chemical structures of these two bioactive ingredients are reported in **Figure 1.3**.

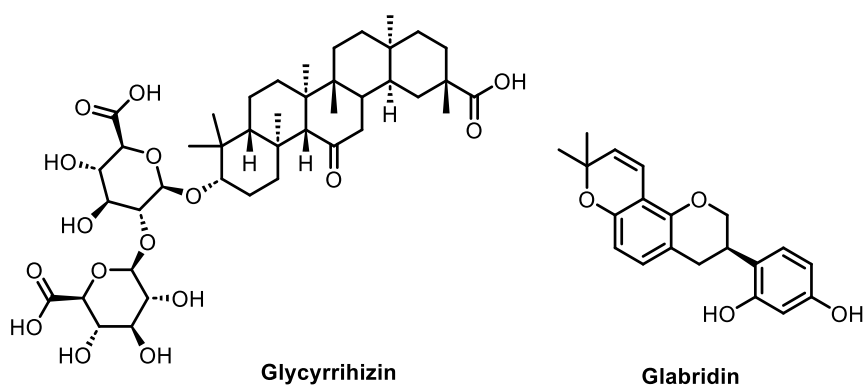


Figure 1. 3 Chemical structure of antidiabetic bioactive ingredients in *Glycyrrhiza* sp.

1.2 Amorphutins

Amorfrutins are an intriguing class of secondary metabolites that were named after *Amorpha fruticosa*, a shrub from Fabaceae family native to North America, where these compounds were firstly isolated from in 1981 (Mitscher et al., 1981). In the same year, amorfrutins were also isolated from *G. acanthocarpa*, an Australian species of genus *Glycyrrhiza* (Ghisalberti et al., 1981). These class of compounds were also described in other two *Glycyrrhiza* species including the leaves of *G. lepidota*, native to America (Manfredi et al., 2001) and from the edible roots of *G. foetida*, native to northern Africa (Weidner et al., 2012) and later in the south African *Helichrysum umbraculigerum* (Pollastro et al., 2017). Interestingly, amorfrutins have not been reported in *G. glabra* and *G. uralensis*. Moreover, only a handful of natural amorfrutins had been characterized and named with a confused structure-independent and partly overlapping dual system of numbers and letters ([A (= 1), B, 2, 3, 4, C, D]). Amorfrutins structure clearly resemble the one of phytocannabinoids, from which, in fact, they share some features in their biogenetic origin. Other source of amorfrutins are plants belonging to genus *Radula* (Asakawa et al., 1995), and some analogues were isolated from peanut (*Arachis hypogea* L.) seeds infected with an *Aspergillus flavus* fungal strain (Sobolev et al., 2016).

1.2.1 Amorphutins biogenesis

Amorphutins are non-cannabis-derived phytocannabinoids characterized by a salicylate core variously substituted. Biogenetically, they have a hybrid biosynthetic nature based on a 2,4-dihydroxybenzoic acid scaffold of polyketide derivation decorated with two para-orientated substituent, a prenyl moiety and a phenethyl or pentyl side chain, whose nature mirror the ketide starter (**Figure 1.4**). Thus, amorphutins can be sorted out into two main classes, namely pentyl-type (cannabis-type) and phenethyl-type, also called bibenzyl amorphutins.

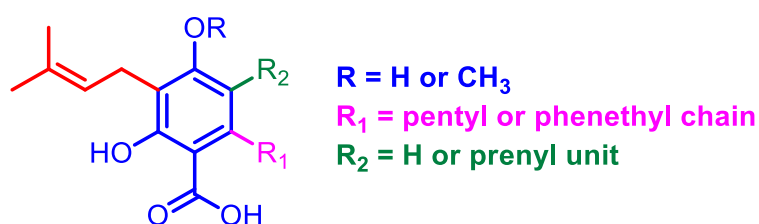


Figure 1. 4 Amorphutin core.

Pentyl-type amorphutins are similar to the acidic phytocannabinoids from *Cannabis*, and are presumably derived from an hexanoyl starter, while phenethyl-type phytocannabinoids derive from a dihydrocinnamyl starter and have never been reported from *Cannabis*. With the exception of amorphutin 2 and its analogues, the majority of amorphutins belong to the phenethyl type and are structurally related to pre-cannabigerol-O-methyl ether (**Figure 1.5**). Although a few decarboxylated analogues have been described, the majority of amorphutins are in their acidic form. In fact, similarly to phenethyl cannabinoids, for unclear reasons amorphutins seem to be more resistant to decarboxylation compared to the alkyl phytocannabinoids. O-prenylation, along with meta-hydroxylation, has also been reported in a bibenzyl cannabinoid (Srebnik et al., 1984) from *Glycyrrhiza lepidota* (Manfredi et al., 2001).

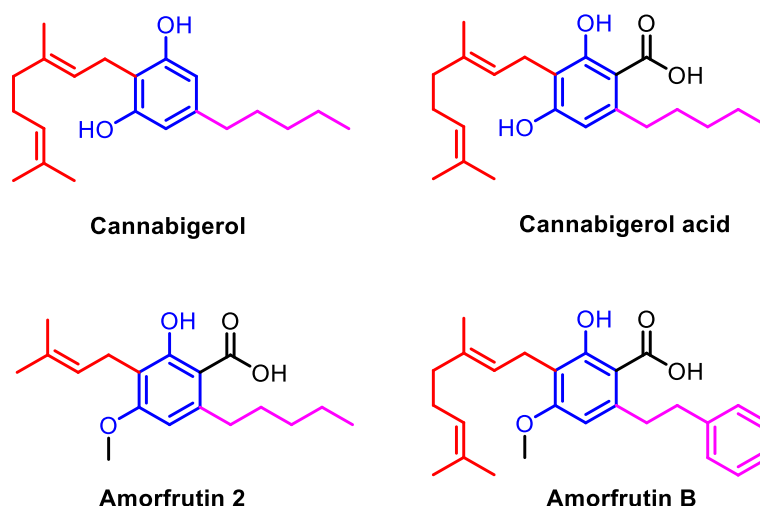


Figure 1. 5 Structural similarities between phytocannabinoids and amorfrutins.

1.2.2 Amorfrutins bioactivity

Although amorfrutins were firstly described as mild antibacterial agents (Mitscher et al., 1981), they have lately demonstrated a great potential on PPAR γ , a challenging target for medicinal chemistry that plays a crucial role in type 2 diabetes mellitus treatment. In fact, the only drug currently available on the market are thiazolidinediones (TZDs), a controversial class of drugs that have several downsides, including the high cost and the severe side effects, that limits their clinical application (Kung et al., 2012). In particular, a crystallographic analysis shed light on the interaction of amorfrutins with PPAR γ , whose helix H3 and the β -sheet, placed at the entry side of the ligand binding domain, interact with amorfrutins (de Groot et al., 2013). This binding mode is distinct from that of glitazones and is associated with the activation of a different subset of genes resulting in a different profile of protein expression, typical of PPAR γ partial agonists. This crucial difference seems to explain the different safety profile that amorfrutins exhibit, as they appear to lack the undesirable side effects often seen with commercial drugs or other experimental synthetic ligands, including fluid retention, cardiovascular complications and bladder cancer (Weidner et al., 2012, de Groot et al., 2013). Thus, amorfrutins are highly promising selective PPAR γ activators in white adipose tissue and improve insulin resistance. Among the compounds reported in the literature, the most powerful compound on PPAR γ is amorfrutin B, an analogue embedding a geranyl unit, which is characterized by a K_i of 19 nM, and shows a great insulin-sensitivity effect in vivo. Other amorfrutins, like amorfrutin A and for the diprenylated analogue amorfrutin C, exhibit a

reduced potency compared to amorfrutin B, showing an affinity constant on PPAR γ around 0.2 μ M. However, amorfrutins seem to be more selective on PPAR γ since their K_i values on PPAR α and PPAR δ/β are around 27 μ M (Weidner et al., 2012). Other properties ascribed to amorfrutins are the inhibition of the glycolytic enzyme glyceraldehyde 3-phosphate dehydrogenase (GAPDH) which is involved in neuronal death (Choi et al., 2015), the inhibition of NF- κ B and iNOS activities related to inflammatory pathways (Fuhr et al., 2015), and of the cysteine protease ATG4B, a key regulator of autophagy (Sauer, 2014). These additional targets make amorfrutins particularly suitable for a multitarget approach for the treatment not only of diabetes, but also other pathologies related to chronic inflammation.

1.3 *Glycyrrhiza foetida* phytochemical composition

Glycyrrhiza foetida is a less investigated species of genus *Glycyrrhiza* that grows in Southern Spain and northern Africa (**Figure 1.6**). It has been used to treat hepatic and metabolic disorders, and is one of the few species of its genus from where amorfrutins were isolated from (Weidner et al., 2012). Given the great potential of this class of compounds on diabetes, we had a strong interest in investigating this species. Moreover, the minor members of amorfrutins had not been described yet, leaving unexplored the structure-activity-relationship of these metabolites on PPARs. In addition, in the literature a comprehensive phytochemical characterization of *G. foetida* was missing. Thanks to our collaboration with the University of Monastir, we have had the opportunity to study a Tunisian sample of *G. foetida*, conducting a detailed analysis of its aerial parts organic extract. Firstly, we have focused our attention on the amorfrutins composition of the plant and their PPAR γ/α agonism (Serino et al., 2023). Afterwards, we expanded our investigation on the polyphenol composition of the plant, along with deepening in the modulatory activity of all the isolated secondary metabolites on other targets involved in MetS (Al-Hmadi et al., 2024).



Figure 1.6 *Glycyrrhiza foetida* distribution

1.3.1 Amorfrutin profiling

Air-dried, ground, and powdered aerial parts of *G. foetida* were exhaustively extracted with MeOH/CH₂Cl₂ 1:1 and the obtained extract was subjected to a preliminary LC-MS/MS analysis to carry out a first dereplication step (**Figure 1.7**). A surprising profligacy of amorfrutins analogues was detected in the extract that were tentatively assigned through the fragmentation pattern analysis and their mass weight, and later were isolated and structurally characterized by NMR. The putative identification of the amorfrutins is reported in **Table 1.1** (for details see Appendix D) The most abundant compound of the extract was annotated as amorfrutin A (*R_t* 10.82 min) based on its molecular weight (*m/z* 341 [M+H]⁺), the MS/MS fragmentation pattern (Chen et al., 2016) and the co-injection with a standard available in our lab. Its decarboxylated analogue was also annotated, along with amorfrutin 2 (*R_t* 11.57 min), which appeared to be the second most abundant compound of the extract, thanks to the comparison with a standard. Also amorfrutin 3 and decarboxyamorfrutin 2 were confidently annotated thanks to the fragmentation pattern and molecular weights, while the peak at 15.61 min could be either assigned to the geranylated amorfrutin B or to the bisisoprenylated amorfrutin C. Most of the other detected peaks were likely to be new amorfrutins, as suggested by their fragment ions, but a detailed NMR analysis was required to undoubtedly characterize them. The annotation of the five known amorfrutins was also confirmed by isolation and NMR-based structural identification.

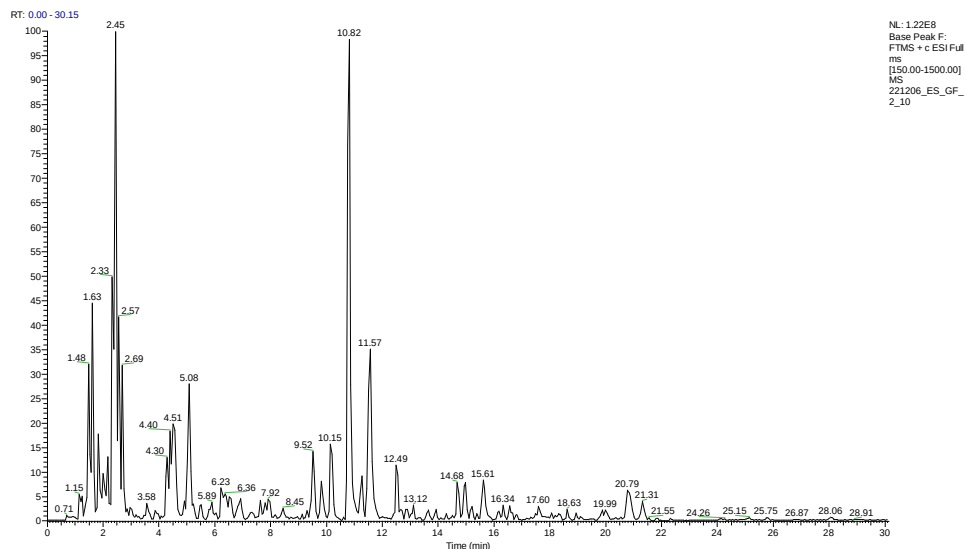


Figure 1. 7 BPC of *G. foetida* aerial parts aerial parts extract

Taking advantage of the information coming from the LC-MS profile, we carried out a series of chromatographic purifications on MPLC and HPLC (see Experimental Section) that led to the isolation of 29 pure compounds. All the compounds were analyzed and identified through NMR and HRESI-MS experiments, identifying the known through spectra data comparisons with the ones reported in literature, or characterizing for the first time the undescribed compound with in-depth 2D NMR analysis. We have firstly focused our attention on the amorfrutins compounds, which appeared to be the most representative class of compounds of the extract. The phytochemical characterization led to the isolation of 17 amorfrutins, 11 belonging to the phenethyl series (**1-11**), while 6 exhibiting the pentyl side chain (**12-17**) (**Figure 1.8**). The isolated known metabolites included four phenethyl type amorfrutins, amorfrutin A (**1**) (Mitscher et al., 1981), decarboxyamorfrutin A (**2**) (Mitscher et al., 1981), amorfrutin 3 (**3**) (Weidner et al., 2012), amorfrutin C (**11**) (Weidner et al., 2016), together with two pentyl type amorfrutins, amorfrutin 2 (**12**) and decarboxyamorfrutin 2 (**13**) (Ghisalberti et al., 1981). Among the phenethyl series, 7 compounds were unprecedented in the literature, therefore we have conducted a complete set of 2D NMR analysis to determine their structural features.

PART 3

Table 1.1 Identified components of the *G. foetida* aerial parts analyzed via LC-MS/MS. Compounds are listed in order of LC-MS elution. All mass peaks are [M+H]⁺ adducts

Assignment	Formula	R _t (min, Fig.1)	Precursor ion (m/z)	Δppm	Fragments
Not Assigned (tentatively assignable to compounds 6 or 9)	C ₂₁ H ₂₄ O ₅	2.33	357.1694	-0.617	325.1409
Not Assigned (identified after isolation as compound 16)	C ₁₈ H ₂₆ O ₅	2.45	323.1850	-1.022	305.1747; 287.1635 251.1285
Amorfrutin 3 (3)	C ₂₁ H ₂₄ O ₅	2.57	357.1695	-0.533	339.1592; 321.1487; 285.1117; 325.1407
Not assigned (identified after isolation as compound 15)	C ₁₈ H ₂₆ O ₅	2.69	323.1849	-1.208	305.1744 291.1948 287.1644
Not assigned (identified after isolation as compound 8)	C ₁₈ H ₁₆ O ₄	3.58	297.1119	-0.658	279.1014; 253.1222; 207.1381 175.0754 105.0695
Decarboxyamorfrutin A (2)	C ₂₀ H ₂₄ O ₂	10.15	297.1847	-0.729	265.1586; 241.1224; 229.1227; 181.1015
Amorfrutin A (1)	C ₂₁ H ₂₄ O ₄	10.82	341.1741	-0.729	323.1638, 285.1119; 267.1012; 241.1218
Decarboxyamorfrutin 2 (13)	C ₁₇ H ₂₆ O ₂	11.26	263.2003	-0.899	231.1747 207.1378 195.1379 137.0594
Not assigned (identified after isolation as compound 10)	C ₂₄ H ₃₀ O ₂	11.45	351.2317	-0.560	295.1693; 283.1693; 227.1073 239.1083;
Amorfrutin 2 (12)	C ₁₈ H ₂₆ O ₄	11.57	307.1901	-0.865	289.1796; 251.1279; 233.1171
Not assigned (identified after isolation as compound 17)	C ₂₁ H ₃₂ O ₂	12.49	317.2471	-1.282	261.1847; 249.1848; 205.1222 193.1222
Amorfrutin C or B (identified after isolation as amorfrutin C)	C ₂₆ H ₃₂ O ₄	15.61	409.2370	-0.772	391.2269; 353.1741; 309.1869

PART 3

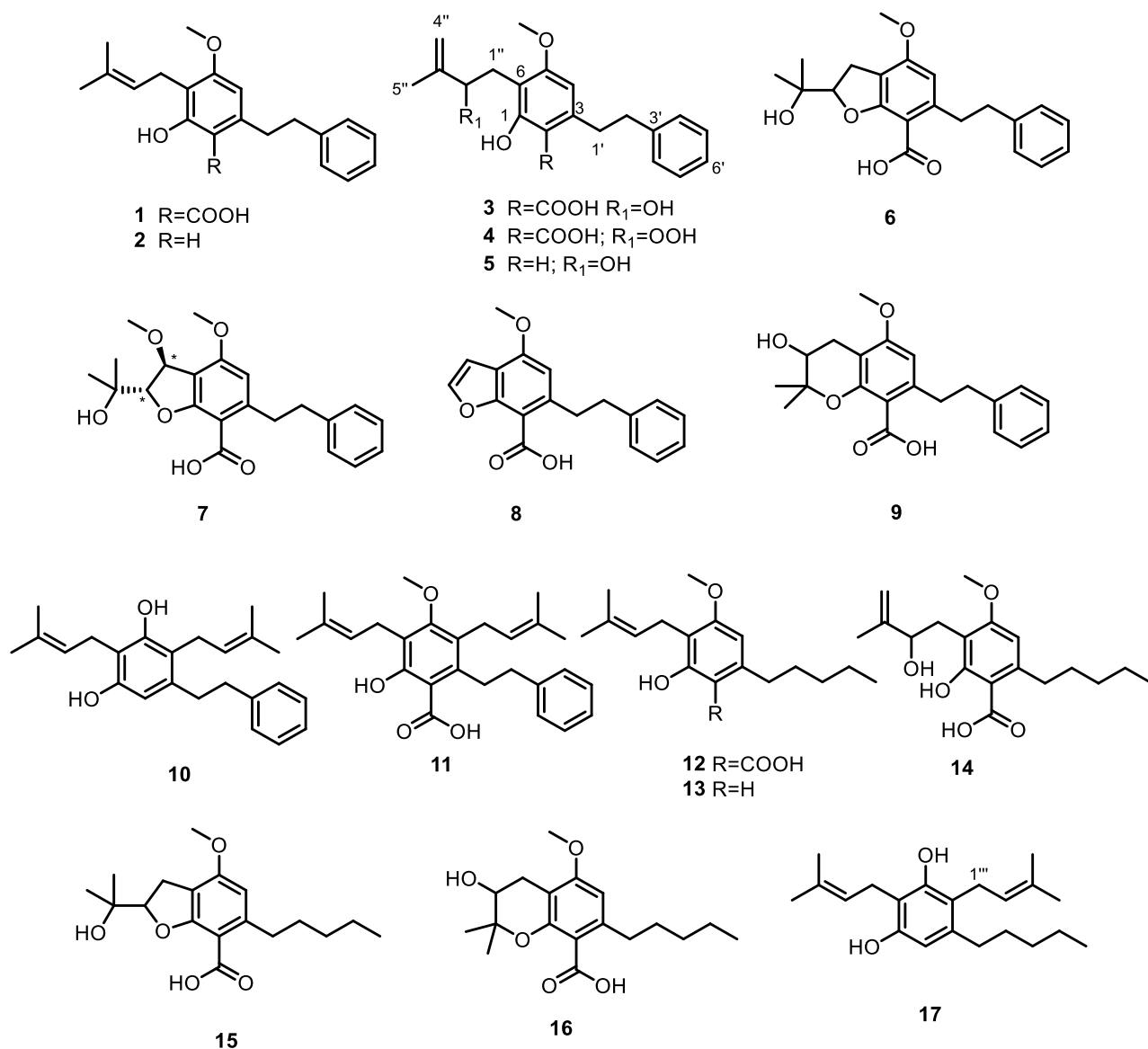


Figure 1.8 Amorfrutins isolated from *G. foetida* aerial parts

Amorfrutin E (**4**) showed a molecular formula of C₂₁H₂₄O₆ (HRESIMS, *m/z* 395.1469). In the ¹H NMR spectrum of **4** a single aromatic proton (δ_H 6.20) was detected, suggesting a carboxylated amorfrutin derivative, and those of the phenylethyl and monoterpenyl units. In particular, this latter moiety signals paralleled exactly that of amorfrutin 3 (**3**), but the evidently deshielded H-2'' (δ_H 4.44) and especially C-2'' (δ_C 86.5) pointed to the presence of a hydroperoxide group, in line with the additional oxygen atom indicated by the molecular formula. Similarly to amorfrutin 3 (**3**), the molecule was optically inactive and, therefore, racemic.

Amorfrutin N (**5**), molecular formula $C_{20}H_{24}O_3$ deduced by HR-ESIMS, was isolated in minute amounts and its structure inferred by spectral data and comparison with known amorfrutins. In particular, its ^1H NMR spectrum showed signals of two singlets in the aromatic region (δ_H 6.26 and 6.49) and one -OMe (δ_H 3.74), suggesting that **5** is a monomethoxylated and decarboxylated amorfrutin. In addition, typical signals of the phenylethyl moiety and side chain signals overlapping with those of amorfrutin 3 (**3**) were detected. The structure of **5** was secured and fully characterized by COSY, HSQC, and HMBC experiments. Thus, amorfrutin N (**5**) is a decarboxylated analogue of amorfrutin 3. Since **5** showed no optical rotation, it is racemic at C-2".

Amorfrutin F (**6**), an optically inactive compound with the molecular formula $C_{21}H_{24}O_5$ (HRESIMS, m/z 357.1699), was clearly a phenethyl amorfrutin analogue, exhibiting a deshielded methylene (δ_H 3.12), an oxymethylene (δ_H 4.84) and two methyl singlets, suggesting a modification of the terpene moiety. These signals could be accommodated into a dihydrobenzofuran moiety based on the HMBC correlations shown in **Figure 1.9**. Particularly significant were those from H-2" to C-1, indicating the presence of the oxygen bridge, and from H₃-4" and H₃-5" to C-2". Thus, structure **6** was assigned to amorfrutin F.

Amorfrutin G (**7**) was characterized by $C_{22}H_{26}O_6$ as molecular formula (HRESIMS m/z 409.1622). Its ^1H and ^{13}C NMR spectra indicated that **7** was a methoxy-substituted analogue of compound **6**. The presence of an additional methoxy group was clearly visible from the ^1H NMR spectrum, that showed an additional 3-proton singlet along with the lack of the methylene H₂-1", replaced by an oxymethine (δ_H 5.10). The diagnostic HMBC cross-peak between the additional methoxyl group with C-1" (**Figure 1.9**) unambiguously confirmed its position. The terpene moiety of **7** displayed two adjacent stereogenic centers whose stereochemical features needed to be investigated. The NOESY correlations of H-1" with H₃-4"/H₃-5" and of H-2" with OMe-1" indicated the *trans* orientation of the two substituents on the five-membered ring and, given the absence of optical activity, the compound must be in racemic form.

Amorfrutin H (**8**), a colorless amorphous solid with a molecular formula of $C_{18}H_{16}O_4$ (HRESIMS, m/z 297.1126), is characterized by the presence of a benzofuran skeleton, likely derived from the cleavage of the terpene moiety. The ^1H NMR spectrum of **8** showed the typical signals of the phenylethyl moiety, the H-4 singlet of the resorcinol moiety, but no signals attributable to a

terpene portion. However, two broad singlets at δ_{H} 7.62 and 6.87 which could be attributed to a furan moiety, in agreement with the molecular formula, were detected. This structural assignment was secured by HMBC correlations of H-1" with C-1, C-2, and C-3 and of H-2" with C-1 and C-2. This compound has been reported as a synthetic product in a patent (Sauer et al., 2014), but no NMR data were available.

The racemic amorfrutin I (**9**), $\text{C}_{21}\text{H}_{24}\text{O}_5$ by HRESIMS (m/z 357.1695), shares with amorfrutin F (**6**) the same molecular formula and the phenylethyl resorcinol core, but clear differences in the resonances of the isoprenoid moiety were evident. In particular, the deshielded C-3" (δ_{C} 80.0) and the shielded H-2" (from δ_{H} 4.84 to 3.88) pointed to a hydroxylated chroman-type skeleton. The inspection of 2D NMR spectra (COSY, HSQC, and HMBC) led to the full characterization of this previously undescribed amorfrutin (**Figure 1.9**).

Decarboxydemethoxyamorfrutin C (**10**), $\text{C}_{24}\text{H}_{30}\text{O}_2$ by HRESIMS (m/z 351.2312), is a close analogue of diprenylated amorfrutin C (**11**) lacking the carboxylic group at C-6 (δ_{H} 6.30) and with an -OH group in place of the -OMe at C-3. The full NMR assignment of **10** was guided by comparison with data of **11** (Weidner et al., 2016) and inspection of a full set of 2D NMR spectra.

Among the six pentyl-type amorfrutins (**12-17**), we characterized four new analogues (**14-17**) embedding terpenoid moieties already detected in the phenethyl analogues.

Amorfrutin J (**14**), is the racemic pentyl analogue of amorfrutin 3 (**3**) (Weidner et al., 2012), with $\text{C}_{18}\text{H}_{26}\text{O}_5$ as molecular formula (HRESIMS m/z 345.1664). Accordingly, full NMR assignment based on the COSY, HSQC, and HMBC spectra indicated that the two compounds differed only in the side chain, with the pentyl H/C signals replacing those of the phenylethyl one. This compound has been mentioned in the proprietary literature as a synthetic derivative, but its NMR data were not available for comparison.

Based on the same arguments, we assigned amorfrutin K (**15**) (m/z 323.1852) as the pentyl analogue of amorfrutin F (**6**), amorfrutin L (**16**) (m/z 345.1684) as the pentyl analogue of amorfrutin I (**9**), and amorfrutin M (**17**) (m/z 317.2469) as the pentyl analogue of the diprenylated **10**.

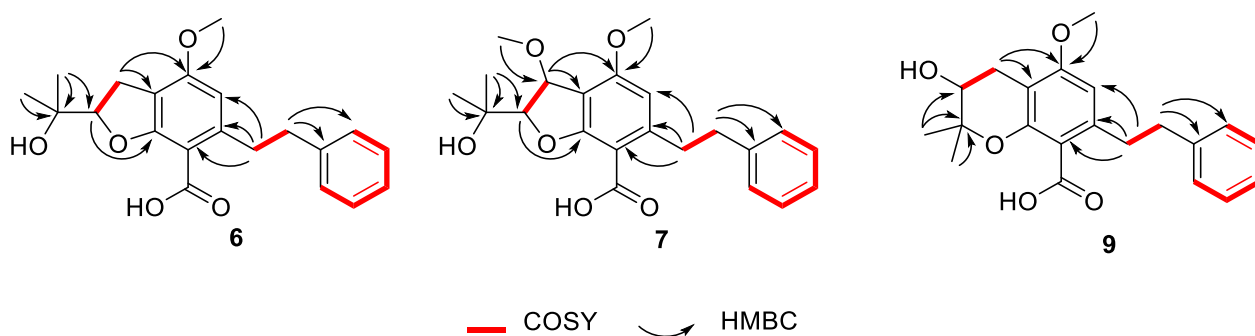


Figure 1.9 COSY and key HMBC correlations of amorfrutins F (6), G (7) and I (9).

We can assume that all the newly described amorfrutin origin from the archetypal member of amorfrutin A and amorfrutin 2 for the analogues of the phenethyl and pentyl series, respectively (**Figure 1.10**). The ring opening of the epoxide and the concomitant nucleophilic attack at position 2'' or 3'' led to the formation of a five- or six-membered ring. The formation of the benzofuran ring could result from the loss of an acetone unit. The oxidation in the terpenic moiety was associated with the formation of stereogenic centers, but all the chiral compounds were obtained in racemic form. This could be due either to the poor enantioselectivity of the oxidase biosynthetic steps or to a non-enzymatic autoxidation of the prenyl group.

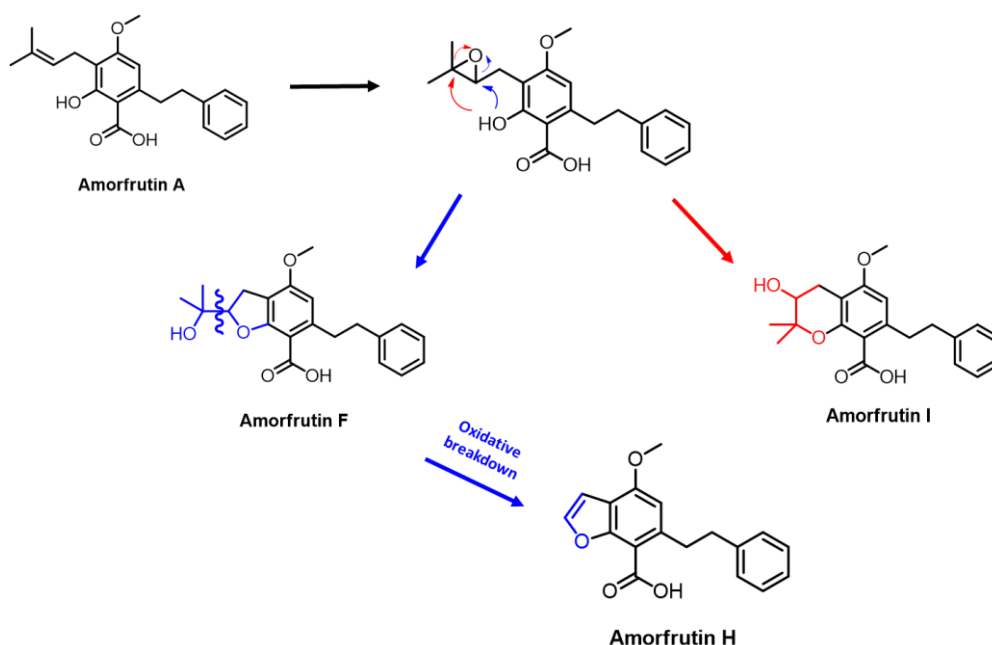


Figure 1.10 New amorfrutins postulated biogenesis

1.3.2 Polyphenols profiling

In order to expand our knowledge on the phytochemical composition of *G. foetida*, we have carried out also an LC-UV-MS/MS analysis in negative mode to have a preliminary profile of the polyphenols of the plant (**Figure 1.11**). From the data in literature, genus *Glycyrrhiza* is rich in different kind of flavonoids that could be tentatively annotated not only with the fragmentation pattern analysis, but also comparing their UV profiles. In fact, some polyphenols share a similar fragmentation pattern, like flavanones and chalcones, but they significantly differ in their UV profiles (Xu et al., 2013). An elaborate analysis of the different UV spectra and fragmentation patterns allowed us to annotate the main polyphenols present in the extract. In fact, besides amorfrutins, which exhibit a typical UV profile with λ_{max} 220, 270, and 305 nm, in agreement with the literature (Chen et al., 2016), the presence of flavanones was strongly inferred by the acquired UV adsorption, with λ_{max} 220, 295, and 345 nm (**Table 1.2**).

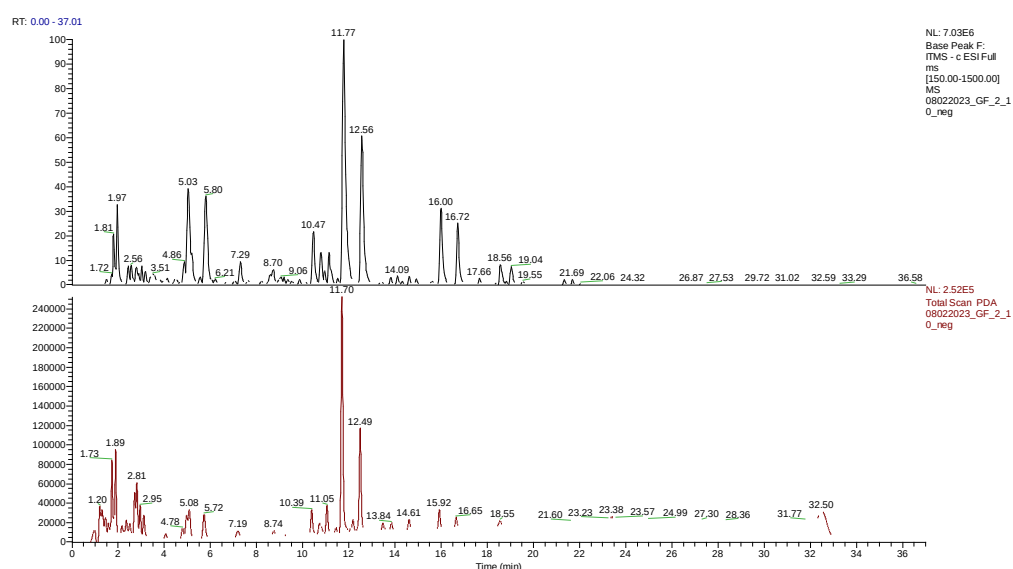


Figure 1.11 LC-MS base ion peak chromatogram (up) and LC-UV (bottom) chromatogram of the extract obtained from the aerial parts of *G. foetida*.

The flavanones lonchocarpol A (R_t 10.47 min) and hiravanone (R_t 10.77 min), and the isoflavone biochanin A (R_t 3.17 min) were annotated, along with the most abundant amorfrutins. All the annotated polyphenols were then confirmed after isolation through an extensive NMR analysis.

Table 1.2 Annotation of the major peaks of the *G. foetida* extract through UV analysis

Assignment	R_t (min)	Precursor ion (m/z)	Fragments	UV (λ_{max} in nm)
Amorfrutin 3	3.02	355.30	311.28; 293.24;	220.00; 260.00; 300.00
Biochanin A	3.17	283.28	268.06; 265.20; 240.29	220.00; 260.00;
Exiguaflavanone K	9.06	369.33	339.12; 325.18; 293.23 284.17;	220.00; 265.00; 290.00; 375.00
Lonchocarpol A	10.47	407.33	398.22; 301.11; 313.21; 287.24; 261.22; 243.18	220.00; 295.00; 345.00
Hiravanone	10.77	437.43	422.15; 313.20; 301.18; 287.11; 261.25; 243.23	220.00; 295.00; 345.00 (220; 290; 335; 375; 425)
Amorfrutin A	11.77	339.28	295.15; 225.19	225.00; 270.00; 305.00; 375.00; 425.00
Amorfrutin 2	12.56	305.24	290.24; 275.25; 261.18; 247.22	220.00; 270.00; 305.00; 375.00; 420.00
Amorfrutin C	16.00	407.33	363.30; 348.25; 331.27; 305.12; 285.29	220.00; 315.00; 375.00; 410.00

Thus, along the previously described amorfrutins, ten additional polyphenols and the triterpene oleanolic acid (**29**) (Errichiello et al., 2023), were also isolated from extract fractionation. In particular, six flavanone derivatives were obtained and identified as sakuranetin (**18**) (Vasconcelos et al., 1998), sophoraflavanone B (**19**) (Chen et al., 2013), exiguaflavanone K (**20**) (Iinuma et al., 1994), the unprecedented exiguaflavone L (**21**), lonchocarpol A (**22**) (Roussis et al., 1987), and shiravanone (**23**) (Ito et al., 1989). In addition, the isoflavonoid biochanin A (**24**) (Chang et al., 1994), the chalcone phlorizin (**25**) (Xu et al., 2013), the acetophenone myrciaphenone A (**26**) (Gurbuz et al., 2018), 6-dihydroxyphenyl-1-butanone-2- β -D-O-glucopyranoside (**27**) (Ko et al., 2009), and benzyl- β -D-glucopyranoside (**28**) (Seigler et al., 2002) were also obtained.

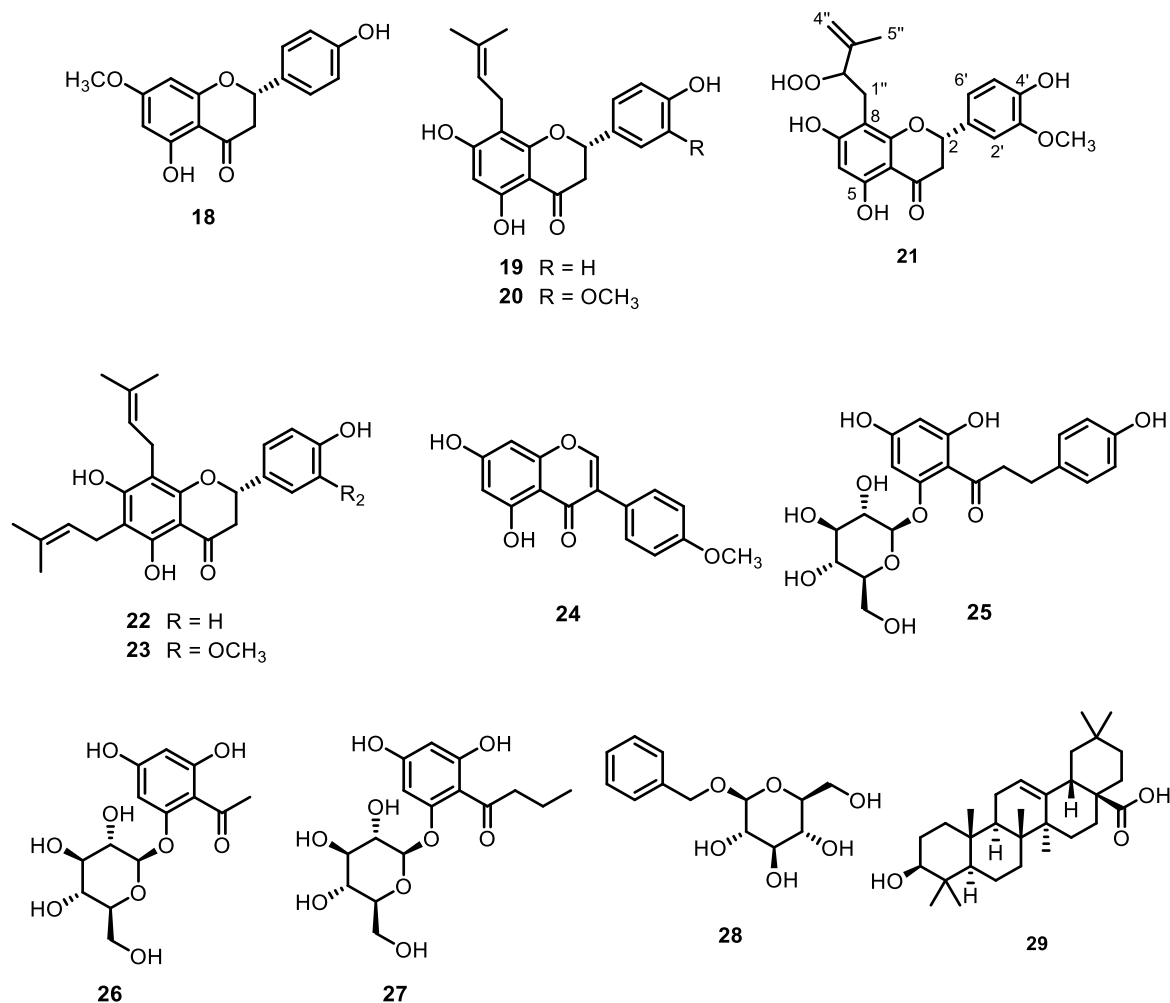


Figure 1.12 Chalcones, flavonoids and other metabolites isolated from *G. foetida*

Exiguaflavone L (**21**) was obtained as a pale yellow amorphous solid with the molecular formula C₂₁H₂₂O₈, as deduced by the positive [M+Na]⁺ HR-ESIMS peak at *m/z* 425.1211 (calculated for C₂₁H₂₂O₈Na 425.1212). From the combined analysis of the ¹H NMR and the COSY spectrum of **21**, the typical pattern of a prenylated flavanone of the exiguaflavone family were revealed, including the ring A singlet at δ_H 6.07, the CH₂-CH spin system of ring B (δ_H 5.32, 3.09 and 2.79), and three almost completely overlapped methine signals around δ_H 6.93. The prenylated side chain included a -CH₂CH- moiety (δ_H 3.08, 2.79, 4.48), a *sp*² methylene at δ_H 5.07, and a methyl singlet at δ_H 1.87. All the proton signals were associated with those of the directly linked carbon atoms thanks to the correlations of the HSQC spectrum. The carbon skeleton of exiguaflavone L (**21**) was built up with the aid of the HMBC spectrum, with the crucial correlations of H-1 with C-1', C-2', and C-6', of H-6 with C-5, C-10, C-7, and C-8, and of H-1'' with C-3, C-7, C-8, and C-9. The placement of the methoxy group was secured by the HMBC cross-peak of the singlet at δ_H 3.93 with C-3'. Finally, a hydroperoxy group, in agreement

with the molecular formula, was placed at C-2'' on the basis of the marked downfield shift of the carbon atom (δ_c 88.8). The minute amounts isolated prevented determination of the absolute configuration at C-2''.

1.4 Bioactivity of *G. foetida* secondary metabolites

From our phytochemical investigation, *G. foetida* appeared to be a rich source of bioactive ingredients. Polyphenols are currently in the spotlight for their range of benefits on metabolic syndrome (Santa et al., 2023), and we aimed to examine how *G. foetida* secondary metabolites can modulate some targets involved in this pathology. Firstly, we have decided to direct our efforts in evaluate the amorfrutins agonism on PPAR α and PPAR γ nuclear receptors. Moreover, given the library of obtained amorfrutins, we were able to deepen in their structure-activity relationships and conducting in silico analysis to study how they interact with the targets. Secondly, we have chosen to test the compounds effect on dysfunctional pathways commonly involved in MetS, such as mitochondrial activity and glucose uptake.

1.4.1 Amorfrutins modulation on PPARs target

Relying upon the promising data available in the literature on the amorfrutin potential as PPARs agonist, we have tested the obtained library of natural amorfrutins on PPAR α and PPAR γ nuclear receptors. Before testing the pure compounds, their cytotoxicity was tested in COS-7 cells, a common fibroblast-like clonal cell line, using the MTT assay, thus treating the cells with increasing concentrations of each compound (1, 10, and 30 μ M) for 24 h. As shown in **Figure 1.13**, decarboxyamorfrutin A (**2**), its C-prenyl derivative (**10**), and amorfrutin 2 (**12**) showed a significant cytotoxic effect. Interestingly, in some cases (e.g., **1** vs **2** or **10** vs **11**), the cytotoxic effect appeared to be associated with the lack of the carboxylic group. Based on these results, all compounds showing a cytotoxic effect were excluded from the following experiments, as well as amorfrutin N (**5**) and amorfrutin G (**7**) that could not be tested due to the an insufficient amount isolated. COS-7 cells were then used for the luciferase assays. For these assay, subconfluent cells (60-70%) were transiently transfected with distinct plasmids corresponding to PPAR α or PPAR γ (ligand binding domain, LBD)/GAL4 (dimerization binding domain, DBD), the UAS enhancer MH100, and Renilla. After transfection, cells were then treated with test compounds at a concentration of 10 μ M. The selective agonists GW7647 (0.1 μ M, for PPAR α)

and rosiglitazone (0.1 μ M, for PPAR γ) were used as positive controls. Amorfrutin A (**1**), amorfrutin 3 (**3**), amorfrutin E (**4**), amorfrutin H (**8**), amorfrutin J (**14**), amorfrutin K (**15**), and amorfrutin M (**17**) significantly promoted PPAR α activity (between 5- and 10-fold) compared to vehicle-treated cells (**Figure 1.14**, panel A shows that). Notably, some of these compounds (**1**, **3**, **4**, **14**) proved to be active also on PPAR γ , with amorfrutin A showing the highest effect (fold >20) and with amorfrutin C (**11**) being active only on PPAR γ (**Figure 1.14**, panel B).

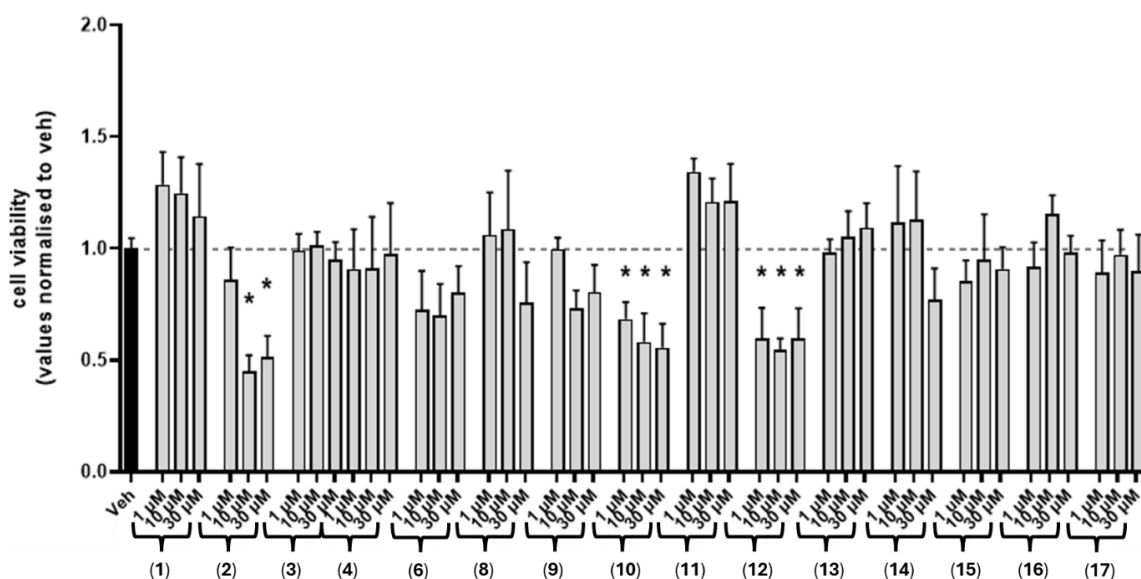


Figure 1.13 Effect of natural and synthetic amorfrutin derivatives on COS7 cell viability. The day after plating when cells reached a confluency of about 70% were treated with indicated concentrations of the tested amorfrutins for 24 h (the absence of serum was maintained during the treatments). Cell viability was assessed using the MTT assay. Each bar is the mean $\pm$ S.E.M. from 3-4 independent biological samples. **** $P \leq 0.0001$; ** $P \leq 0.003$; * $P \leq 0.05$ vs. the indicated experimental group calculated using ANOVA using GraphPad Prism 7 software.

To evaluate more in-depth the dose-response relationship, transfected COS-7 cells were treated with increasing concentrations of the active compounds (0.1 to 30 μ M). A linear concentration-response curve (**Figure 1.14**, panel C) for PPAR α activation was observed for all compounds except amorfrutin H (**8**), which produced a bell-shaped response, most likely due to the cytotoxic effect observed at the highest concentration (30 μ M) (**Figure 1.14**). All compounds, except **1**, also showed a linear concentration-response curve for the activation of PPAR γ , plateauing at the two highest concentrations assayed (10 and 30 μ M, **Figure 1.14**, panel D). Taken together, these results show that amorfrutin M (**17**), amorfrutin 3 (**3**), amorfrutin J (**14**), and amorfrutin E (**4**) behave as dual PPAR α/γ agonists, amorfrutins C (**11**) and A (**1**) are more potent as PPAR γ agonists, while amorfrutin H (**8**) is a selective PPAR α agonist. The molecular bases for these interactions were explored with *in silico* strategy.

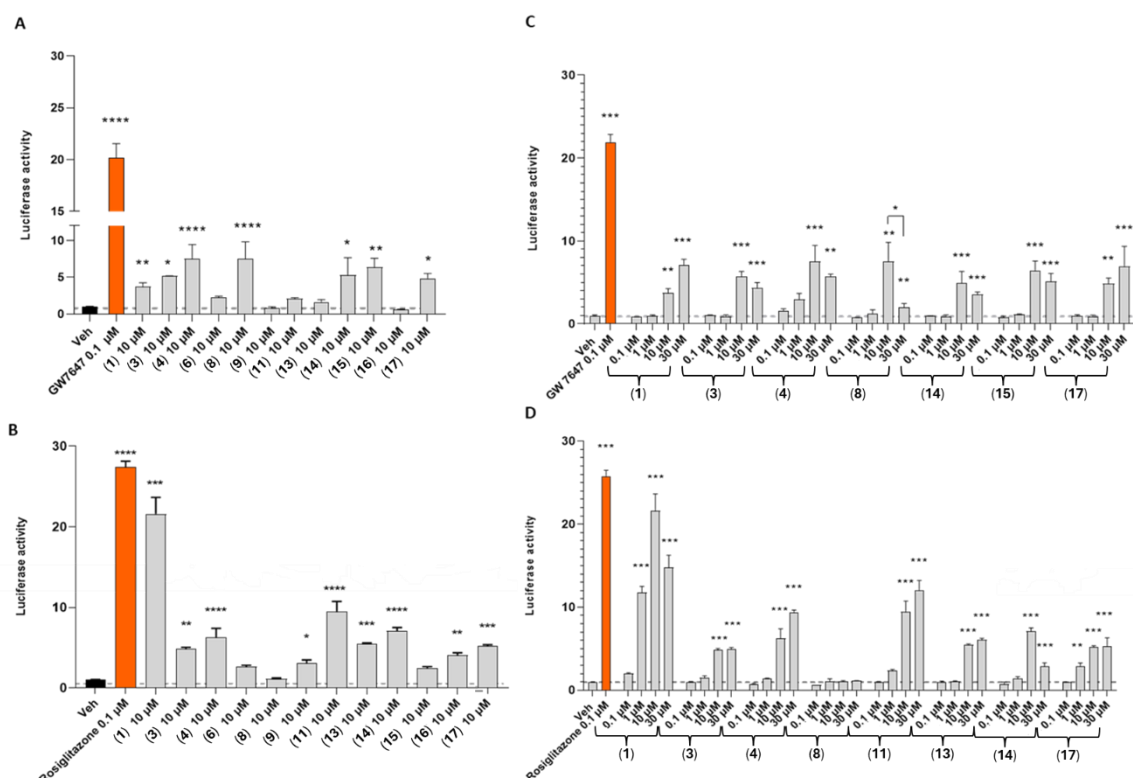


Figure 1.14 Effect of natural amorfrutin derivatives on PPAR α and/or PPAR γ activity in COS-7 cells. Relative Luciferase Units in response to 10 μ M of test compounds on PPAR α (A) and PPAR γ (B). Concentration-response effect of test compounds on PPAR α (C) and PPAR γ (D). The activity of the vehicle control was set at 1 and the relative luciferase activities obtained for each compound and dose tested are presented as fold induction to the vehicle control. Each bar is the mean $\pm$ S.E.M. from 3-4 independent biological samples. **** $P \leq 0.0001$; *** $P \leq 0.0003$; ** $P \leq 0.003$; * $P \leq 0.05$ vs. the indicated experimental group calculated using ANOVA or unpaired Student's t-test using GraphPad Prism 7 software.

1.4.2 In Silico Study and Structure-Activity Relationships on PPARs

To provide insights on the structure-activity relationships of this class of compounds, the PPAR α/γ binding mode of the most active amorfrutins was obtained by a combined approach of molecular docking and molecular dynamics (MD) simulations. For each compound, the most relevant binding poses from docking, selected on the basis of binding energy, cluster population, and visual inspection, underwent 100 ns of MD to assess their stability. If rearrangements occurred during MD, the stability of the new pose was evaluated for an additional 50-100 ns. Only docking poses remaining stable for at least 60 ns during MD were considered for structure-activity relationships (SARs) (see Appendix D). Their main overall features are described in this section, while interactions characterizing each pose are given in **Table 1.3** and **Table 1.4**.

Table 1.3 Pattern of relevant H-bonds between amorfrutins and PPAR α residues. Secondary structure elements involving each residue are also reported. Ligand chirality is shown in parenthesis for interactions observed only for a single stereoisomer.

Compound	Cys276 (H3)	Gln277 (H3)	Ser280 (H3)	Tyr314 (H4/5)	His440 (H10/11)	Tyr464 (H12)
1 (pose A)				x	x	x
1 (pose B)			x	x	x	x
3			x	x	x	x
4 (pose up)	X	x	x(S)	x(S)	x(S)	x(S)
4 (pose down)		x	x	x	x	x
8		x	x	x	x	
14		x(R)	x	x	x	x
15			x	x	x	x
17			x			

Table 1.4 Pattern of relevant H-bonds between amorfrutins and PPAR γ residues. Secondary structure elements involving each residue are also reported.

Compound	His266 Ω -loop	Arg288 H3	Ser289 H3	Ser342 β -sheet	His323 H4/5	His 449 H10/11	Tyr473 H12
3				x			
4		x		x			
8	X			x			
11		x					
13			x				
14			x		x	x	x
17			x				

The crystal structure of amorfrutin A (**1**) complexed to PPAR γ has been reported (de Groot et al., 2013); conversely, no experimental data are available for its interaction with PPAR α , where it is also active, although with lower potency than on PPAR γ . Thus, its binding mode was investigated in silico, and two alternative binding poses were found (pose 1A and 1B), characterized by a reversed orientation of the phenylethyl moiety (**Figure 1.15**)

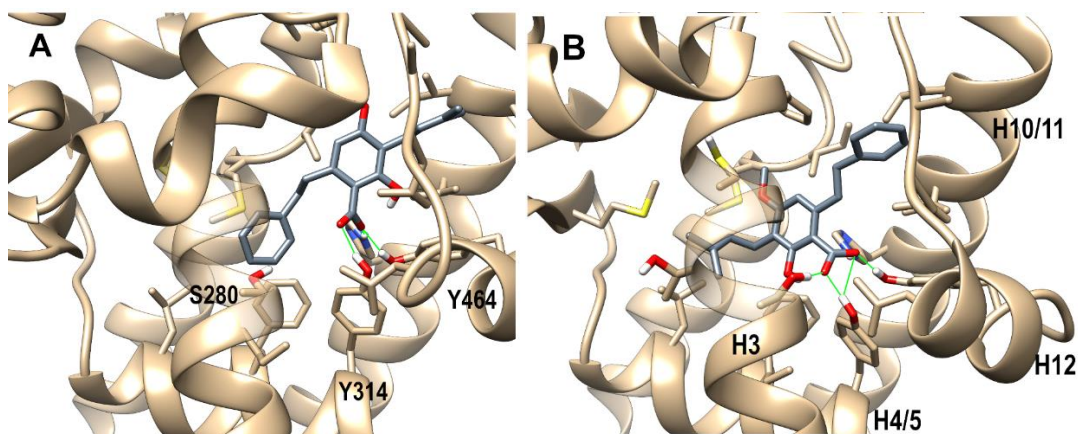


Figure 1.15 Representative MD frames of PPAR α LBD (tan) in complex with amorfrutin A (**1**, slate gray). Poses A and B discussed in the text are shown in panel A and B, respectively. Ligand and protein sidechains within 4 Å from the ligand are shown in stick representation. Oxygen, nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively. H-bonds are shown as green sticks.

In both poses, the carboxylate group forms a network of Hbonds involving residues lying in the polar cleft, herein after referred to as “canonical network” (see **Table 1.6**), while in pose B an additional H-bond with the Ser280 (H3) side chain is observed. The network of H-bonds engaged by the ligand within the LBD is compatible with full agonism, but the pattern of interaction confines the ligand in a polar subpocket, with an overall limited occupancy of the LBD. Amorfrutin C (**11**) was the second most active PPAR γ agonist with negligible activity on PPAR α . Its stable binding modes at PPAR γ are shown in **Figure 1.16**. In analogy to amorfrutin A (**1**) interaction, the binding site encompasses helix H3, the β -sheet, and the Ω -loop. Two main orientations are adopted by the phenethyl chain, which points toward Phe264 on the Ω -loop in orientation A and in the opposite direction in orientation B. Amorfrutin C (**11**) differs from amorfrutin A (**1**) only for an additional isoprene group decorating the resorcinol ring. This additional substituent exerts a detrimental effect on the activity at both receptors. At PPAR α , it induces steric hindrance preventing the binding at the hydrophilic subpocket, which hosts instead amorfrutin A (**1**). At PPAR γ , it increases both the entropic cost upon binding, due to the flexible additional pendant, and the number of suboptimal binding modes in terms of receptor activation.

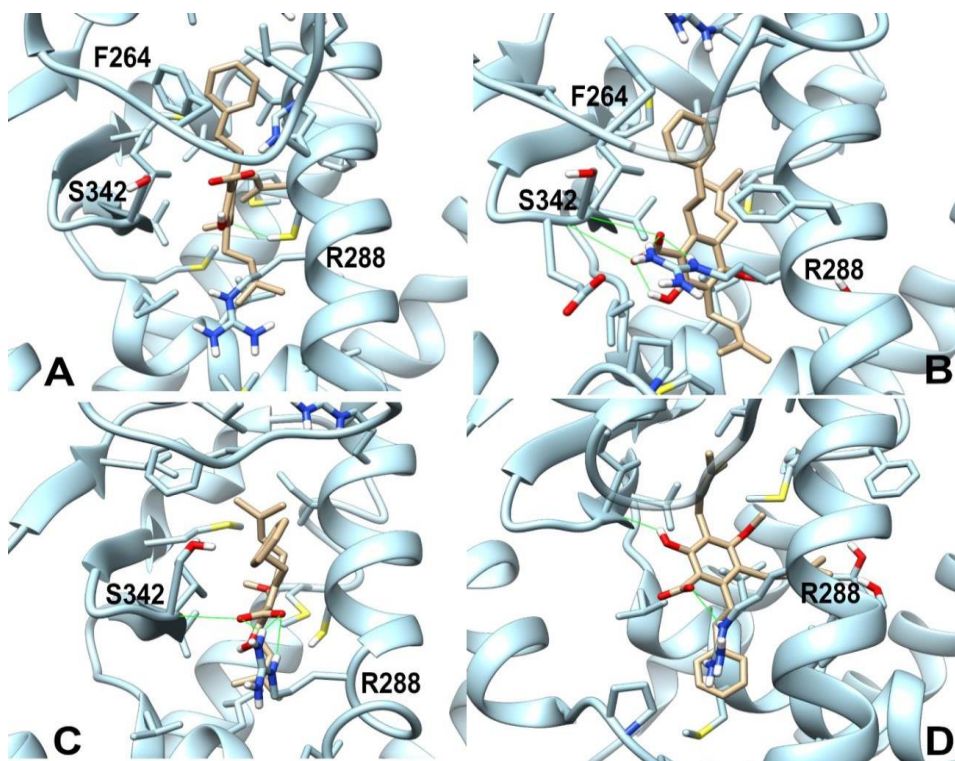


Figure 1.16 Representative MD frames of PPAR γ LBD (light blue) in complex with amorfrutin C (**10**, tan). Poses A1, A2 and A3 discussed in the text are shown in panel A-C, pose B in panel D. Ligand and protein sidechains within 4 Å from the ligand are shown in stick representation. Oxygen, nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively. H-bonds are shown as green sticks.

The racemic amorfrutin J (**14**) behaves as a dual PPAR α/γ agonist. The binding modes of the two enantiomers, hosted in the polar subpocket which comprises helix H12, are similar in the two receptor isoforms. In PPAR α/γ LBDs, the carboxylate group forms the canonical network of hydrogen bonds (see **Table 1.6** and **Table 1.7**). However, the ligand adopts a flipped orientation for PPAR α , its pentyl chain now pointing toward helix H10/11. This reversed orientation is probably due to the occurrence of Phe363 (H7) in place of PPAR α Ile354, which changes the accessible surface size/shape of this hydrophobic cleft.

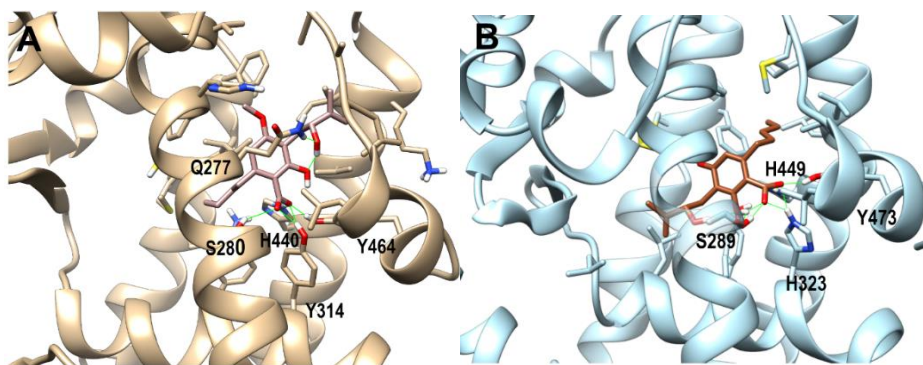


Figure 1.17 Panel A: representative MD frame of PPAR α LBD (tan) in complex with the R enantiomer of amorfrutin J (**14**, rose brown). Panel B: representative MD frame of PPAR γ LBD (light blue) in complex with the S enantiomer of amorfrutin J (**14**, sienna). Ligand and protein sidechains within 4 Å from the ligand are shown in stick representation. Oxygen, nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively. H-bonds are shown as green sticks.

The situation of amorfrutin 3 (**3**), again a dual PPAR α/γ agonist, is quite similar to that of **14**, since the two compounds only differ in the side chains. The binding poses of the two enantiomers of **3** at the PPAR α are similar to those of **14**. Conversely, the preferred binding site at PPAR γ is located between the β -sheet and helix H3, recapping the binding mode of the other amorfrutins with the same side chain (**Figure 1.18**).

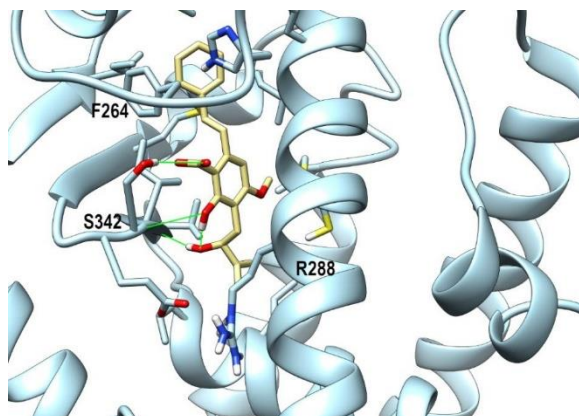


Figure 1. 18 Representative MD frame of PPAR γ LBD (light blue) in complex with the S enantiomer of compound **3** (khaki). Ligand and protein sidechains within 4 Å from the ligand are shown in stick representation. Oxygen, nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively. H-bonds are shown as green sticks. The binding poses of both enantiomers of **3** are stabilized by π - π stacking between the phenyl group and Phe264 on Ω -loop and by H-bonds engaged by the carboxylate group with Ser342 sidechain on the β -sheet. The S enantiomer forms an additional H-bond involving the hydroxyl group of the stereogenic center with the backbone atoms of Leu340 and Ser342.

The racemic amorfrutin E (**4**) shows a dual agonist profile. Both enantiomers can adopt two binding poses within the LBD of PPAR α , depending on the orientation of the phenyl group, which can alternatively point toward the helix H10/11 (up pose) or the β -sheet (down pose). In the up pose, the carboxylate group of the *S* enantiomer is involved in the canonical H-bonds listed in **Table 1.6** and the hydroperoxide group is involved in a H-bond with the Ser280 (H3) side chain. In this pose, the *R* enantiomer moves away from helix H12 due to the formation of a H-bond between the peroxide group and Cys276 (H3) (data not shown). In the down pose, both enantiomers engage the same pattern of canonical H-bonds, but the peroxide group is now involved in a H-bond with the Gln277 (H3) side chain (**Figure 1.19**).

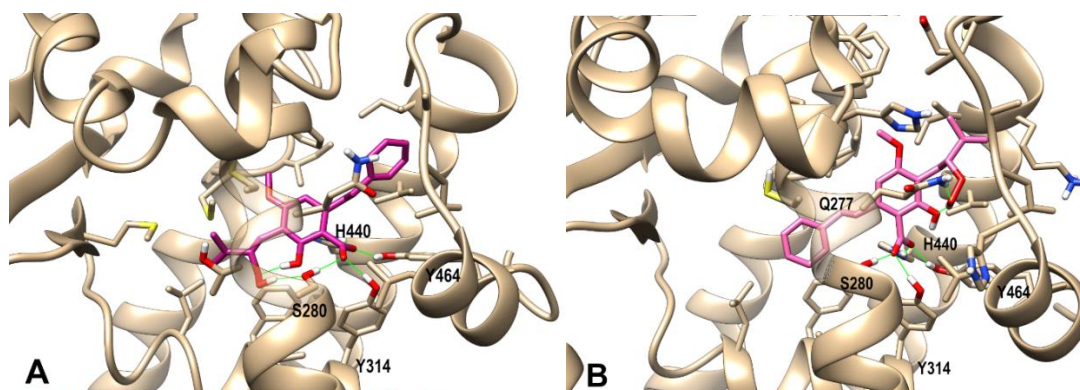


Figure 1.19 Representative MD frames for the up pose (panel A) and down pose (panel B) of PPAR α LBD (tan) in complex with the *S* (magenta, panel A) and *R* (pink, panel B) enantiomers of amorfrutin E (**4**). Ligand and protein sidechains within 4 Å from the ligand are shown in stick representation. Oxygen, nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively. H-bonds are shown as green sticks.

In the LBD of PPAR γ , both enantiomers are hosted in the cleft between the β -sheet and helix H3, where they form π - π stacking between their phenyl ring and Phe264 on the Ω -loop. Alternative poses have been obtained for both enantiomers (see **Figure 1.20**).

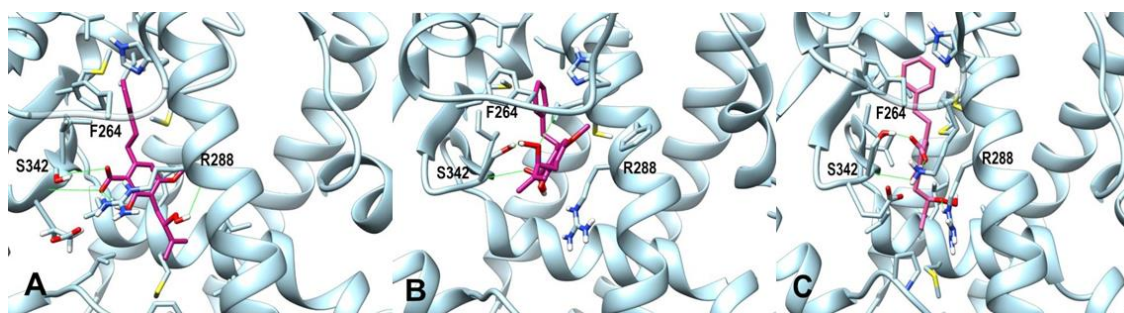


Figure 1. 20 Representative MD frames of PPAR γ LBD (tan) in complex with the S (violet red, panels A and B) and R (pink, panel C) enantiomers of amorfrutin E (**4**). Ligand and protein sidechains within 4 Å from the ligand are shown in stick representation. Oxygen, nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively. H-bonds are shown as green sticks. Alternative poses have been obtained for both enantiomers. The carboxylate group is involved in H-bonds with backbone atoms and/or Ser342 sidechain of the β -sheet and/or in H-bond/ionic interactions with Arg288 (H3). The peroxide group of the R enantiomer can alternatively form H-bonds with Arg288 (H3) or Cys285 (H3) sidechains, while that of the S enantiomer is H-bonded to either Ser342 sidechain and Ω -loop backbone atoms, or backbone atoms of helix H3.

While devoid of PPAR γ agonist activity, amorfrutin H (**8**) was, along with amorfrutin E (**4**), the most potent PPAR α agonist. To rationalize this behavior, a computational study was undertaken for both receptor isoforms. At PPAR α , amorfrutin H (**8**) binds in the polar subpocket of the LBD encompassing helix H12, where it engages the canonical network of stable H-bonds with its carboxylate group. The phenyl group is hosted in a hydrophobic cleft formed by Phe273 (H3), Leu344 (H6), Ile354, and Met355 (H7). The representative MD frame for this complex is shown in **Figure 1.21**, panel A. A stable complex was also obtained for PPAR γ (**Figure 1.21**, panel B), where amorfrutin H binds near the β -sheet and is involved in both H-bonds with Ser342 (β -sheet) and His266 (Ω -loop) and π - π stacking with Phe264 and His266 (Ω -loop). This stable binding mode prompted us to evaluate by luciferase assay a possible activity as a PPAR γ antagonist for this compound. As shown in the graphic of **Figure 1.21**, amorfrutin H is indeed able to significantly antagonize the effect of rosiglitazone.

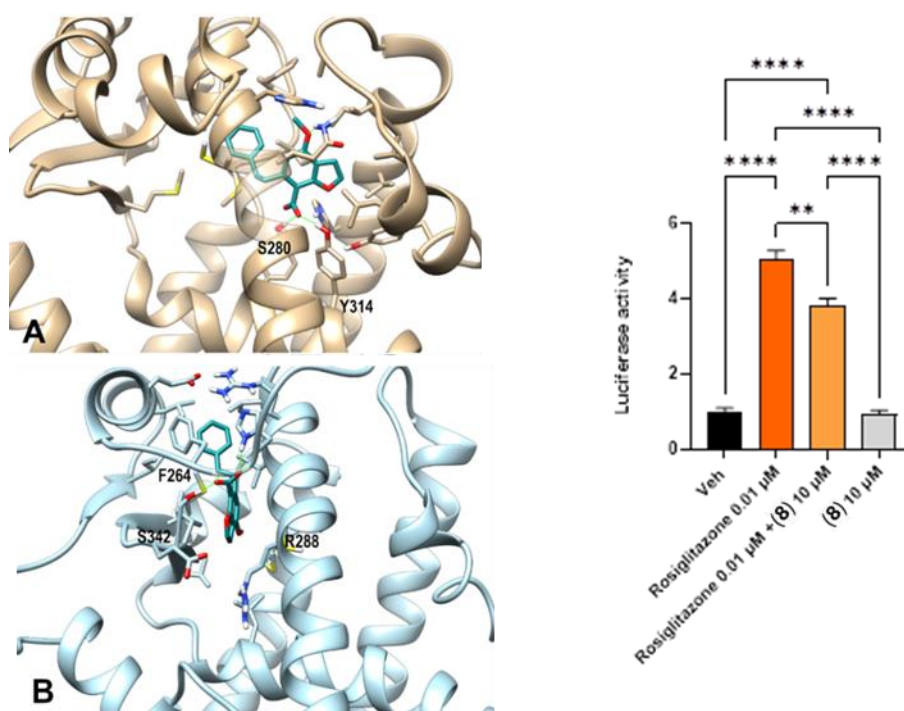


Figure 1. 21 Panel A: representative MD frames of PPARα (tan, panel A) and PPARγ LBD (light blue, panel B) in complex with amorfrutin H (**8**) (dark cyan). Ligand and protein sidechains within 4 Å from the ligand are shown in stick representation. Oxygen, nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively. H-bonds are shown as green sticks. Panel C: Effect of amorfrutin H (**8**) on PPARγ activity induced by rosiglitazone 0.01 μM. The activity of the vehicle control was set at 1 and the relative luciferase activities obtained for each compound and dose tested are presented as fold induction to the vehicle control. Each bar is the mean ± S.E.M. from 3-4 independent biological samples. **** $P \leq 0.0001$; ** $P \leq 0.003$ vs. the indicated experimental group calculated using ANOVA using GraphPad Prism 7 software.

The racemic amorfrutin K (**15**) showed stronger activity toward PPARα, being a weak PPARγ agonist. At PPARα, the enantiomers adopt a similar binding pose and engage H-bonds between the carboxylate group and residues lying in the polar cleft encompassing helix H12 (see **Figure 1.22**), but only the *R* enantiomer forms an additional H-bond between its hydroxy group and Ser280 (H3), as shown in **Figure 1.22**. Intriguingly, **15** shares a similar structure of the isoprenoid region compared to that of amorfrutin H (**8**), thus suggesting a detrimental effect of this moiety on PPARγ agonism.

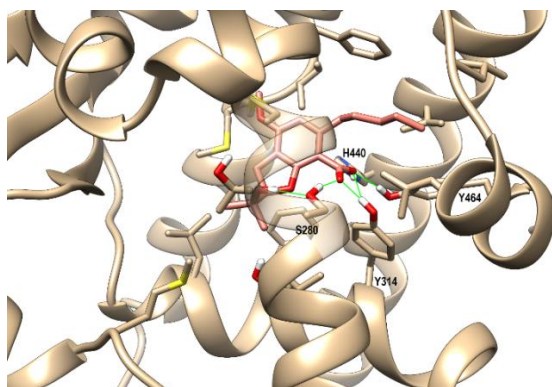


Figure 1.22 Representative MD frame of PPAR α (tan) LBD in complex with the R enantiomer of amorfrutin K (**15**, coral). Ligand and protein sidechains within 4 Å from the ligand are shown in stick representation. Oxygen, nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively. H-bonds are shown as green sticks.

The binding mode of amorfrutin M (**17**) (see **Figure 1.23**, panel A) is similar at both receptors, with the two isoprene arms adopting a horseshoe arrangement around helix H3, and the resorcinol moiety engaged in a H-bond with Ser280(PPAR α)/289 (PPAR γ), accounting for its activity as a dual PPAR α/γ agonist. The binding mode of compound **13**, a selective PPAR γ agonist, is shown in **Figure 1.23**, panel B. As for compound **17**, the resorcinol moiety is involved in an H-bond with Ser289 (H3). However, in this case, the lack of one isoprene substituent allows π - π stacking between the resorcinol moiety and Phe363 (H7). This stabilizing interaction is missing in PPAR α , where a residue of isoleucine occurs in place of Phe363 (H7), and this could explain the lack of activity at this receptor.

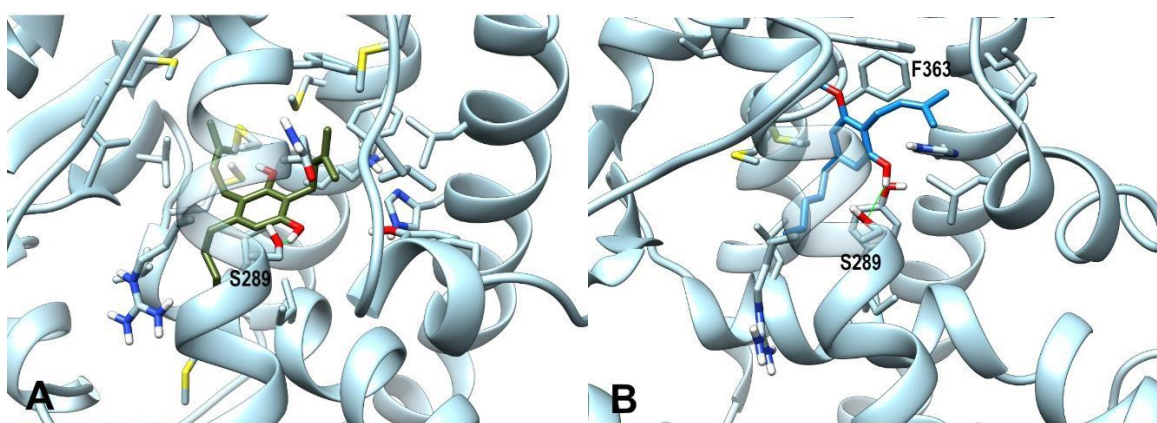


Figure 1.23 Representative MD frames of PPAR γ (light blue) LBD in complex with compound **17** (dark green, panel A) and **13** (dodger blue, panel B). Ligand and protein sidechains within 4 Å from the ligand are shown in stick representation. Oxygen, nitrogen, and sulfur atoms are colored red, blue, and yellow, respectively. H-bonds are shown as green sticks.

1.4.3 Effect of mitochondrial activity

In the first pipeline, we investigated the effect of *G. foetida* metabolites on the mitochondrial activity of C2C12 myoblasts grown in vitro. Mitochondrial activity was measured using the selective probe MitoTracker CMXRos (de Brito Monteiro et al., 2020). The probe accumulates in the mitochondrial intermembrane space and emits fluorescence with an intensity well correlated to cellular mitochondrial activity. C2C12 cells were cultured in the presence of *G. foetida* pure metabolites. As shown in **Figure 1.24**, most of the amorfrutins tested (**1–3, 6, 8, 9–17**) boosted mitochondrial activity in C2C12, although with different potencies. Amorfrutin M (**17**) and decarboxyamorfrutin A (**2**) clearly presented the highest potencies. Sophoraflavanone B (**19**) also showed a good activity. Confocal microscopy confirmed mitochondrial stimulation of the amorfrutin-containing fraction (AF) obtained from the total extract (E2 and E3) (**Figure 1.25**).

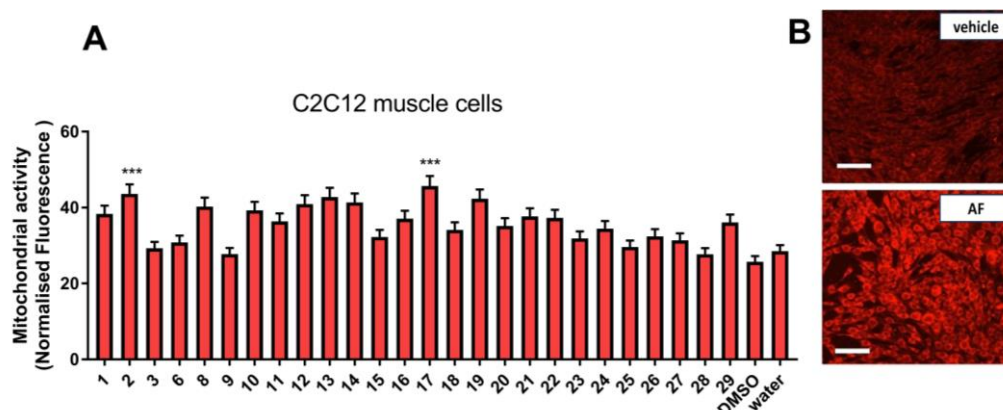


Figure 1. 24 Effect of *G. foetida* pure metabolites on mitochondrial metabolism in C2C12 cells. A) Normalized Mito-tracker Red CMX-ROS fluorescence measured upon incubation of C2C12 cells with the indicated *G. foetida* metabolites (3 μ M) for 48 hours, (values are reported as mean $\pm$ SEM and are representative of three independent experiments; *** $p < 0.001$); B) Confocal microscopy of C2C12 cells incubated with pooled amorfrutins (Amorfrutin Fraction (AF), 10 mg/L) or with an equivalent amount of vehicle for 48 hours and stained with Mito-tracker Red CMX-ROS. Magnification bar = 100 μ m

To verify the mitochondrial-stimulating activity of amorfrutins on a different cell line, we tested *G. foetida* pure metabolites on human hepatoma HuH7 cells (**Figure 1.25**). Similar to myoblasts, amorfrutins also exerted mitochondrial stimulatory activity on the hepatoma cell line. Amorfrutin A (**1**), its decarboxylated analogue **2**, amorfrutin C (**11**), its decarboxylated and demethoxylated analogue **10**, and amorfrutin M (**17**), presented the highest potencies.

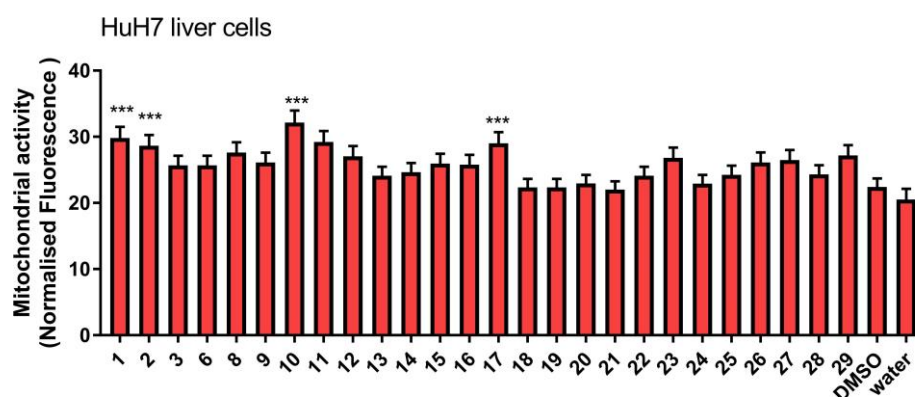


Figure 1. 25 Effect of *G. foetida* pure metabolites on mitochondrial metabolism in HuH7 liver cells. Normalized Mito-tracker Red CMX-ROS fluorescence measured upon incubation of HuH7 cells with the indicated *G. foetida* metabolites (3 μ M) for 48 hours. (values are reported as mean $\pm$ SEM and are representative of three independent experiments; *** $p < 0.001$).

Investigate at the transcriptional level the effect of *G. foetida* metabolites on mitochondrial metabolism, mRNA was extracted from HuH7 cells treated with the amorfrutin-rich fraction AF. Analyzed by qPCR, mRNA levels of PGC1- α (PPAR γ coactivator 1alpha), a key player in lipid metabolism and mitochondrial activity (Esterbauer et al., 1999), were upregulated upon incubation with the AF (**Figure 1.26**). On the contrary, the enzyme involved in cholesterol conversion into bile acids, CYP7A1, was downregulated in treated cells, suggesting that anabolic reactions involving cholesterol were hampered in cells treated with amorfrutins. A reduction in lipid content (fatty acids and cholesterol) was confirmed by GC-MS. Lipids were extracted from HuH7 cells treated for 48 h with the AF. GC-MS analysis revealed decreased intracellular levels of cholesterol, palmitic acid, and myristic acid, probably as consequence of augmented β -oxidation promoted by PPAR γ .

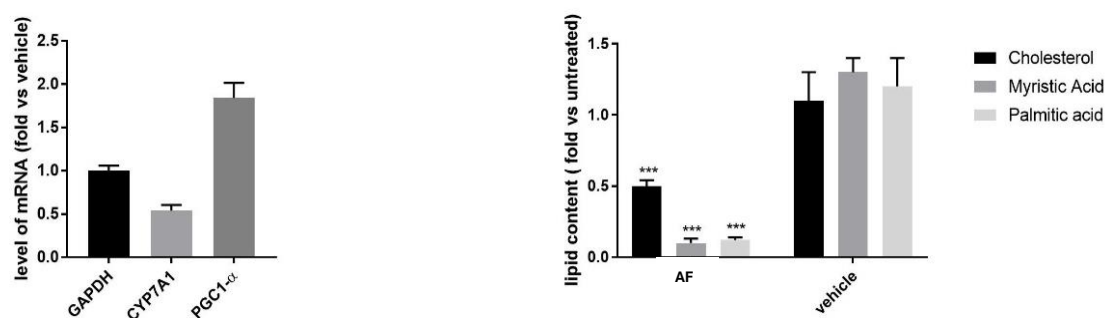


Figure 1.26 Left panel) Intracellular level of PGC-1 α and CYP7A1 mRNA in HuH7 cells treated with *G. foetida* amorfrutin enriched fraction AF (10 mg/L) for 48 hours (values are reported as mean $\pm$ s.d. and are representative of three independent experiments); Right panel) Intracellular levels of cholesterol, palmitic acid and myristic acid in HuH7 cells treated with *G. foetida* fraction AF (10 mg/L) for 48 hours (value are reported as mean $\pm$ s.d. and are representative of three GC-MS independent experiments).

1.4.4 Promotion of glucose uptake via glucose transporters (GLUTs)

In a second pipeline, *G. foetida* metabolites were tested for their capacity to promote glucose uptake via glucose transporters (GLUTs). Dysmetabolic disorders, which are frequently characterized by insulin resistance and hyperglycemia, have impaired control over the handling of circulating glucose levels. GLUTs are essential for removing glucose from the bloodstream, and many of them are controlled by insulin receptors and react to insulin stimulation via PI3K signaling. Insulin stimulation can increase the number of GLUT transporters (mostly GLUT4) on the plasma membrane of muscle cells, boost glycolysis, and ultimately enhance the activity of passive transporters (GLUT1). The effect of *G. foetida* metabolites on glucose uptake was measured by monitoring the uptake of NBDG [2-(7-nitro-2,1,3-benzoxadiazol-4-yl)-D-glucosamine], a fluorescent derivative of deoxyglucose covalently bound to the fluorescent probe nitro blue tetrazolium (NBT) (Bala et al., 2021). An intracellular increase in fluorescence upon NBDG uptake indicates activation of the GLUT transporters. Most of the amorfrutins exhibited glucose uptake stimulatory activity, with amorfrutin 2 (**12**) presenting the highest potency (**Figure 1.27**).

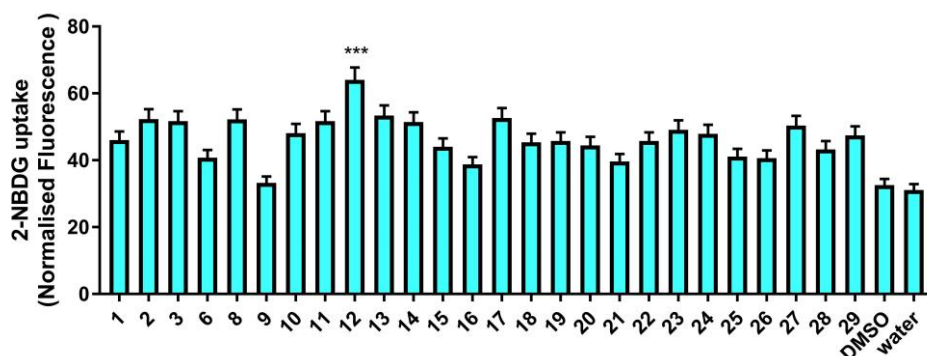


Figure 1. 27 Normalized 2-NBDG fluorescence measured upon incubation of C2C12 cells with the indicated *G. foetida* metabolites (3 μ M) for 48 hours; the dotted red line indicates the basal glucose uptake activity of the cells (values are reported as mean $\pm$ SEM and are representative of three independent experiments; *** $p < 0.001$).

1.5 Conclusions

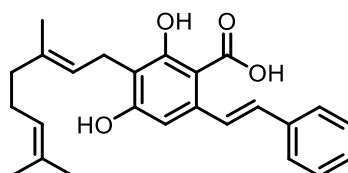
Our detailed investigation of the aerial parts of *G. foetida* led to the isolation of 29 pure compounds mainly belonging to the amorfrutin class and to other polyphenol classes, including 11 previously undescribed amorfrutins and an unprecedented prenylated flavanone. The biosynthetic profligacy of *G. foetida* provided a set of amorfrutins that included many compounds with an oxygenated or truncated isoprenyl moiety, an element of the biological space of amorfrutins still substantially unexplored. In some cases, oxidation was associated with the formation of stereogenic centers, and all the chiral compounds were obtained in racemic form. All amorfrutins lacking cytotoxicity were assayed for agonist activity on the PPAR α and PPAR γ nuclear receptors and computational studies shed light on the binding of amorfrutins to PPAR α/γ receptors, rationalizing the strict structure-activity relationships that characterize these interactions. Afterwards, all the isolated compounds were investigated for modulation of mitochondrial activity and stimulation of glucose uptake via GLUT transporters, two endpoints related to glucose homeostasis and, thus, to metabolic syndrome.

In conclusion, amorfrutins seem to be the most promising class of compounds, as supported by the following evidences:

- 1) Amorfrutin A (**1**) proved to be the most potent agonist of PPAR γ , followed by amorfrutin C (**11**).

- 2) Amorfrutin H (**8**), the single compound including a benzofuran moiety, selectively activated PPAR α and significantly antagonized the effect of rosiglitazone on PPAR γ .
- 3) Amorfrutins E (**4**), 3 (**3**), and J (**14**) behaved as dual agonists. Interestingly, they practically share the same 2'-oxygenated prenyl moiety, regardless the pentyl or phenethyl chains.
- 4) Amorfrutin M (**17**) and decarboxyamorfrutin A (**2**) emerged as mitochondrial stimulators. Moreover, **17** displayed an interesting activity also as on PPAR α/γ agonist.
- 5) Amorfrutin 2 (**12**) emerged as a glucose uptake promoter.

Thus, amorfrutins represent promising new leads for the development of a new class of selective PPARs modulators. Indeed, their peculiar activation of PPAR γ , resulting in the absence of side effects showed by the glitazones, could pave the way to solve one of the medicinal chemistry challenges in the pharmacological treatment of type 2 diabetes mellitus. Furthermore, the co-activation of PPAR α/γ demonstrated by the dual agonists allows the simultaneous modulation of hyperglycaemia and hypertriglycaemia, typical co-morbidities of patients suffering from type 2 diabetes mellitus or obesity. In fact, certain amorfrutin analogs sharing the same pharmacological profile are currently in preclinical evaluation. In particular, amorphastibol (**Figure 1.28**), isolated from *Amorpha fruticosa*, has been shown to improve glucose and lipid dysregulation in vivo without significant adverse effects such as weight gain or hepatomegaly (Lee et al., 2013; Balakumar et al., 2019). Long-term studies are needed to assess the safety profile of this class of compounds to support their clinical application (Lavecchia et al., 2015), but the preliminary results are very encouraging and suggest that it is worthwhile to pursue further investigation of amorfrutins.



Amorphastibol

Figure 1.28 Chemical structure of amorphastibol

However, the rich polyphenol components of this extract clearly contribute to the beneficial effect on the plant on MetS, with sophoraflavanone B (**19**) and 6-dihydroxyphenyl-1-butanone-2- β -D-O-glucopyranoside (**27**) efficacy in mitochondrial and GLUT stimulation, respectively. Moreover, it is worth mentioning that the extract also contains the glycosylated dihydrochalcone phlorizin (**25**), a sodium-glucose transport (SGLT) protein inhibitor in the nephron with a significant effect on inhibiting reabsorption of sugar in the kidney and a consequent lowering effect on plasma glucose (Ehrenkranz et al., 2005). The activity of phlorizin served as inspiration for the development of the synthetic class of SGLT2 inhibitors collectively named glifozins (Blaschek et al., 2017).

In conclusion, since the *G. foetida* total extract contains several biomolecules acting on different targets potentially involved in the management of metabolic syndrome symptoms, its traditional use is fully supported by our investigation, which, hopefully, will inspire further analysis to support a commercial use of this herbal product helpful in MetS management.

1.6 Experimental Section

1.6.1 Plant material and extraction

G. foetida aerial parts were collected in Fernana (northwestern Tunisia) in July 2019. The taxonomic identification was carried out by Dr. Ridha El Mokni, of the Department of Pharmaceutical Sciences “A”, Laboratory of Botany Faculty of Pharmacy of Monastir, Tunisia, where a voucher specimen (GF 121) has been deposited. 600 g of the aerial parts were dried in the shade at room temperature until constant weight and successively powdered and extracted by maceration in MeOH/CH₂Cl₂ (1:1 v/v) for 3 days (2.5 $\times$ 3 L). After removal of the solvent under vacuum 80 g of organic extract were obtained.

1.6.2 LC-MS/MS Analysis

The qualitative profile of *G. foetida* aerial parts was obtained by LC-MS/MS analysis carried out using an LTQ Orbitrap XL FT-MS mass spectrometer (Thermo Fisher Scientific Spa) equipped with an Ultimate 3000 HPLC (Agilent Technology). Chromatographic separation was achieved on a Kinetex 2.6 μ m polar C18 100 Å (100 $\times$ 3 mm) column (Phenomenex). The LC-MS/MS experiment was carried out using a combination of A (0.1% formic acid in H₂O, v/v) and B (0.1%

formic acid in CH₃CN, v/v) as mobile phase; the gradient elution was optimized as follows: 50% B for 3 min, 50% to 95% B over 20 min, held 10 min before returning to the initial conditions. The total run time, including the column wash and the equilibration, was 30 min, flow rate 0.5 mL/min, injection volume 5 μ L. The MS and MSⁿ spectra were recorded in data-dependent acquisition mode, inducing fragmentation of the five most intense peaks for each scan. Source conditions were as follows: spray voltage, 3.5 kV (positive mode); capillary voltage, 25 V; source temperature, 320 °C; normalized collision energy, 25. The acquisition range was *m/z* 150–1500. Source conditions: spray voltage 3.5 kV (positive mode); capillary voltage 25 V; source temperature 320 °C; normalized collision energy 25%. The acquisition range was *m/z* 150–1500. The spectra were recorded firstly in positive mode, and then in negative modes coupling with a UV detection system.

1.6.3 Purification of secondary metabolites from *G. foetida*

The crude extract was chromatographed on an open column (90 cm length, 4 cm diameter) packed with silica (230–400 mesh). Compounds were eluted with a gradient starting from *n*-hexane followed by *n*-hexane/EtOAc in the following ratios: 8:2, 7:3, 1:1, 4:6, 2:8, 1:9, and then EtOAc, and continuing with EtOAc/MeOH 9:1, 7:3, 1:1, 3:7, 1:9, and finally MeOH. Equally sized fractions of 200 mL each were collected in clean glass flasks, monitored by TLC, and similar ones were pooled to obtain fractions E1 to E18. Based on their similar TLC profiles, fractions E2 to E5 were combined (2.0 g) and purified via MPLC-DAD (silica cartridge 50 Å 25 μ —size 25, 25 g)—column volume (CV) 30 mL/min. The mobile phase was a mixture of A: hexane, B: EtOAc, and C: MeOH, with the gradient method: A-B 9:1 for CV 0–5; from A-B 9:1 to A-B 1:1 for CV 5–40; isocratic A-B 1:1 for CV 40–45; from A-B 1:1 to B for CV 45–50; isocratic B for CV 50–52; isocratic C for 53–55 CV; the flow rate was 25.0 mL/min. UV detection wavelength = 275 nm. This separation afforded 16 fractions that were further purified by HPLC. Fractions 3, 4, 5, and 7 were purified on a HPLC-RI detector equipped with a Rheodyne injector, on a Luna 5 μ m Silica (2) 100 Å 250 mm $\times$ 4.6 mm column, with a flow rate of 1.0 mL/min, using hexane and EtOAc as the mobile phase. Fraction 3 (68.4 mg) was separated by using *n*-hexane/EtOAc 95:5 as the mobile phase, to afford decarboxyamorfrutin 2 (**13**) (R_t = 6.5 min, 18.4 mg) and amorfrutin M (**17**) (R_t = 9.0 min, 1.6 mg). Fraction 4 (67.2 mg) afforded decarboxyamorfrutin A (**2**) (R_t = 15.0 min, 4.7 mg), amorfrutin 2 (**12**) (R_t = 16.0 min, 8.6 mg), the new amorfrutin N (**5**) (R_t = 21.3 min, 0.8 mg), and decarboxydemethoxyamorfrutin C (**10**) (R_t = 24.8 min, 3.0 mg), while fraction 5

(64.3 mg), purified using *n*-hexane/EtOAc 9:1, led to the isolation of amorfrutin C (**11**) (Rt = 8.6 min, 2.0 mg) and amorfrutin A (**1**) (Rt = 9.5 min, 15.4 mg). Fraction 7 (50.3 mg) was separated using *n*-hexane/EtOAc 85:15 as the eluent to yield amorfrutin H (**8**) (Rt = 19.7 min, 2.0 mg), hiravanone (**23**) (Rt = 21.0 min, 6.8 mg), and lonchocarpol A (**22**) (Rt = 22.0 min, 16.9 mg). Fraction 8 (19.9 mg) was purified in the same conditions, leading to the isolation of oleanolic acid (**29**) (Rt = 20.0 min, 1.7 mg). Fraction 9 (75.5 mg) was purified using *n*-hexane/EtOAc 8:2 to yield sakuranetin (**18**) (Rt = 17.8 min, 1.6 mg) and biochanin A (**24**) (Rt = 18.9 min, 1.0 mg). Fraction 10 (15.7 mg) was purified using *n*-hexane/EtOAc 70:30, affording sophoraflavanone B (**19**) (Rt = 19.5 min, 1.4 mg) and exiguaflavanone K (**20**) (Rt = 21.2 min, 1.6 mg), while fraction 11 (31.3 mg) was purified in the same conditions as fraction 10, yielding the new exiguaflavone L (**21**) (Rt = 16.5 min, 2.7 mg). Fraction E7 (40.7 mg) was purified using *n*-hexane/EtOAc 7:3 to afford amorfrutin J (**14**) (Rt = 4.2 min, 2.1 mg), amorfrutin L (**16**) (Rt = 8.6 min, 1.9 mg), and amorfrutin I (**9**) (Rt = 17.5 min, 2.0 mg). Fraction E9 was purified by RP-HPLC-RI using MeOH/H₂O 7:3 + 0.1% of formic acid on a Kinetex 264 C18 100 Å 100 mm × 4.6 mm column, with a flow rate of 1.0 mL/min, to afford amorfrutin F (**6**) (Rt = 8.2 min, 2.4 mg), amorfrutin K (**15**) (Rt = 9.5 min, 1.9 mg), and amorfrutin 3 (**3**) (Rt = 14.9 min, 1.7 mg). Fraction E13 (429.1 mg) was first purified by MPLC-DAD on a Silica C18 Cartridge 50 Å 25 μ—size 40 (55 g)—column volume (CV) 30 mL/min, eluting with *n*-hexane/EtOAc 70:30, with a flow rate of 20 mL/min. This purification led to the isolation on eight fractions. Fractions 13_2 and 13_4 were purified by HPLC-RI on a Luna C18 (2) 250 mm × 4.6 mm column, 5 μm, by using 65:35 H₂O + 0.1% of formic acid/methanol as the mobile phase. Fraction E13_2 (15.7 mg) led to the isolation of myrciaphenone A (**26**) (Rt = 6.7 min, 1.6 mg) and benzyl-β-D-glucopyranoside (**28**) (Rt = 8.3 min, 2.1 mg), while fraction E13_4 afforded 4,6-dihydroxyphenyl-1-butanone-2-β-D-O-glucopyranoside (**27**) (Rt = 10.6 min, 5.6 mg) and phlorizin (**25**) (Rt = 14.2 min, 6.8 mg).

1.6.4 Spectral Data

Amorfrutin E (**4**). Colorless amorphous solid, $[\alpha]_D^{20}$ 0 (c 0.21, CHCl₃); UV (MeOH) $\lambda_{\max}$ (log ϵ) 302 (3.12) 265 (3.63) nm; IR (KBr) $\nu_{\max}$ 3240, 2967, 1643, 1579; ¹H NMR (CDCl₃, 600 MHz) 7.28 (2H, overlapped, H-5', H-7'), 7.18 (1H, t, *J* 7.1 Hz, H-6'), 7.17 (2H, d, *J* 7.1 Hz, H-4', H-8'), 6.20 (1H, s, H-4), 4.99 (1H, s, H-4''a), 4.95 (1H, s, H-4''b), 4.44 (1H, dd, *J* 8.8, 4.1 Hz, H-2''), 3.78 (3H, s, OMe), 3.26 (2H, m, H-1'), 3.14 (1H, dd, *J* 8.8, 5.1 Hz, H-1a''), 3.02 (1H, dd, *J* 13.7, 4.1 Hz, H-1''b), 2.91 (2H, t, *J* 7.8 Hz, H-2'), 1.89 (3H, s, H-5''); ¹³C NMR (CDCl₃, 150 MHz) 173.4 (COOH), 162.0 (C-3),

146.7 (C-5), 144.7 (C-1), 144.6 (C-3''), 142.1 (C-3'), 128.5 (C-4', C-8'), 128.3 (C-5', C-7'), 126.0 (C-6'), 112.6 (C-4''), 111.5 (C-2), 106.5 (C-4), 103.8 (C-6), 86.5 (C-2''), 55.5 (OMe), 39.6 (C-1'), 37.6 (C-2'), 24.0 (C-1''), 19.5 (C-5''); ESIMS m/z 395 $[M+Na]^+$; HRESIMS m/z 395.1469 $[M+Na]^+$ (calcd for $C_{21}H_{24}O_6Na^+$ 395.1471).

Amorfrutin N (**5**). Colorless amorphous solid $[\alpha]_D 0$ (c 0.16, $CHCl_3$). HRESI-MS, m/z 313.1814 $[M+H]^+$ (calc. for $C_{20}H_{25}O_3$ m/z 313.1804). ^1H NMR ($CDCl_3$, 600 MHz) 7.28 (2H, overlapped, H-5', H-7'), 7.20 (3H, overlapped, H-4', H-6', H-8'), 6.49 (1H, s, H-4), 6.26 (1H, s, H-2), 4.89 (1H, s, H-4b''), 4.86 (1H, s, H-4a''), 4.29 (1H, dd, J 8.0, 2.3, H-2''), 3.74 (3H, s, OMe), 3.10 (1H, dd, J 14.6, 2.3, H-1a''), 2.91 (2H, m, H-2'), 2.85 (2H, m, H-1'), 2.78 (1H, dd, J 14.6, 8.0, H-1b''), 1.84 (3H, s, H-5''); ^{13}C NMR ($CDCl_3$, 150 MHz) 158.2 (C-5), 155.8 (C-1), 146.9 (C-3''), 142.7 (C-3), 141.5 (C-3'), 128.6 (C-5', C-7'), 128.4 (C-4', C-8'), 125.9 (C-6'), 112.6 (C-6), 110.3 (C-4, C-4'), 103.5 (C-2), 76.8 (C-2''), 55.5 (OMe), 37.9 (C-1'), 37.8 (C-2''), 30.8 (C-1''), 18.1 (C-5'').

Amorfrutin F (**6**). Colorless amorphous solid, $[\alpha]_D 0$ (c 0.24, $CHCl_3$); UV (MeOH) λ_{max} ($\log \epsilon$) 275 (2.81) nm; IR (KBr) ν_{max} 3540, 3470, 2927, 1656; ^1H NMR ($CDCl_3$, 600 MHz) 7.26 (2H, overlapped, H-4', H-8'), 7.25 (2H, overlapped, H-5', H-7'), 7.17 (1H, t, J 6.8 Hz, H-6'), 6.22 (1H, s, H-4), 4.85 (1H, t, J 8.6 Hz, H-2''), 3.78 (3H, s, OMe), 3.31 (2H, m, H-1'), 3.13 (2H, d, J 8.6 Hz, H-1''), 2.88 (2H, t, J 8.0 Hz, H-2'), 1.39 (1H, s, H-5''), 1.25 (1H, s, H-4''); ^{13}C NMR ($CDCl_3$, 150 MHz) 172.9 (COOH), 160.1 (C-1), 158.5 (C-3), 148.5 (C-5), 142.0 (C-3'), 128.4 (C-5', C-7'), 126.0 (C-4', C-8'), 125.7 (C-6'), 112.1 (C-2), 108.1 (C-4), 104.5 (C-6), 92.3 (C-2''), 71.8 (C-3''), 55.5 (OMe), 37.9 (C-2'), 37.8 (C-1'), 27.5 (C-1''), 26.2 (C-5''), 24.2 (C-4''); HRESIMS m/z 357.1699 $[M+H]^+$ (calcd for $C_{21}H_{25}O_5^+$ 357.1702).

Amorfrutin G (**7**). Colorless amorphous solid, $[\alpha]_D 0$ (c 0.10, $CHCl_3$); UV (MeOH) λ_{max} ($\log \epsilon$) 271 (2.77) nm; IR (KBr) ν_{max} 3549, 3481, 2940, 1651; ^1H NMR ($CDCl_3$, 600 MHz) 7.26 (2H, overlapped, H-5', H-7'), 7.21 (2H, overlapped, H-4', H-8'), 7.19 (1H, overlapped, H-6'), 6.27 (1H, s, H-4), 5.10 (1H, d, J 3.0 Hz, H-1''), 4.59 (1H, d, J 3.0 Hz, H-2''), 3.82 (3H, s, OMe), 3.44 (3H, s, OMe), 3.39 (2H, m, H-1'), 2.90 (2H, m, H-2'), 1.32 (3H, s, H-5''), 1.30 (3H, s, H-4''); ^{13}C NMR ($CDCl_3$, 150 MHz) 173.3 (COOH), 162.3 (C-1), 159.7 (C-3), 145.1 (C-5), 141.3 (C-3'), 128.8 (C-5', C-7'), 128.2 (C-4', C-8'), 125.8 (C-6'), 111.8 (C-2), 108.3 (C-4), 104.7 (C-6), 96.3 (C-2''), 79.0 (C-1''), 71.2 (C-3''), 56.4 (OMe-1''), 55.7 (OMe-3), 37.9 (C-1'), 37.6 (C-2'), 25.5 (C-5''), 25.2 (C-4''); HRESIMS m/z 409.1622 $[M+Na]^+$ (calcd for $C_{22}H_{26}O_6Na^+$ 409.1627).

Amorfrutin H (**8**). Colorless amorphous solid; UV (MeOH) λ_{max} (log ϵ) 314 (2.49) nm; ^1H NMR (CDCl_3 , 600 MHz) 7.62 (1H, s, H-2''), 7.27 (2H, t, J 7.4 Hz, H-5', H-7'), 7.22 (2H, d, J 7.4 Hz, H-4', H-8'), 7.19 (1H, t, J 7.4 Hz, H-6'), 6.87 (1H, s, H-1''), 6.41 (1H, s, H-4), 3.88 (3H, s, OMe), 3.40 (2H, t, J 7.8 Hz, H-1'), 2.97 (2H, t, J 7.8 Hz, H-2'); ^{13}C NMR (CDCl_3 , 150 MHz): 172.9 (COOH), 156.7 (C-1), 156.3 (C-3), 145.1 (C-5), 143.5 (C-2''), 141.8 (C-3'), 128.7 (C-4', C-8'), 128.3 (C-5', C-7'), 125.9 (C-6'), 116.7 (C-2), 107.8 (C-4), 106.2 (C-6), 104.2 (C-1''), 56.2 (OMe), 38.1 (C-2'), 37.2 (C-1'); ESIMS m/z 297 $[\text{M}+\text{H}]^+$; HRESIMS m/z 297.1126 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{18}\text{H}_{17}\text{O}_4^+$ 297.1127).

Amorfrutin I (**9**). Colorless amorphous solid, $[\alpha]_{\text{D}}^0$ (c 0.20, CHCl_3); UV (MeOH) λ_{max} (log ϵ) 302 (2.28) 280 (2.72) nm; IR (KBr) ν_{max} 3370, 2967, 1641; ^1H NMR (CDCl_3 , 600 MHz) 7.28 (2H, overlapped, H-4', H-8'), 7.28 (2H, overlapped, H-5', H-7'), 7.18 (1H, br s, H-6'), 6.29 (1H, s, H-4), 3.88 (1H, d, J 5.3 Hz, H-2''), 3.78 (3H, s, OMe), 3.29 (2H, t, J 7.3 Hz, H-1'), 2.92 (1H, overlapped, H-1a''), 2.91 (2H, overlapped, H-2'), 2.67 (1H, dd, J 17.4, 5.3 Hz, H-1''b), 1.46 (1H, s, H-5''), 1.44 (1H, s, H-4''); ^{13}C NMR (CDCl_3 , 150 MHz) 175.0 (COOH), 160.0 (C-3), 152.5 (C-1), 147.7 (C-5), 142.1 (C-3'), 128.8 (C-4', C-8'), 128.8 (C-5', C-7'), 125.9 (C-6'), 110.5 (C-6), 107.2 (C-4), 106.5 (C-2), 80.0 (C-3''), 68.6 (C-2''), 55.9 (OMe), 38.7 (C-1'), 37.5 (C-2'), 26.3 (C-1''), 24.9 (C-4''), 21.8 (C-5''); HRESIMS m/z 357.1695 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{21}\text{H}_{25}\text{O}_5^+$ 357.1702).

Decarboxydemethoxy amorfrutin C (**10**). Colorless amorphous solid; ^1H NMR (CDCl_3 , 600 MHz) 7.29 (2H, t, J = 7.1 Hz, H-5', H-7'), 7.20 (1H, overlapped, H-6'), 7.19 (2H, overlapped, H-4', H-8'), 6.30 (1H, s, H-6), 5.26 (1H, t, J 6.9 Hz, H-2'''), 5.10 (1H, t, J 6.4 Hz, H-2''), 3.40 (2H, d, J = 6.9 Hz, H-1'''), 3.30 (2H, d, J 6.4 Hz, H-1''), 2.82 (4H, overlapped, H-1', H-2'), 1.82 (3H, s, H-4'''), 1.80 (3H, s, H-4''), 1.75 (3H, s, H-5'''), 1.73 (3H, s, H-5''); ^{13}C NMR (CDCl_3 , 150 MHz) 153.9 (C-3), 152.9 (C-1), 142.0 (C-3'), 138.9 (C-5), 134.8 (C-3'''), 134.3 (C-3''), 128.5 (C-5', C-7'), 128.4 (C-4', C-8'), 126.1 (C-6'), 122.9 (C-2''), 122.4 (C-2'''), 117.4 (C-2), 111.9 (C-4), 109.3 (C-6), 37.7 (C-2'), 35.6 (C-1'), 26.1 (C-5'''), 26.1 (C-5''), 25.6 (C-1''), 25.6 (C-1'''), 18.4 (C-4''), 18.4 (C-4'''); ESIMS m/z 351 $[\text{M}+\text{H}]^+$; HRESIMS m/z 351.2312 $[\text{M}+\text{H}]^+$ (calcd for $\text{C}_{24}\text{H}_{31}\text{O}_2^+$ 351.2324).

Amorfrutin J (**14**). Colorless amorphous solid; $[\alpha]_{\text{D}}^0$ (c 0.19, CHCl_3); UV (MeOH) λ_{max} (log ϵ) 300 (3.00) 267 (3.41) nm; IR (KBr) ν_{max} 3255, 2977, 1648; ^1H NMR (CDCl_3 , 600 MHz) 6.35 (1H, s, H-4), 4.99 (1H, s, H-4''a), 4.83 (1H, s, H-4''b), 4.26 (1H, dd, J 9.0, 3.1 Hz, H-2''), 3.88 (3H, s, OMe), 3.03 (1H, dd, J 13.9, 3.2 Hz, H-1''a), 2.92 (2H, overlapped, H-1'), 2.91 (1H, overlapped, H-

1''b), 1.85 (3H, s, H-5''), 1.58 (2H, overlapped, H-2'), 1.35 (2H, overlapped, H-3'), 1.34 (2H, overlapped, H-4'), 0.90 (3H, t, J 6.7 Hz, H-5'); ^{13}C NMR (CDCl_3 , 150 MHz) 173.8 (COOH), 162.4 (C-3), 148.3 (C-5), 148.0 (C-1), 147.7 (C-3''), 112.2 (C-2), 110.0 (C-4''), 106.7 (C-4), 104.0 (C-6), 76.0 (C-2''), 55.8 (OMe), 37.2 (C-1'), 32.1 (C-3'), 31.7 (C-2'), 29.6 (C-1''), 22.5 (C-4'), 18.1 (C-5''), 14.1 (C-5'); HRESIMS m/z 345.1664 [$\text{M} + \text{Na}$] $^+$ (calcd for $\text{C}_{18}\text{H}_{26}\text{O}_5\text{Na}^+$ 345.1678).

Amorfrutin K (**15**). Colorless amorphous solid, $[\alpha]_D$ 0 (c 0.12, CHCl_3); UV (MeOH) λ_{max} (log ϵ) 302 (3.32) 260 (3.62) nm; IR (KBr) ν_{max} 3300, 2667, 1640; ^1H NMR (CDCl_3 , 700 MHz) 6.36 (1H, s, H-4), 4.84 (1H, t, J 8.8 Hz, H-2''), 3.87 (3H, s, OMe), 3.12 (2H, d, J 8.8 Hz, H-1''), 3.02 (2H, m, H-1'), 1.57 (2H, m, H-2'), 1.37 (3H, s, H-5''), 1.35 (2H, overlapped, H-3'), 1.32 (2H, overlapped, H-4'), 1.24 (3H, s, H-4''), 0.89 (3H, t, J = 7.0 Hz, H-5'); ^{13}C NMR (CDCl_3 , 175 MHz) 173.1 (COOH), 160.3 (C-1), 158.7 (C-3), 111.6 (C-2), 149.5 (C-5), 107.5 (C-4), 104.9 (C-6), 92.0 (C-2''), 71.8 (C-3''), 55.5 (OMe), 35.3 (C-1'), 32.0 (C-3'), 31.6 (C-2'), 27.4 (C-1''), 26.0 (C-5''), 24.0 (C-4''), 22.4 (C-4'), 14.1 (C-5'); HRESIMS m/z 323.1852 [$\text{M} + \text{H}$] $^+$ (calcd for $\text{C}_{18}\text{H}_{27}\text{O}_5^+$ 323.1858).

Amorfrutin L (**16**). Colorless amorphous solid, $[\alpha]_D$ 0 (c 0.12, CHCl_3); UV (MeOH) λ_{max} (log ϵ) 299 (3.30) 259 (3.60) nm; IR (KBr) ν_{max} 3310, 2666, 1636; ^1H NMR (CDCl_3 , 600 MHz) 6.41 (1H, s, H-4), 3.86 (1H, overlapped, H-2''), 3.86 (3H, s, OMe), 2.98 (2H, m, H-1'), 2.91 (1H, dd, J 17.7, 5.1 Hz, H-1''a), 2.67 (1H, dd, J 17.4, 5.6 Hz H-1''b), 1.60 (2H, overlapped, H-2'), 1.37 (2H, overlapped, H-3'), 1.37 (3H, s, H-5''), 1.36 (2H, overlapped, H-4'), 1.35 (3H, s, H-4''), 0.89 (3H, t, J 7.0 Hz, H-5'); ^{13}C NMR (CDCl_3 , 150 MHz), 174.9 (COOH), 159.9 (C-3), 152.4 (C-1), 148.8 (C-5), 113.2 (C-2), 110.8 (C-6), 106.4 (C-4), 78.5 (C-3''), 68.5 (C-2'') 55.7 (OMe), 36.0 (C-1'), 32.2 (C-3'), 29.5 (C-2'), 26.2 (C-1''), 24.6 (C-4''), 22.3 (C-4'), 21.5 (C-5''), 14.0 (C-5'); HRESIMS m/z 345.1684 [$\text{M} + \text{Na}$] $^+$ (calcd for $\text{C}_{18}\text{H}_{26}\text{O}_5\text{Na}$ 345.1678).

Amorfrutin M (**17**). Colorless amorphous solid, $[\alpha]_D$ 0 (c 0.16, CHCl_3); UV (MeOH) λ_{max} (log ϵ) 283 (3.53) 228 (3.82) nm; ^1H NMR (CDCl_3 , 600 MHz): 6.26 (1H, s, H-6), 5.25 (1H, tt, J 7.0, 1.3 Hz, H-2''), 5.13 (1H, tt, J 6.7, 1.3 Hz, H-2'''); 3.38 (2H, d, J 7.0 Hz, H-1''), 3.28 (2H, d, J 6.7 Hz, H-1'''), 2.50 (2H, t, J 7.7 Hz, H-1'), 1.81 (6H, s, H-4'', H-4'''), 1.74 (6H, s, H-5'', H-5'''), 1.51 (2H, m, H-2'), 1.33 (2H, overlapped, H-3'), 1.33 (2H, overlapped, H-4'), 0.89 (3H, t, J 7.2 Hz, H-5'); ^{13}C NMR (CDCl_3 , 150 MHz): 153.7 (C-3), 152.8 (C-1), 140.0 (C-5), 134.9 (C-3''), 133.9 (C-3'''), 122.9 (C-2'''), 122.6 (C-2''), 117.1 (C-4), 111.5 (C-2), 108.8 (C-6), 33.5 (C-1'), 31.9 (C-3'), 31.1 (C-2'), 25.9 (C-

5"), 25.9 (C-5^{'''}), 25.4 (C-1^{'''}), 22.6 (C-4'), 22.6 (C-1"), 17.8 (C4^{'''}), 17.4 (C-4"), 14.0 (C-5'); HRESIMS m/z 317.2469 [M+H]⁺ (calcd for C₂₁H₃₃O₂⁺ 317.2481).

Exiguaflavone L (**21**) Colorless amorphous solid [α]_D - 4.2 (c 0.19, CHCl₃). HRESI-MS, m/z 403.1390 [M + H]⁺ (calc. for C₂₁H₂₃O₈ m/z 403.1393). ¹H NMR (600 MHz, CDCl₃): 6.96 (1H, overlapped, H-2'), 6.94 (1H, overlapped, H-5'), 6.93 (1H, overlapped, H-6'), 6.07 (1H, s, H-6), 5.32 (1H, m, H-2), 5.07 (1H, s, H-4''), 4.48 (1H, m, H-2''), 3.93 (3H, s, OMe), 3.08 (1H, m, H-1''a), 3.09 (1H, m, H-3a), 2.79 (2H, m, H-3b and H-1''b), 1.87 (3H, s, H-5''); ¹³C NMR (125 MHz, CDCl₃): 196.0 (C-4), 163.7 (C-7), 163.5 (C-9), 161.9 (C-5), 146.4 (C-3'), 146.3 (C-4'), 143.4 (C-3''), 135.9 (C-1'), 119.6 (C-6'), 114.4 (C-5'), 113.5 (C-4''), 108.8 (C-2'), 105.3 (C-8), 102.8 (C-10), 96.4 (C-6), 88.8 (C-2''), 79.2 (C-2), 55.9 (OMe), 43.4 (C-3), 23.8 (C-1''), 18.9 (C-5'').

1.6.5 Computational Methods

Ligands were built with UCSF Chimera 1.1714 followed by initial energy minimization (EM) at the molecular mechanics level using AM1-BC charges. The molecules were then fully optimized using the GAMESS program¹⁵ at the Hartree-Fock level with the STO-3G basis set and subjected to HF/ 6-31G*/STO-3G single-point calculations to derive the partial atomic charges using the RESP procedure (Morris et al., 2009). Docking studies were performed with AutoDock 4.2 and Autodock Vina by using PPAR α and PPAR γ crystallographic structures (PDB id: 2P54 and 6LX4 for PPAR α and 2F4B and 2YFE for PPAR γ). Two X-ray structures have been used for each receptor isoform to cope with differences both in Ω -loop and residue side chain conformations. Both proteins and ligands were processed with AutoDock Tools (ADT) package version 1.5.6rc117 to merge nonpolar hydrogens and calculate Gasteiger charges. Grids for docking evaluation with a spacing of 0.375 Å and 60 × 60 × 70 points, centered on the ligand binding site, were generated using the program AutoGrid 4.2 included in Autodock 4.2 distribution, following the docking protocol already published (Iannotti et al., 2020). The complexes, selected on the basis of binding energy and cluster population, were completed by addition of all hydrogen atoms, and they underwent EM and then MD simulations with the Amber20 pmemd.cuda module (Fox et al., 1998), using the 14SB version of the AMBER force field for the protein and gaff parameters for the ligand (Wang et al., 2004). To perform MD simulations in the solvent, the complexes were confined in TIP3P water periodic boxes exhibiting a minimum distance between solute atoms and box surfaces of 10 Å using the tleap module of the AmberTools20 package. The systems were then neutralized by addition of

counterions (Na^+) and subjected to 1000 steps of EM, with solute atoms harmonically restrained to their starting positions ($K_r = 10 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-1}$). Then, 500 ps of restrained MD ($K_r = 5 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-1}$) at constant pressure was run on each solvated complex, gradually heating the system to 300 K, followed by 500 ps of restrained MD ($K_r = 5 \text{ kcal}\cdot\text{mol}^{-1}\cdot\text{\AA}^{-1}$) at constant temperature (300 K) and pressure (1 atm) to adjust system density. Production MD simulations were carried out at a constant temperature (300 K) and pressure (1 atm) for 100 ns, with a time step of 2 fs. Bonds involving hydrogens were constrained by using the SHAKE algorithm. The cpptraj module of AmberTools20 and UCSF Chimera 1.17 were used to perform MD analysis and to draw the figures, respectively.

1.6.6 Cell Culture, Transfection, and Luciferase Assay

COS-7 cells (monkey kidney fibroblast-like cells) were grown in DMEM supplemented with 10% fetal bovine serum and 1% Pen/Strep under the standard conditions. For transfection, cells were plated in 24-well dishes. The day after, at about 60-70% confluence, cells were transfected using Lipofectamine 2000 reagent (cat. #11668019; Life Technologies) with a mix of three plasmids: (a) CMX-Gal4-hPPAR α ; (b) TK-MH100x4-Luc containing the UAS enhancer; and (c) Renilla luciferase (pRL, Promega, cat. #E2231) according to the manufacturer's instruction. Cells transfected with an equal amount of pcDNA3.1 (cat. #V79020; Life Technologies) were used as the control condition. On day 3, cells were treated for 24 h with test compounds. Rosiglitazone $0.01 \mu\text{M}$ (cat. #5325; Tocris UK) and GW7647 (cat. #1677; Tocris UK) were used as positive internal controls for PPAR γ and PPAR α . At the end of treatment, the luciferase reporter activity was detected using the Dual-Glo luciferase assay kit (Promega, cat. #E2920) using a GloMax Discover microplate reader device (Xu et al., 2011).

1.6.7 MTT Assay

COS-7 cells were plated in 24-well plates with 2×10^5 cells per well. On the next day, cells were treated with test compounds for 24 h. Cells treated with DMSO ($<0.03\%$), used to dissolve all compounds, were used as the control group (vehicle). To assess cell viability using this method, for the entire procedure we strictly followed the recommendation reported by Ghasemi et al.²¹ Briefly, the ability of cells to reduce MTT to formazan salts was assumed as a vitality sign, being dependent on functional mitochondrial integrity. After 3 h of incubation,

formazan salts were dissolved in three volumes of 2-methyl-2-propanol and absorbance was read at 620 nm using a GloMax Discover microplate reader device (Xü et al., 2011).

1.6.8 Mitochondrial Activity

Mitochondrial activity was measured in HuH7 and C2C12 cells using a MitoTracker® Red CMXRos probe (Thermo Fisher Scientific, Waltham, MA, USA). The probe was diluted in DMEM to a final dye working solution of 100 nM. HuH7 and C2C12 cells were washed twice in PBS before being cultivated in the presence of the probe for 60 min in a cell incubator set to 37°C and 5% CO₂. Upon incubation, cells were extensively washed in DMEM and once in PBS, and then fixed in 3.7% formaldehyde for 30 min. MitoTracker fluorescence was measured using a Perkin Elmer Envision 2105 Multiplate Reader (Perkin Elmer, Waltham, MA, USA) with the following parameters: $\lambda_{\text{excitation}}$ at 579 nm and $\lambda_{\text{emission}}$ at 599 nm. Cells were then permeabilized in 0.1% Triton X-100 in PBS and stained with the nuclear dye DAPI. DAPI fluorescence was measured using a $\lambda_{\text{excitation}}$ at 351 nm and $\lambda_{\text{emission}}$ at 450 nm. In order to normalize mitochondrial activity to the total number of cells, Mitotraker fluorescence was normalized to DAPI fluorescence.

1.6.9 NBDG Glucose Uptake Assay

C2C12 cells were seeded (5×10^3 /well) on a 96-well black bottom microtiter plate (Perkin Elmer, Waltham, MA, USA) in a final volume of 100 μ L/well of growth media. When the cells achieved 80–90% confluency, the culture medium was carefully removed and replaced with 100 μ L of HBSS containing 100 μ M 2-deoxyglucose, 0.4 g/L Bovine Serum Albumin, and 1.27 mM CaCl₂ (without any growth factors or FBS). When indicated, test molecules were added to the wells at the indicated concentration in DMSO and cells incubated for 1 h at 37°C. Upon incubation, the culture medium was further supplemented with 64 μ M of the fluorescent glucose tracer 2-NBDG. After 45 min of incubation with the fluorescent probe, plates were washed twice in PBS. Uptake of 2-NDBG was determined using a Perkin Elmer Envision 2105 Multiplate Reader (Perkin Elmer) with the following parameters: $\lambda_{\text{excitation}}$ at 471 nm, $\lambda_{\text{emission}}$ at 529 nm, and monochromator cut-off at 360 nm. Following the determination of 2-NDBG, cells were fixed in 3.7% paraformaldehyde for 30 min, permeabilized in 0.1% Triton X-100 in PBS, and stained with the nuclear dye DAPI (100 μ M). The parameters of $\lambda_{\text{excitation}}$ at 351 nm and $\lambda_{\text{emission}}$ at 450 nm were used to measure the fluorescence of DAPI. Normalized

glucose uptake is indicated as the ratio between intracellular 2-NDBG fluorescence and DAPI fluorescence.

1.6.10 RNA Purification and mRNA Quantification by Real-Time RT-PCR

Total RNAs were purified from Huh7 cell lines with an RNeasy®Plus Mini Kit (QIAGEN, Hilden, Germany) according to the manufacturer's protocol. A mass of 0.7 µg of purified RNAs were converted to cDNAs with a High-Capacity cDNA Reverse Transcription Kit (Applied biosystems, Thermo Fisher Scientific, Waltham, MA, USA) as recommended by the manufacturer. The cDNAs were amplified by real-time RT-PCR in a Step One Plus Real-Time PCR System (Applied biosystems, Thermo Fisher Scientific) with the fluorescent double-stranded DNA-binding dye SYBR™ Green PCR Master mix (Applied biosystems, Thermo Fisher Scientific) according to the manufacturer's protocol. Specific primers for each gene were designed to work under the same cycling conditions (95°C for 10 min followed by 40 cycles at 95°C for 15 s, 52°C for 30 s, and 72°C for 30 s), thereby generating products of comparable size (about 100–400 bp for each amplification). Primer combinations were positioned whenever possible to span an exon–exon junction and the RNA was digested with DNase to avoid genomic DNA interference. All samples were run in triplicate and the results, expressed as N-fold differences in target gene expression, were determined as follows: $=2^{-(\Delta CT[\text{target}] - \Delta CT[\text{control}])}$. Primer sequences for each gene analyzed were designed as follows: CYP17A1: FW 5'-GAGCTTGAGGCACGAGAAC-3' RV 5'-TGGAAATGGTGTGTTGCTTGCG-3'; PGC1α: FW 5'-ATTGCCCTCATTTGATGCGC-3' RV: 5'-TAGCTGAGTGTTGGCTGGTG-3'; GAPDH: FW 5'-GCACCACCAACTGCTTA-3' RV: 5'-AGTAGAGGCAGGGATGAT-3'.

1.6.11 GC/MS Analysis

HuH7 cells (2×10^6) were scraped in ice-cold water and centrifuged at $10,000 \times g$ for 5 min at 4 °C. Membrane pellets were dried and dissolved in 1 mL of ice-cold dichloromethane. Insoluble material was removed by centrifugation at high speed for 10 min at 4°C. The supernatants were dried and resuspended in acetonitrile. Samples were solubilized in pyridine (50 µL) and derivatized with 25 µL of *N,O*-bis(trimethylsilyl)(TMS)trifluoroacetamide (BSTFA) with a reaction time of 90 min. A volume of 1 µL was injected, with a split ratio of 1:10. GC-MS analyses were carried out on a Shimadzu GCMS 2010 plus (Kyoto, Japan) with the following parameters: Injection temperature 280°C, ramp 0–1.00 min 100 °C, 1.00–6.00 min 100–320° C, hold for 2.33

min. Column flow 1.10 mL/min. Linear velocity 39 cm/s. Helium gas was used. Ion source temperature 200°C, interface 320°C, solvent cut 5.9 min, scan 35–600 m/z . Detector voltage 0.1 kV. Separation was performed on an Agilent (Santa Clara, CA, USA.) SIL-8, 30 m × 0.25 mm column, 0.25 μm . Cholesterol-TMS (m/z : 458, 368; 329; 129; 73; 75), palmitic acid-TMS (m/z : 313; 117; 73), and myristic acid-TMS (m/z : 285; 145; 117; 73) were identified by comparing their retention time and fragmentation to library standards.

References

- Al-Hmadi, H.B., Serino, E., Pastore, A., Chianese, G., Hammami, S., Stornaiuolo, M. and Taglialatela-Scafati, O., **2024**. Metabolites from Aerial Parts of *Glycyrrhiza foetida* as Modulators of Targets Related to Metabolic Syndrome. *Biomolecules*, *14*(4), p.467.
- Asakawa, Y. and Asakawa, Y., **1995**. Chemical constituents of the bryophytes. *Progress in the chemistry of organic natural products*, pp.1-562.
- Bala, M., Gupta, P., Gupta, S., Dua, A., Injeti, E. and Mittal, A., **2021**. Efficient and modified 2-NBDG assay to measure glucose uptake in cultured myotubes. *Journal of Pharmacological and Toxicological Methods*, *109*, p.107069.
- Balakumar, P., Mahadevan, N. and Sambathkumar, R., **2019**. A contemporary overview of PPAR α / γ dual agonists for the management of diabetic dyslipidemia. *Current Molecular Pharmacology*, *12*(3), pp.195-201.
- Behl, T., Gupta, A., Chigurupati, S., Singh, S., Sehgal, A., Badavath, V.N., Alhowail, A., Mani, V., Bhatia, S., Al-Harrasi, A. and Bungau, S., **2022**. Natural and synthetic agents targeting reactive carbonyl species against metabolic syndrome. *Molecules*, *27*(5), p.1583.
- Blaschek, W., **2017**. Natural products as lead compounds for sodium glucose cotransporter (SGLT) inhibitors. *Planta medica*, *83*(12/13), pp.985-993.
- Chang, Y.C., Nair, M.G., Santell, R.C. and Helferich, W.G., **1994**. Microwave-mediated synthesis of anticarcinogenic isoflavones from soybeans. *Journal of agricultural and food chemistry*, *42*(9), pp.1869-1871.
- Chen, R., Liu, X., Zou, J., Yin, Y., Ou, B., Li, J., Wang, R., Xie, D., Zhang, P. and Dai, J., **2013**. Regio- and stereospecific prenylation of flavonoids by *Sophora flavescens* prenyltransferase. *Advanced Synthesis & Catalysis*, *355*(9), pp.1817-1828.
- Chen, C., Wu, Y. and Du, L., **2016**. Qualitative and quantitative analysis of amorfrutins, novel antidiabetic dietary natural products, by HPLC. *Pharmaceutical Biology*, *54*(3), pp.488-493.
- Choi, H.W., Tian, M., Manohar, M., Harraz, M.M., Park, S.W., Schroeder, F.C., Snyder, S.H. and Klessig, D.F., **2015**. Human GAPDH is a target of aspirin's primary metabolite salicylic acid and its derivatives. *PLoS One*, *10*(11), p.e0143447.
- de Brito Monteiro, L., Davanzo, G.G., de Aguiar, C.F. and Moraes-Vieira, P.M., **2020**. Using flow cytometry for mitochondrial assays. *MethodsX*, *7*, p.100938.

- de Groot, J.C., Weidner, C., Krausze, J., Kawamoto, K., Schroeder, F.C., Sauer, S. and Büssow, K., **2013**. Structural characterization of amorfrutins bound to the peroxisome proliferator-activated receptor γ . *Journal of Medicinal Chemistry*, 56(4), pp.1535-1543.
- Esterbauer, H., Oberkofler, H., Krempler, F. and Patsch, W., **1999**. Human peroxisome proliferator activated receptor gamma coactivator 1 (PPARGC1) gene: cDNA sequence, genomic organization, chromosomal localization, and tissue expression. *Genomics*, 62(1), pp.98-102.
- Errichiello, F., D'Amato, M., Gambuti, A., Moio, L., Pastore, A., Hekmat, A.H., Stornaiuolo, M., Serino, E., Taglialatela-Scafati, O. and Forino, M., **2023**. Oleanolic acid: A promising antidiabetic metabolite detected in Aglianico grape pomace. *Journal of Functional Foods*, 104, p.105548.
- Fox, T. and Kollman, P.A., **1998**. Application of the RESP methodology in the parametrization of organic solvents. *The Journal of Physical Chemistry B*, 102(41), pp.8070-8079.
- Fuhr, L., Rousseau, M., Plauth, A., Schroeder, F.C. and Sauer, S., **2015**. Amorfrutins are natural PPAR γ agonists with potent anti-inflammatory properties. *Journal of natural products*, 78(5), pp.1160-1164.
- Ghisalberti, E.L., Jefferies, P.R. and McAdam, D., **1981**. Isoprenylated resorcinol derivatives from *Glycyrrhiza acanthocarpa*. *Phytochemistry*, 20(8), pp.1959-1961.
- Gurbuz, P., Baran, M.Y., Demirezer, L.O., Guvenalp, Z. and Kuruuzum-Uz, A., **2018**. Phenylacylated-flavonoids from *Peucedanum chryseum*. *Revista Brasileira de Farmacognosia*, 28(2), pp.228-230.
- Hong, M.K., Han, Y., Park, H.J., Shin, M.R., Roh, S.S. and Kwon, E.Y., **2023**. The synergistic action of metformin and *Glycyrrhiza uralensis* Fischer extract alleviates metabolic disorders in mice with diet-induced obesity. *International Journal of Molecular Sciences*, 24(2), p.936.
- Husain, I., Bala, K., Khan, I.A. and Khan, S.I., **2021**. A review on phytochemicals, pharmacological activities, drug interactions, and associated toxicities of licorice (*Glycyrrhiza* sp.). *Food Frontiers*, 2(4), pp.449-485.
- Iannotti, F.A., De Maio, F., Panza, E., Appendino, G., Taglialatela-Scafati, O., De Petrocellis, L., Amodeo, P. and Vitale, R.M., **2020**. Identification and characterization of cannabimovone, a cannabinoid from *Cannabis sativa*, as a novel PPAR γ agonist via a combined computational and functional study. *Molecules*, 25(5), p.1119.

- Iinuma, M., Yokoyama, J., Ohyama, M., Tanaka, T. and Ruangrunsi, N., **1994**. Eight phenolic compounds in root of *Sophora exigua*. *Phytochemistry*, 35(3), pp.785-789.
- Ito, C., Sato, K., Oka, T., Inoue, M., Ju-Ichi, M., Omura, M. and Furukawa, H., **1989**. Two flavanones from Citrus species.
- Ko, R.K., Kang, M.C., Kim, B.S., Han, J.H., Kim, G.O. and Lee, N.H., **2009**. A new phloroglucinol glycoside from *Aster subulatus* Michx. *Bull. Korean Chem. Soc*, 30, pp.1167-1169.
- Kung, J. and Henry, R.R., **2012**. Thiazolidinedione safety. *Expert opinion on drug safety*, 11(4), pp.565-579.
- Lavecchia, A. and Di Giovanni, C., **2015**. Amorphutins are efficient modulators of peroxisome proliferator-activated receptor gamma (PPAR γ) with potent antidiabetic and anticancer properties: a patent evaluation of WO2014177593 A1. *Expert Opinion on Therapeutic Patents*, 25(11), pp.1341-1347.
- Lee, W., Ham, J., Kwon, H.C., Kim, Y.K. and Kim, S.N., **2013**. Anti-diabetic effect of amorphastilbol through PPAR α / γ dual activation in db/db mice. *Biochemical and Biophysical Research Communications*, 432(1), pp.73-79.
- Manfredi, K.P., Vallurupalli, V., Demidova, M., Kindscher, K. and Pannell, L.K., **2001**. Isolation of an anti-HIV diprenylated bibenzyl from *Glycyrrhiza lepidota*. *Phytochemistry*, 58(1), pp.153-157.
- Mitscher, L.A., Park, Y.H., Alshamma, A., Hudson, P.B. and Haas, T., **1981**. Amorphutin A and B, bibenzyl antimicrobial agents from *Amorpha fruticosa*. *Phytochemistry*, 20(4), pp.781-785.
- Morris, G.M., Huey, R., Lindstrom, W., Sanner, M.F. and Belew, R., **2009**. K., Goodsell DS, Olson AJ AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem*, 30(16), pp.2785-2791.
- Pastorino, G., Cornara, L., Soares, S., Rodrigues, F. and Oliveira, M.B.P., **2018**. Liquorice (*Glycyrrhiza glabra*): A phytochemical and pharmacological review. *Phytotherapy research*, 32(12), pp.2323-2339.
- Pollastro, F., De Petrocellis, L., Schiano-Moriello, A., Chianese, G., Heyman, H., Appendino, G. and Taglialatela-Scafati, O., **2017**. Amorphutin-type phytocannabinoids from *Helichrysum umbraculigerum*. *Fitoterapia*, 123, pp.13-17.

- Qureshi, J.A., Memon, Z., Ismail, K., Saher, F., Motiani, V. and Mushtaq, Z., **2020**. Anti-Hyperglycemic and Anti-Dyslipidemic Activities of Glycyrrhiza Glabra Root Extract In Diabetic Rats. *Journal of Islamic International Medical College (JIIMC)*, 15(2), pp.98-103.
- Roussis, V., Ampofo, S.A. and Wiemer, D.F., **1987**. Flavanones from *Lonchocarpus minimiflorus*. *Phytochemistry*, 26(8), pp.2371-2375.
- Santa, K., Kumazawa, Y. and Nagaoka, I., **2023**. Prevention of metabolic syndrome by phytochemicals and vitamin D. *International Journal of Molecular Sciences*, 24(3), p.2627.
- Sauer, S., **2014**. Amorphutins: a promising class of natural products that are beneficial to health. *ChemBioChem*, 15(9), pp.1231-1238.
- Sauer, S., Weidner, C., Kliem, M., Schroeder, F. C., Micikas, R.J., **2014**. Amorphutin analogs as PPAR γ -modulators. PCT Int. Appl WO 2014177593 A1 20141106.
- Seigler, D.S., Pauli, G.F., Nahrstedt, A. and Leen, R., **2002**. Cyanogenic allosides and glucosides from *Passiflora edulis* and *Carica papaya*. *Phytochemistry*, 60(8), pp.873-882.
- Serino, E., Iannotti, F.A., Al-Hmadi, H.B., Caprioglio, D., Moriello, C., Masi, F., Hammami, S., Appendino, G., Vitale, R.M. and Taglialatela-Scafati, O., **2023**. PPAR α/γ -targeting amorphutin phytocannabinoids from aerial parts of *Glycyrrhiza foetida*. *Journal of Natural Products*, 86(11), pp.2435-2447.
- Sen, S. and Singh, R., **2021**. Glycyrrhiza glabra alcoholic root extract ameliorates hyperglycemia, hyperlipidemia, and glycation-induced free iron-mediated oxidative reactions. *Journal of food biochemistry*, 45(12), p.e13970.
- Sharifi-Rad, J., Quispe, C., Herrera-Bravo, J., Belén, L.H., Kaur, R., Kregiel, D., Uprety, Y., Beyatli, A., Yeskaliyeva, B., Kırkın, C. and Özçelik, B., **2021**. Glycyrrhiza Genus: Enlightening Phytochemical Components for Pharmacological and Health-Promoting Abilities. *Oxidative medicine and cellular longevity*, 2021(1), p.7571132.
- Sil, R., Ray, D. and Chakraborti, A.S., **2013**. Glycyrrhizin ameliorates insulin resistance, hyperglycemia, dyslipidemia and oxidative stress in fructose-induced metabolic syndrome-X in rat model.
- Sobolev, V.S., Krausert, N.M. and Gloer, J.B., **2016**. New monomeric stilbenoids from peanut (*Arachis hypogaea*) seeds challenged by an *Aspergillus flavus* strain. *Journal of agricultural and food chemistry*, 64(3), pp.579-584.

- Srebnik, M., Lander, N., Breuer, A. and Mechoulam, R., **1984**. Base-catalysed double-bond isomerizations of cannabinoids: structural and stereochemical aspects. *Journal of the Chemical Society, Perkin Transactions 1*, pp.2881-2886.
- Vasconcelos, J.M., Silva, A.M. and Cavaleiro, J.A., **1998**. Chromones and flavanones from *Artemisia campestris* subsp. *maritima*. *Phytochemistry*, 49(5), pp.1421-1424.
- Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A. and Case, D.A., **2004**. Development and testing of a general amber force field. *Journal of computational chemistry*, 25(9), pp.1157-1174.
- Weidner, C., de Groot, J.C., Prasad, A., Freiwald, A., Quedenau, C., Kliem, M., Witzke, A., Kodelja, V., Han, C.T., Giegold, S. and Baumann, M., **2012**. Amorphutins are potent antidiabetic dietary natural products. *Proceedings of the National Academy of Sciences*, 109(19), pp.7257-7262.
- Weidner, C., Rousseau, M., Micikas, R.J., Fischer, C., Plauth, A., Wowro, S.J., Siems, K., Hetterling, G., Kliem, M., Schroeder, F.C. and Sauer, S., **2016**. Amorphutin C induces apoptosis and inhibits proliferation in colon cancer cells through targeting mitochondria. *Journal of natural products*, 79(1), pp.2-12.
- Xü, K., Deng, R. and Lü, H.T., **2011**. Preparative isolation and purification of phlorizin from apple tree bark by high-speed counter-current chromatography.
- Xu, T., Yang, M., Li, Y., Chen, X., Wang, Q., Deng, W., Pang, X., Yu, K., Jiang, B., Guan, S. and Guo, D.A., **2013**. An integrated exact mass spectrometric strategy for comprehensive and rapid characterization of phenolic compounds in licorice. *Rapid Communications in Mass Spectrometry*, 27(21), pp.2297-2309.
- Zheng, Y., Lee, J., Lee, E.H., In, G., Kim, J., Lee, M.H., Lee, O.H. and Kang, I.J., **2020**. A combination of Korean Red Ginseng Extract and *Glycyrrhiza glabra* L. extract enhances their individual anti-obesity properties in 3T3-L1 adipocytes and C57BL/6J obese mice. *Journal of medicinal food*, 23(3), pp.215-223.

APPENDIX D

Here are reported the spectra of the previously undescribed compounds isolated from *Glycyrrhiza foetida* (described in PART 3 – Chapter 1 of this manuscript).

Table S1: Detailed fragmentation pattern analysis of detected amorfrutins from <i>G. foetida</i> LC-MS/MS analysis.....	201
Figures S1-S4: NMR spectra of compound 4 (CDCl ₃ , 600MHz)	203
Figures S5-S8: NMR spectra of compound 5 (CDCl ₃ , 600MHz).....	205
Figures S9-S12: NMR spectra of compound 6 (CDCl ₃ , 600MHz).....	207
Figures S13-S17: NMR spectra of compound 7 (CDCl ₃ , 600MHz).....	209
Figures S18-S21: ¹ H NMR spectrum of compound 8 (CDCl ₃ , 600MHz)	211
Figures S22-S25: NMR spectra of compound 9 (CDCl ₃ , 600MHz).....	213
Figures S26-S29: NMR spectra of compound 10 (CDCl ₃ , 600MHz)	215
Figures S30-S33: NMR spectra of compound 14 (CDCl ₃ , 600MHz)	217
Figures S34-S37: NMR spectra of compound 15 (CDCl ₃ , 700MHz)	219
Figures S38-41: NMR spectra of compound 16 (CDCl ₃ , 700MHz)	221
Figures S42-S45: NMR spectra of compound 17 (CDCl ₃ , 600MHz)	223
Figure S46-S49: ¹ H NMR spectrum of compound 22 (CDCl ₃ , 600MHz)	225
Figure S50: Root-mean-square deviation (RMSD) plots of amorfrutins in complex with PPARα/γ over MD trajectories after best fitting of protein backbones. The ligand RMSD plot was smoothed with a five-point window running average.	228

Table S1: Detailed fragmentation pattern analysis of detected amorfrutins from *G. foetida* LC-MS/MS analysis.

Compound	Fragment	Fragment formula	Fragment ion (<i>m/z</i>)	Δ ppm
<i>Amorfrutin 3 (3)</i>	M - H ₂ O	C ₂₁ H ₂₄ O ₅	339.1592	0.367
	M - 2H ₂ O	C ₂₁ H ₂₀ O ₃	321.1487	0.495
	M - C ₄ H ₈ O	C ₁₇ H ₁₆ O ₄	285.1117	-1.457
	M - CH ₃ OH	C ₂₀ H ₂₀ O ₄	325.1407	-0.110
Decarboxyamorfrutin A (2)	M - CH ₃ OH	C ₁₉ H ₂₀ O	265.1586	0.031
	M - C ₄ H ₈	C ₁₆ H ₁₆ O ₂	241.1224	0.430
	M - C ₅ H ₈	C ₁₅ H ₁₆ O ₂	229.1227	1.500
	M - C ₆ H ₁₂ O ₂	C ₁₄ H ₁₂	181.1015	1.618
Amorfrutin A (1)	M - H ₂ O	C ₂₁ H ₂₂ O ₃	323.1638	-1.179
	M - C ₄ H ₈	C ₁₇ H ₁₆ O ₄	285.1119	-0.896
	M - C ₄ H ₁₀ O	C ₁₇ H ₁₄ O ₃	267.1012	-1.388
	M - C ₅ H ₈ O ₂	C ₁₆ H ₁₆ O ₂	241.1218	-2.100
Decarboxyamorfrutin 2 (13)	M - CH ₃ OH	C ₁₆ H ₂₂ O	231.1747	1.463
	M - C ₄ H ₈	C ₁₃ H ₁₆ O ₂	207.1378	-0.610
	M - C ₅ H ₈	C ₁₂ H ₁₈ O ₂	195.1379	-0.238
	M - C ₉ H ₁₈	C ₈ H ₈ O ₂	137.0594	-2.452
Amorfrutin 2 (12)	M - H ₂ O	C ₁₈ H ₂₄ O ₃	289.1796	-0.661
	M - C ₄ H ₈	C ₁₄ H ₁₈ O ₄	251.1279	0.336
	M - C ₄ H ₁₀ O	C ₁₄ H ₁₆ O ₃	233.1171	-0.390
Amorfrutin C or B	M - H ₂ O	C ₂₆ H ₃₀ O ₃	391.2269	0.406
	M - C ₄ H ₈	C ₂₂ H ₂₄ O ₄	353.1741	-1.800
	M - C ₅ H ₈ O ₂	C ₂₁ H ₂₄ O ₂	309.1869	-0.021

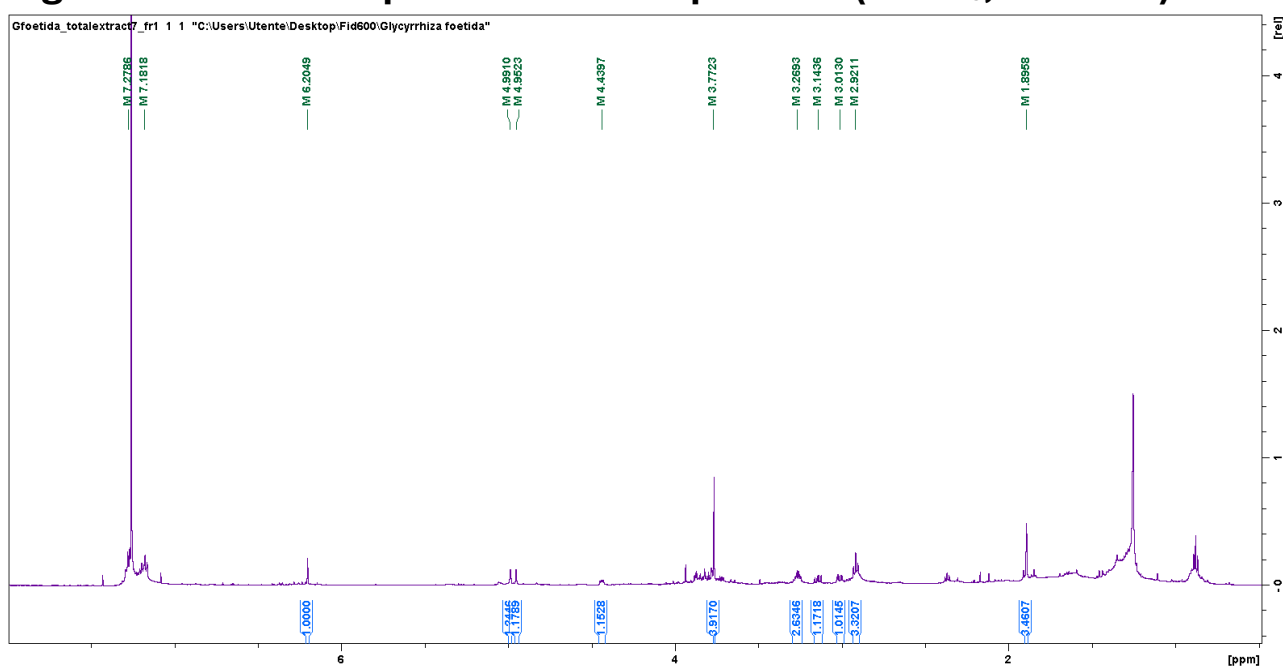
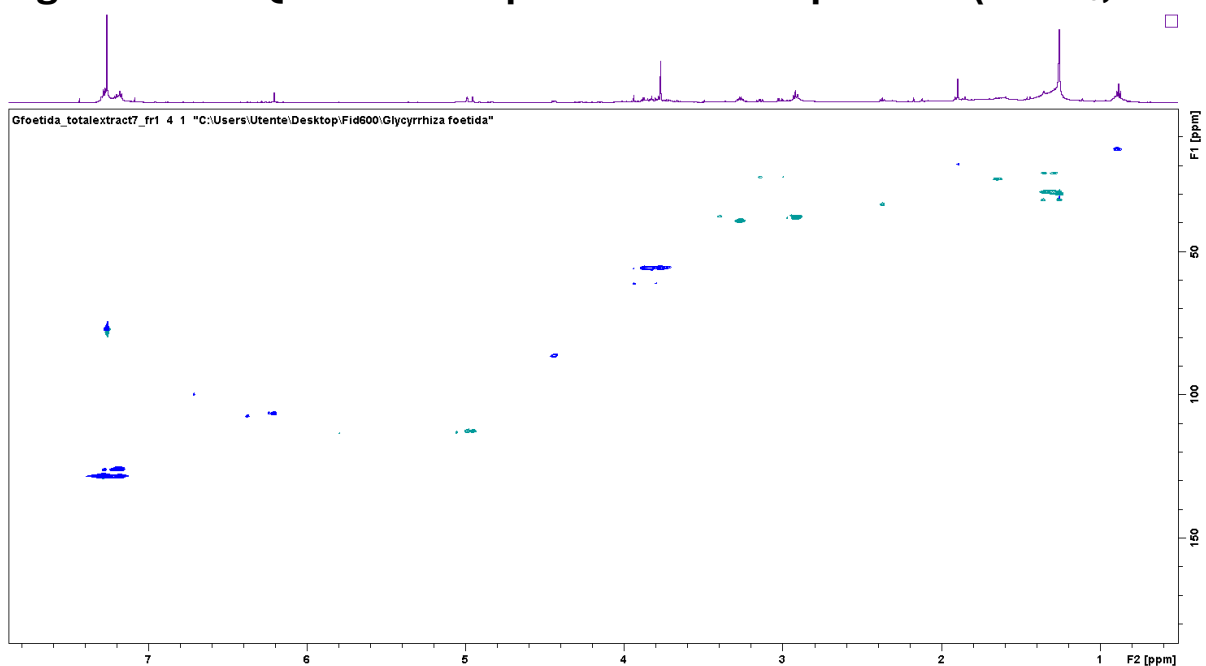
Figure S1: ^1H NMR spectrum of compound 4 (CDCl_3 , 600MHz)**Figure S2: HSQC 2D NMR spectrum of compound 4 (CDCl_3 , 600MHz)**

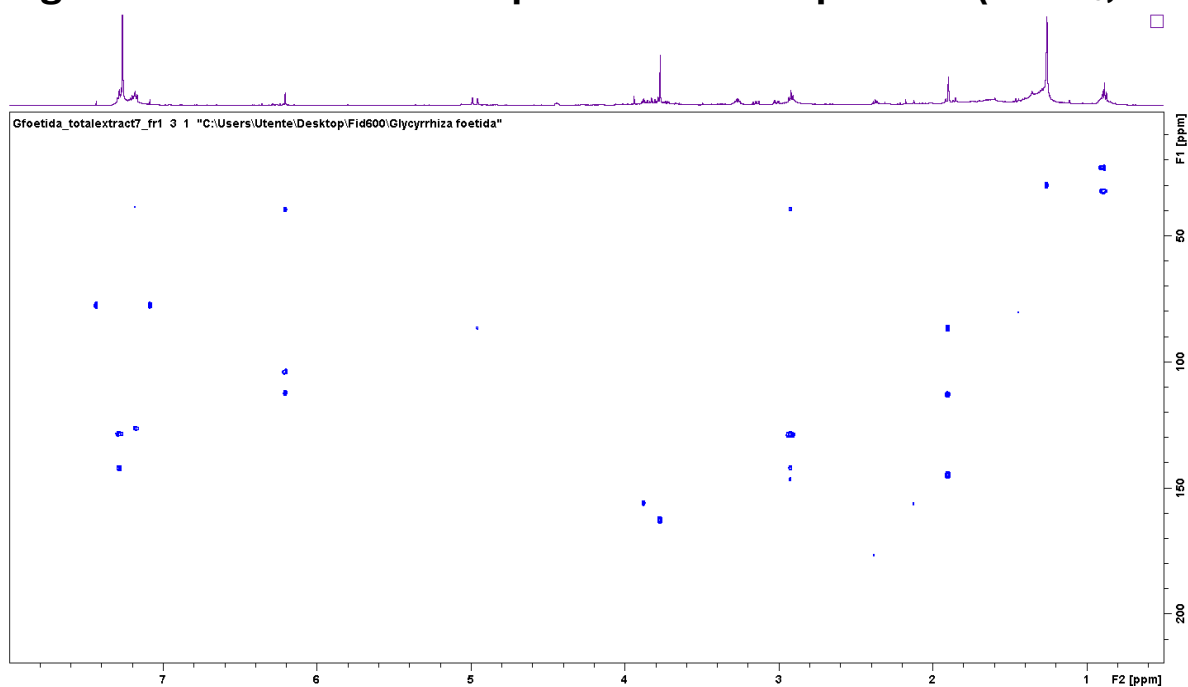
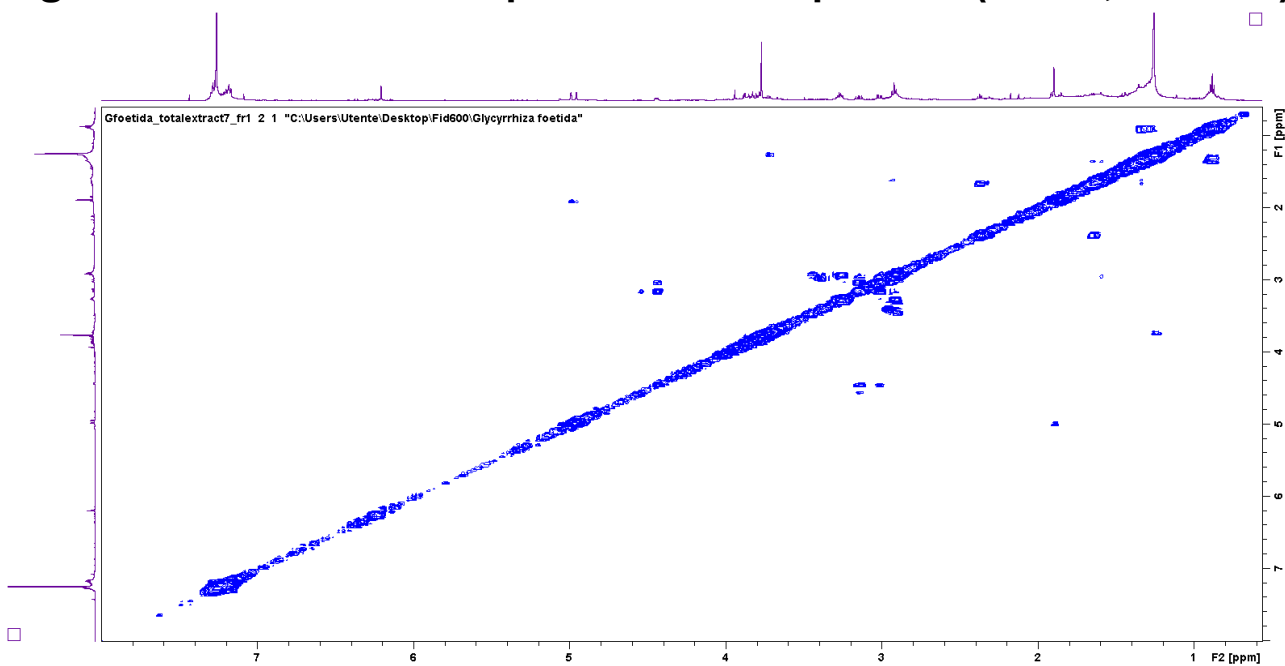
Figure S3: HMBC 2D NMR spectrum of compound 4 (CDCl₃, 600MHz)**Figure S4: COSY 2D NMR spectrum of compound 4 (CDCl₃, 600MHz)**

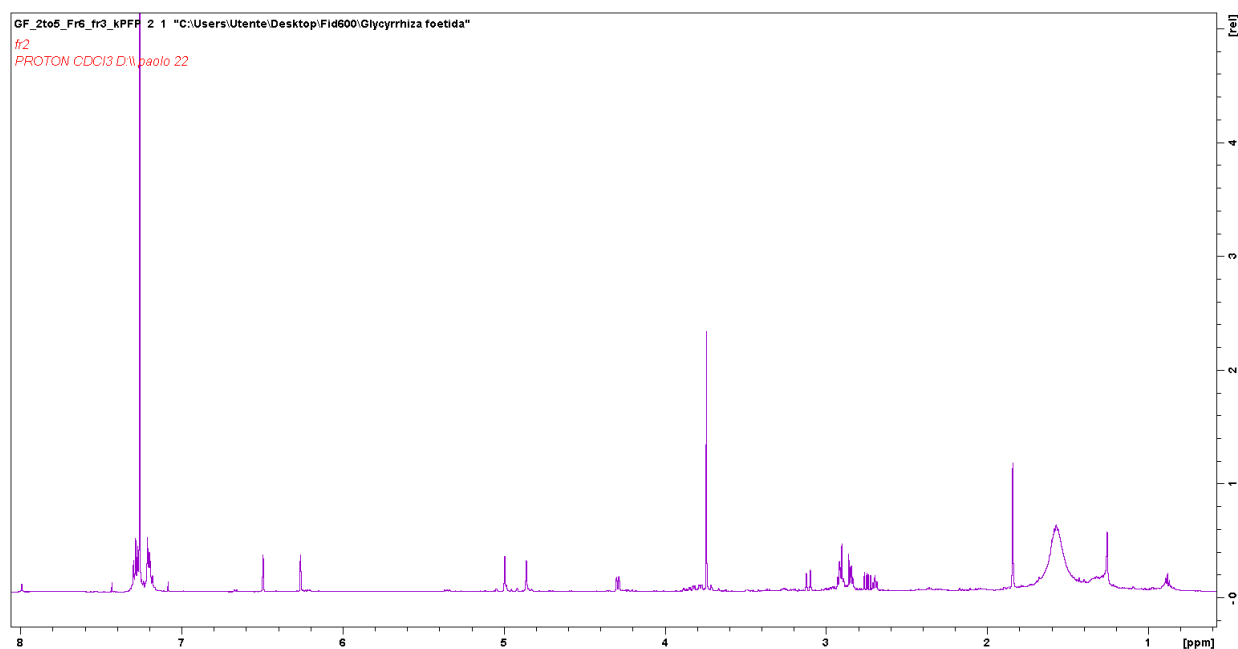
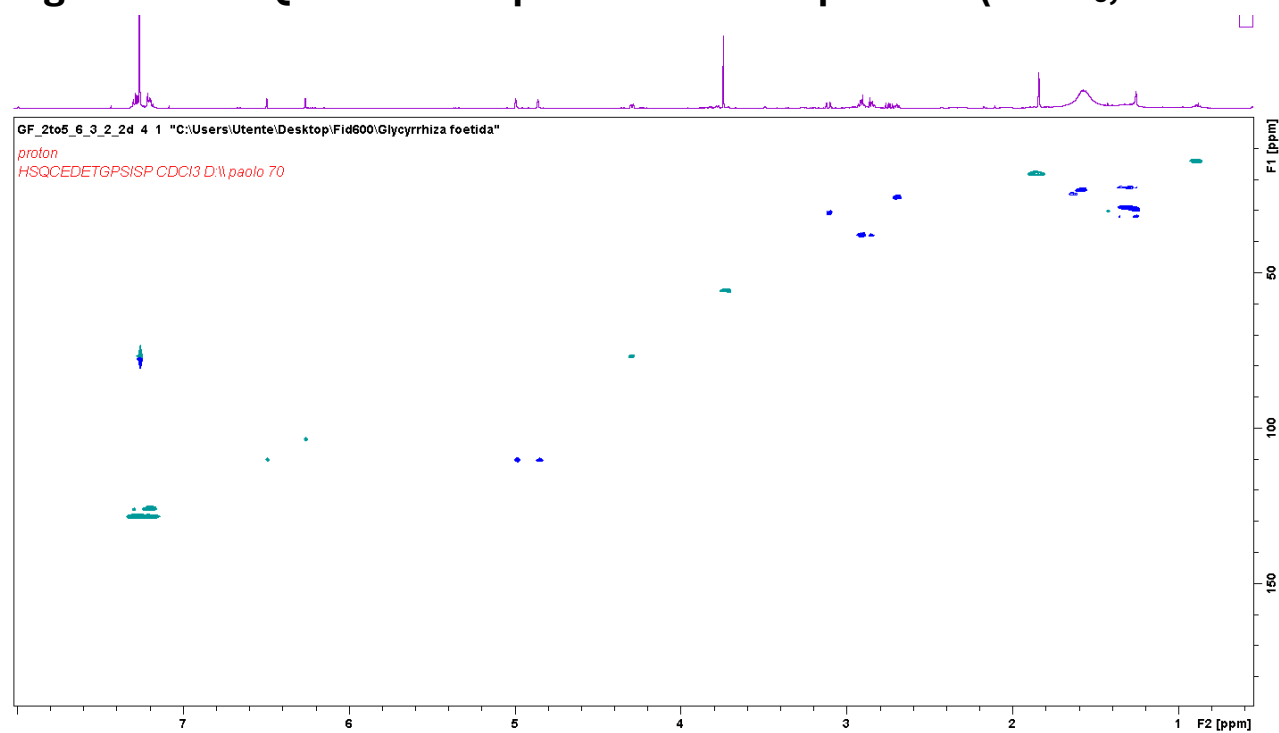
Figure S5: ^1H NMR spectrum of compound 5 (CDCl_3 , 600MHz)**Figure S6: HSQC 2D NMR spectrum of compound 5 (CDCl_3 , 600MHz)**

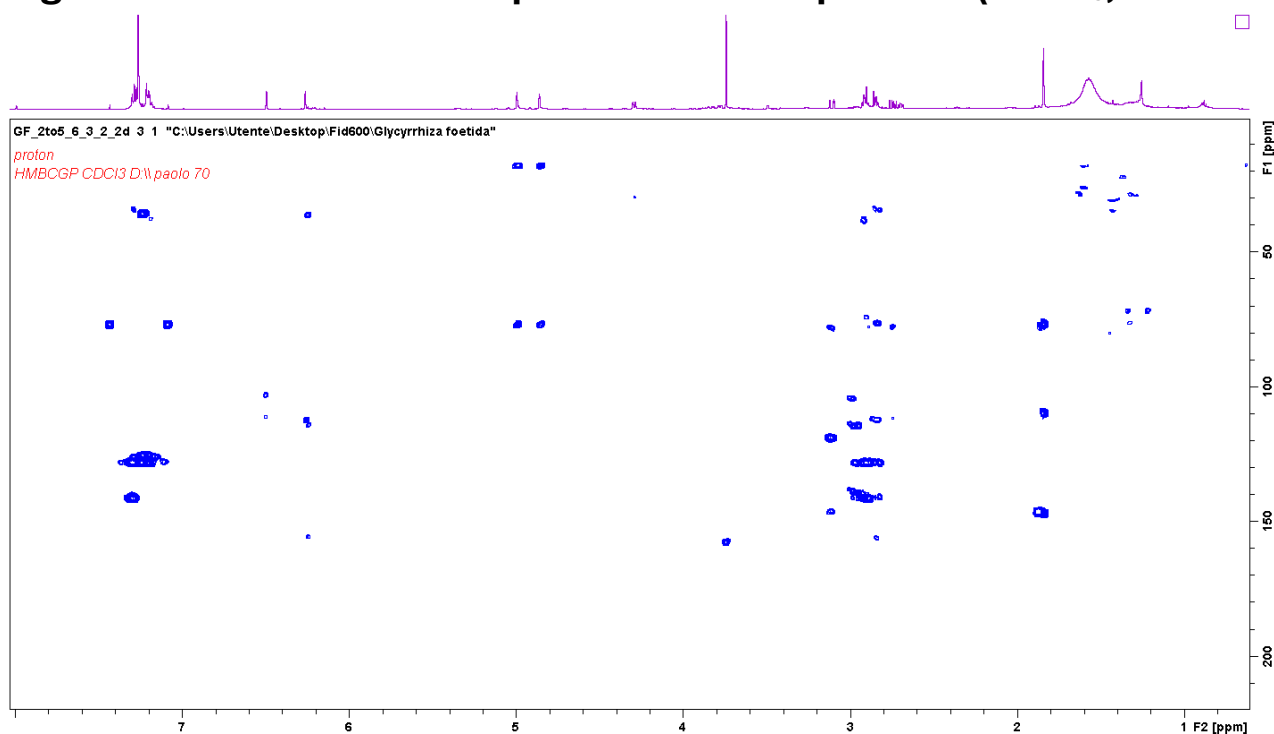
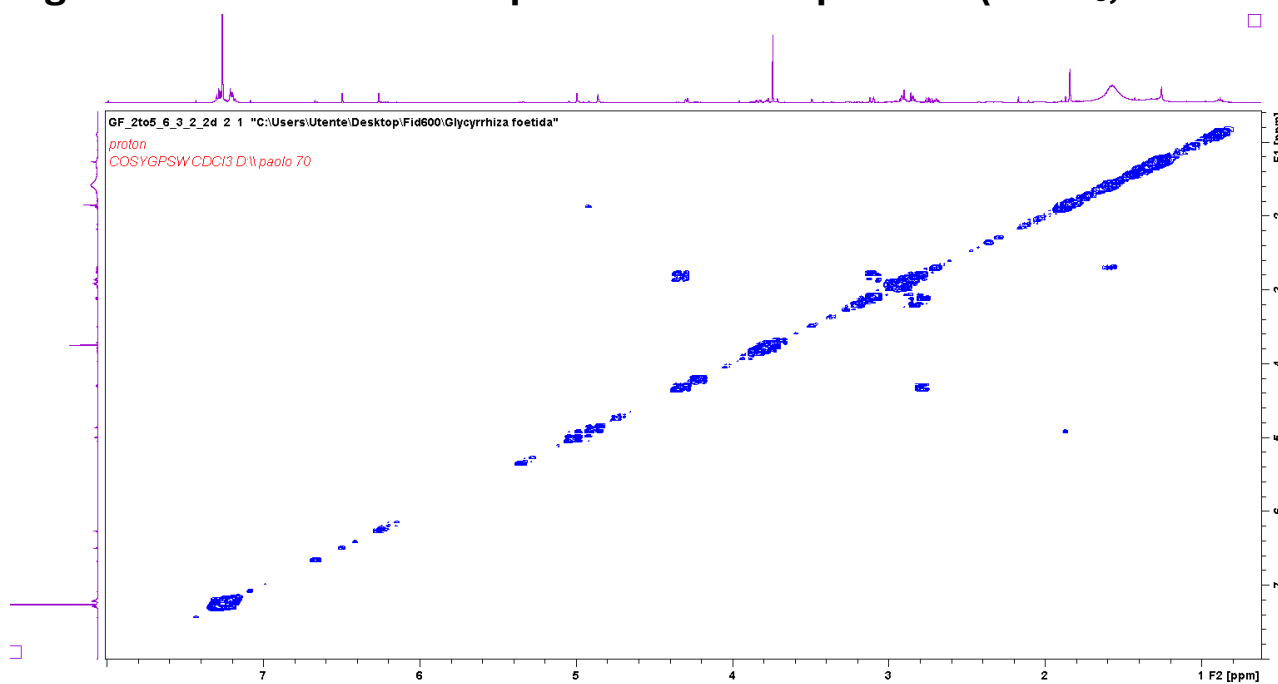
Figure S7: HMBC 2D NMR spectrum of compound 5 (CDCl₃, 600MHz)**Figure S8: COSY 2D NMR spectrum of compound 5 (CDCl₃, 600MHz)**

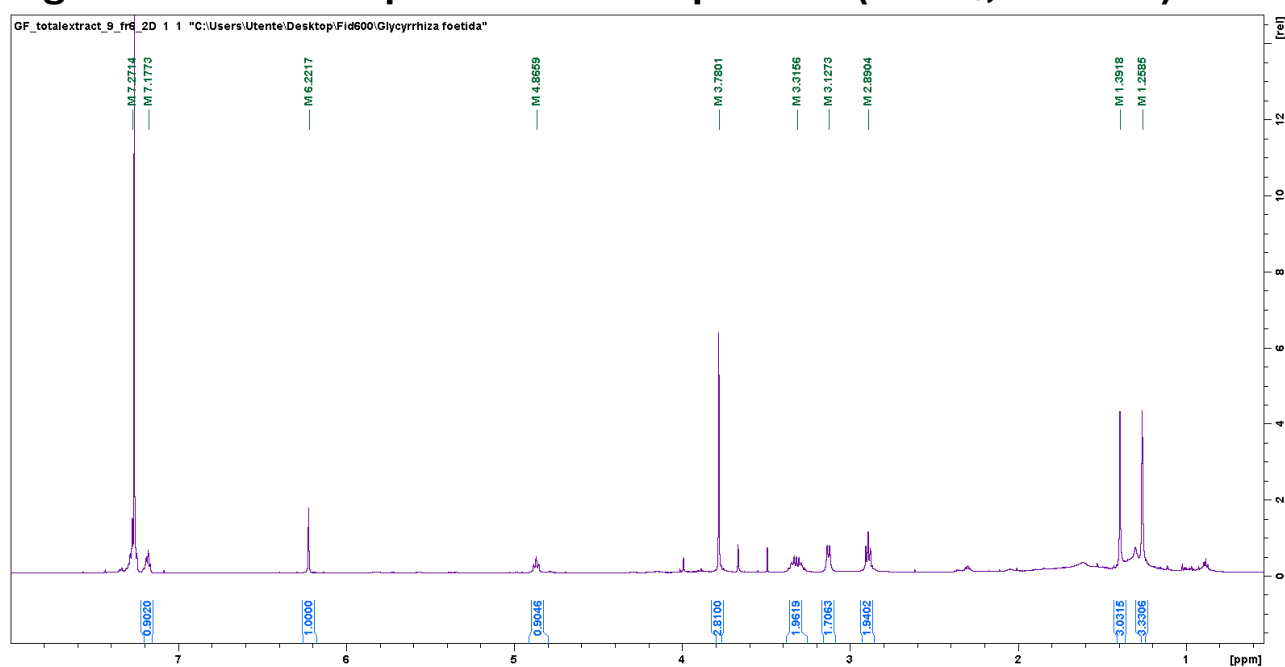
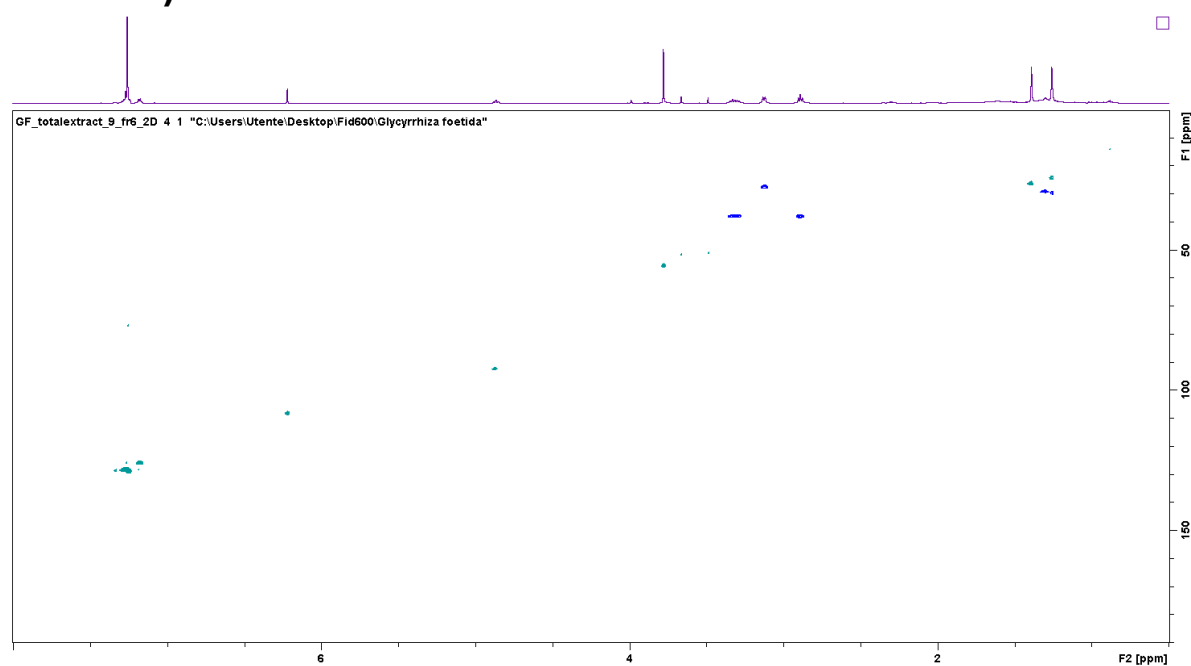
Figure S9: ^1H NMR spectrum of compound 6 (CDCl_3 , 600MHz)**Figure S10: HSQC 2D NMR spectrum of compound 6 (CDCl_3 , 600MHz)**

Figure S11: HMBC 2D NMR spectrum of compound 6 (CDCl₃, 600MHz)

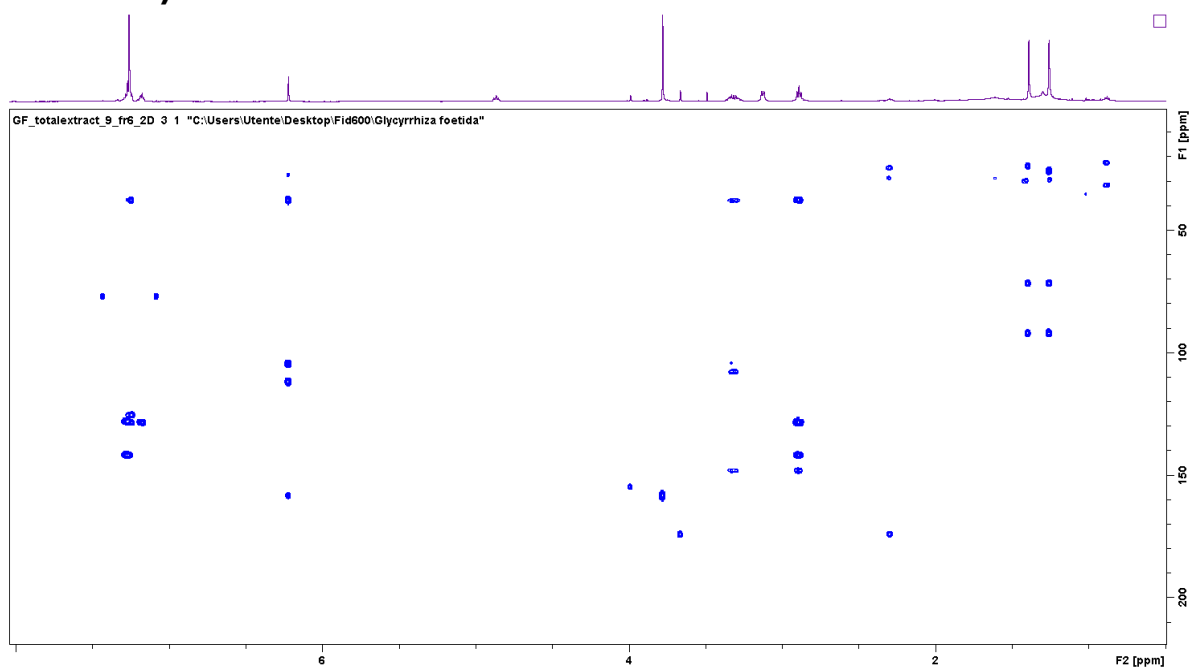


Figure S12: COSY 2D NMR spectrum of compound 6 (CDCl₃, 600MHz)

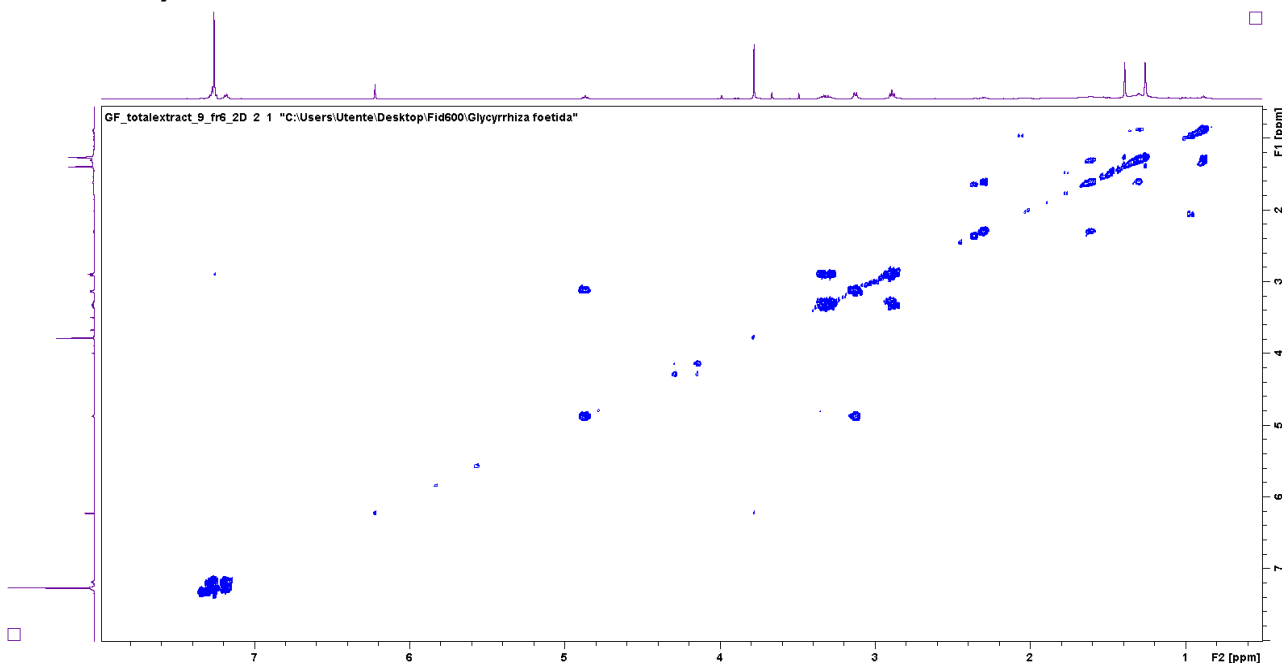


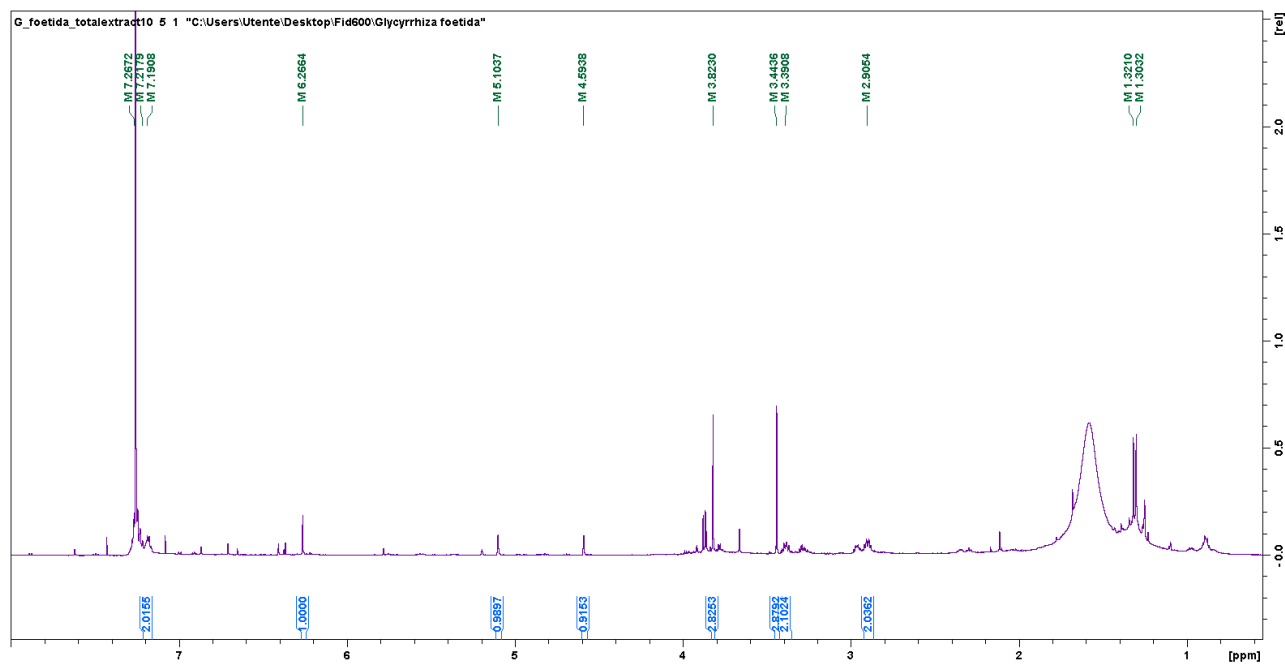
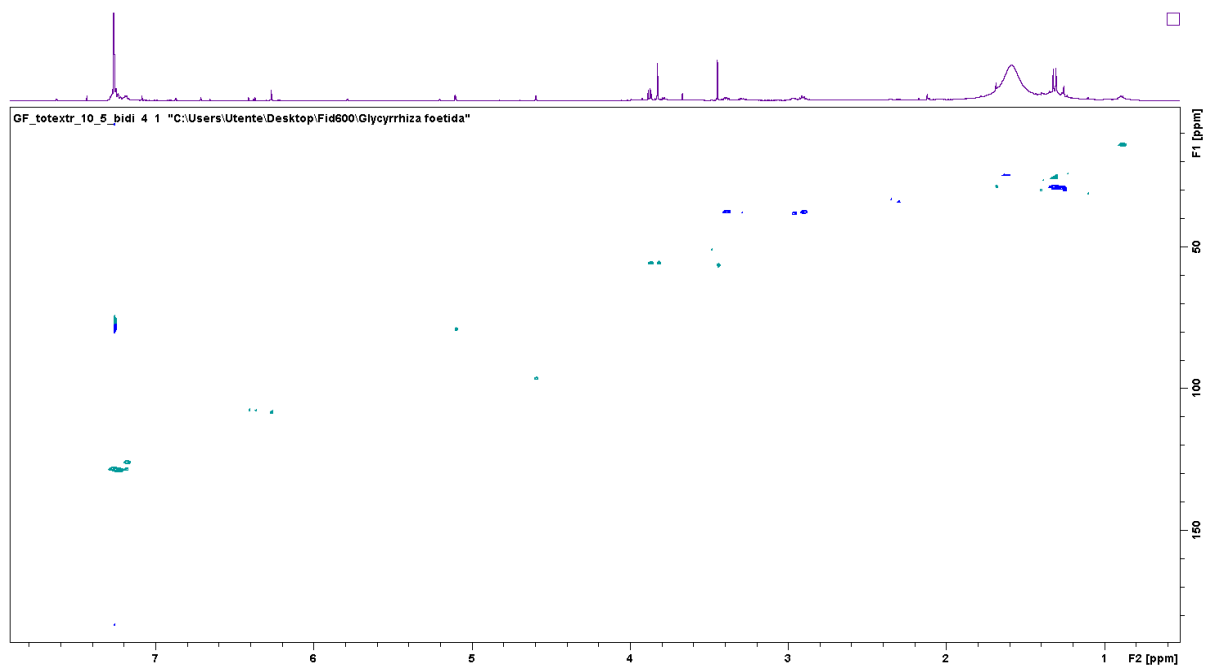
Figure S13: ^1H NMR spectrum of compound 7 (CDCl_3 , 600MHz)**Figure S14: HSQC 2D NMR spectrum of compound 7 (CDCl_3 , 600MHz)**

Figure S15: HMBC 2D NMR spectrum of compound 7 (CDCl₃, 600MHz)

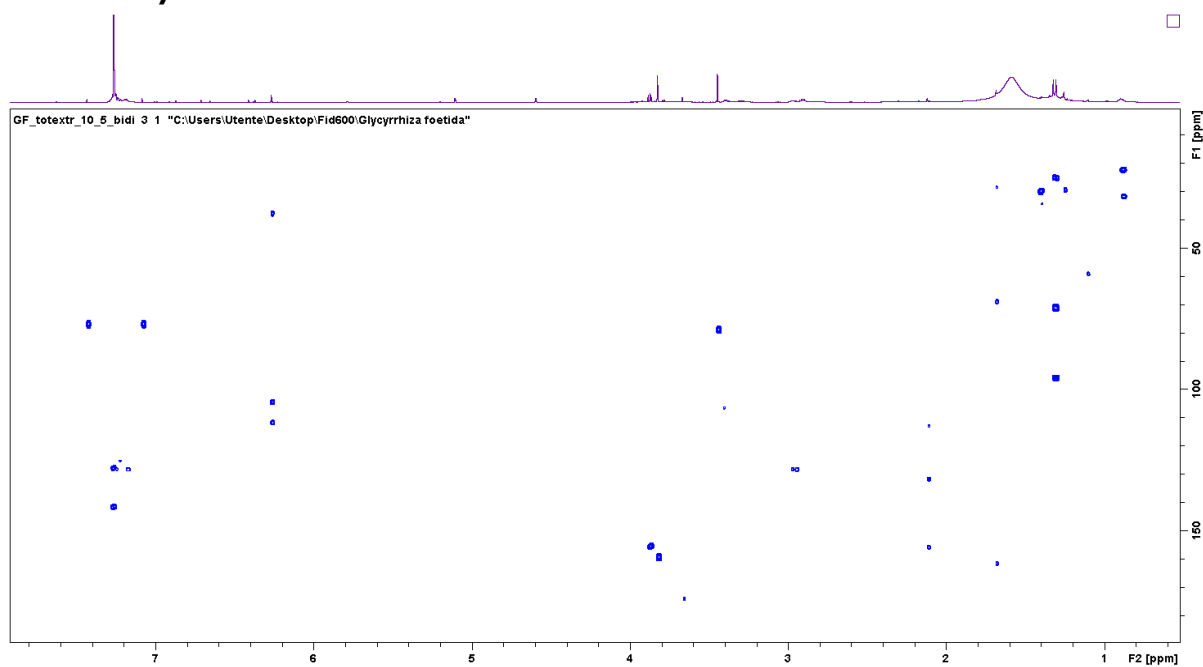
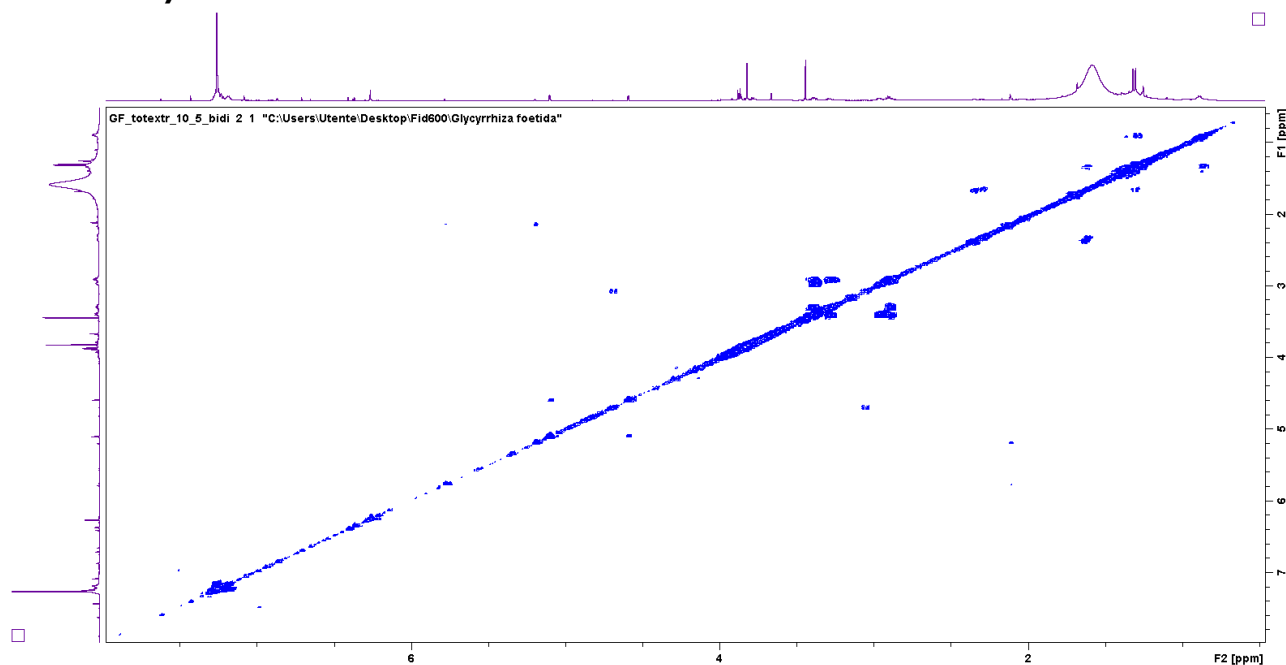


Figure S16: COSY 2D NMR spectrum of compound 7 (CDCl₃, 600MHz)



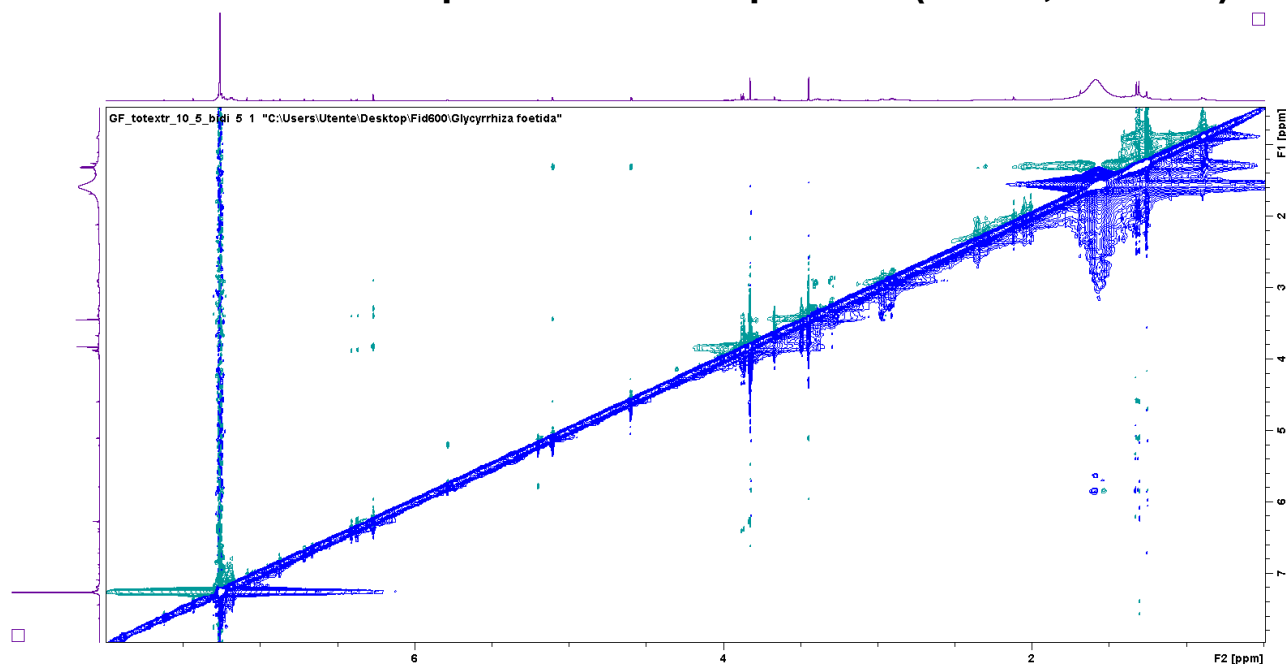
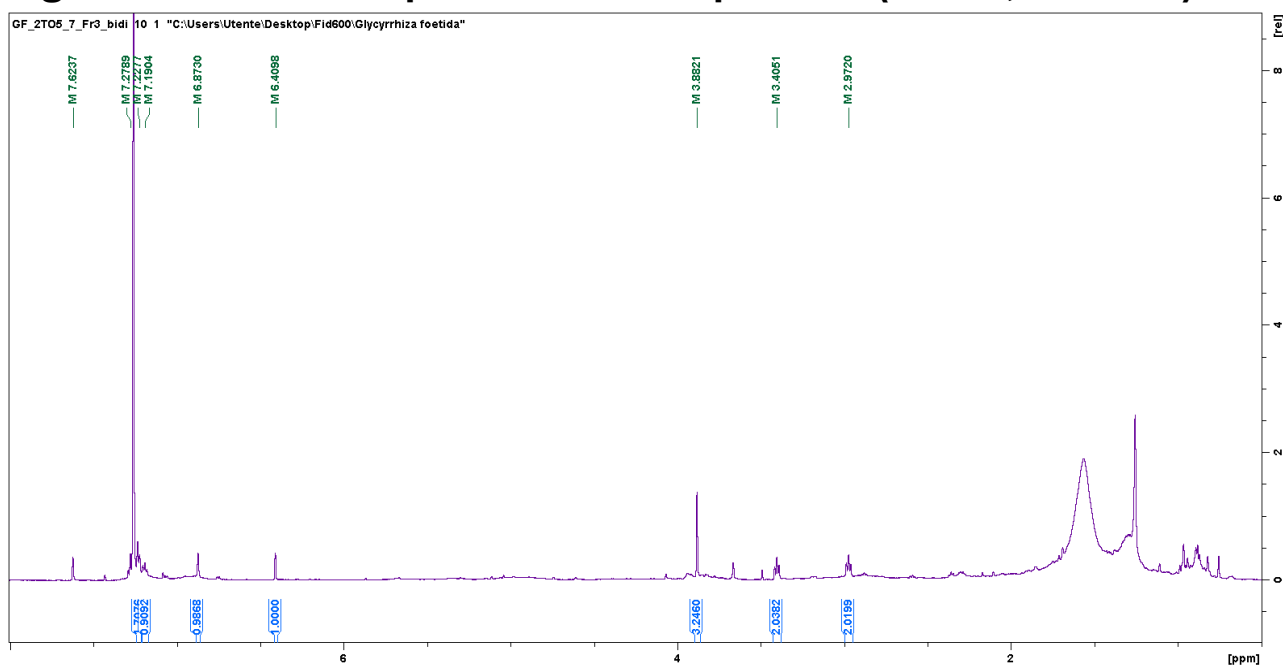
S17: NOESY 2D NMR spectrum of compound 7 (CDCl₃, 600MHz)**Figure S18: ¹H NMR spectrum of compound 8 (CDCl₃, 600MHz)**

Figure S19: HSQC 2D NMR spectrum of compound 8 (CDCl₃, 600MHz)

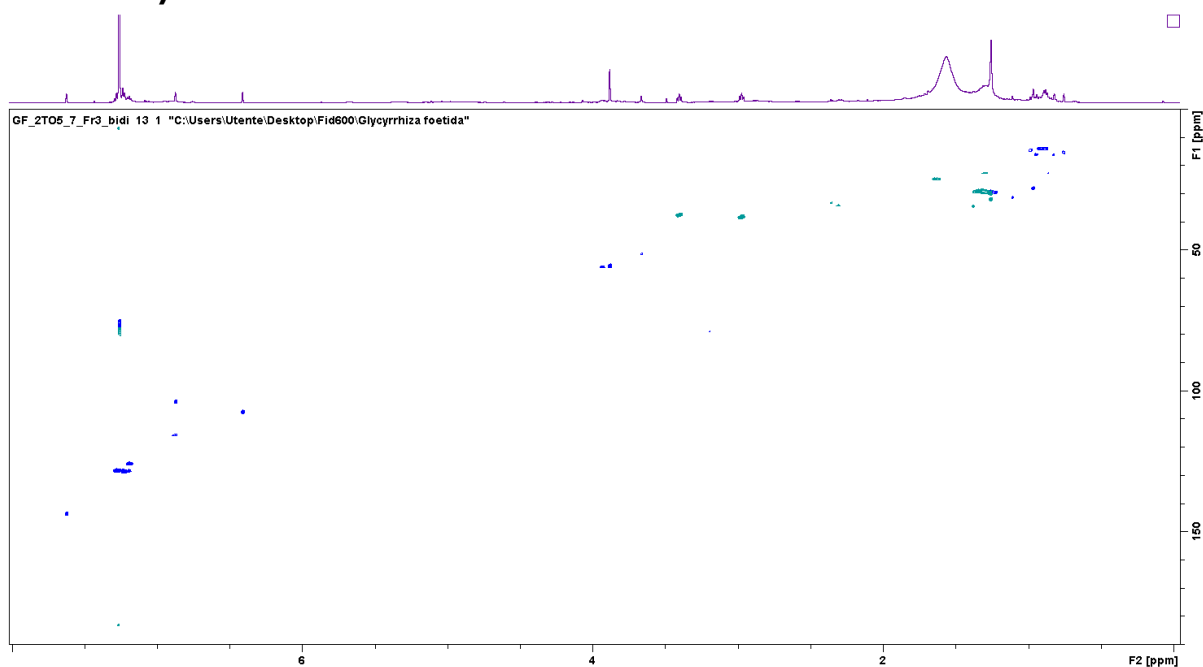


Figure S20: HMBC 2D NMR spectrum of compound 8 (CDCl₃, 600MHz)

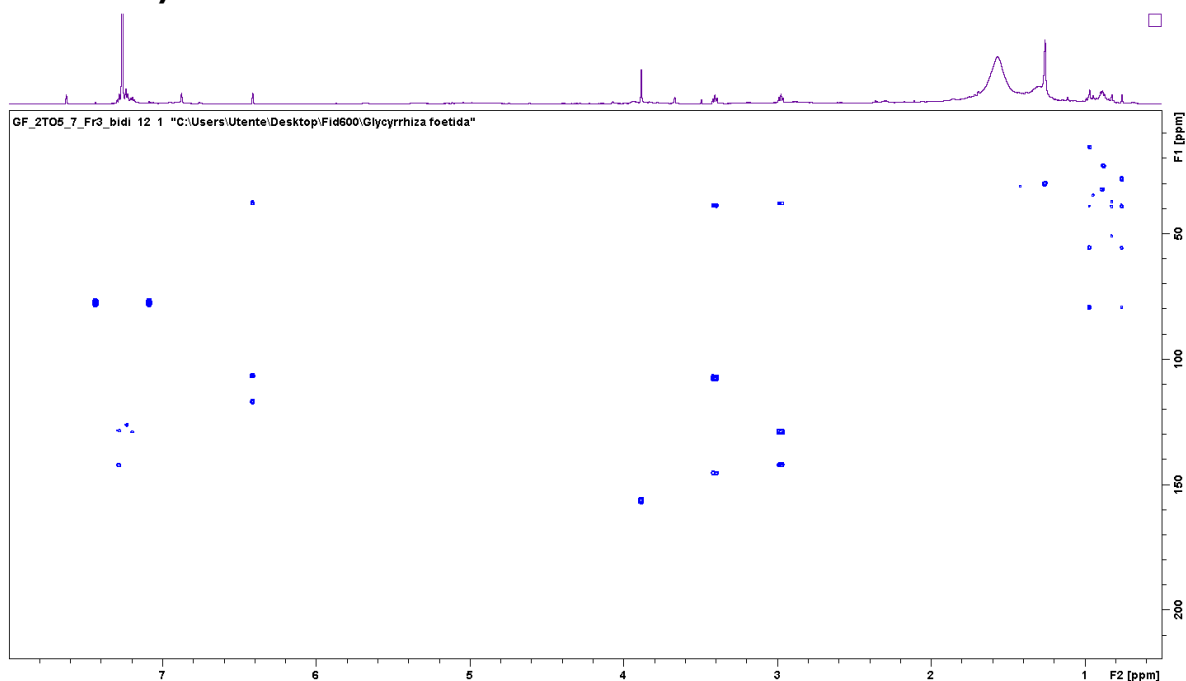


Figure S21: COSY 2D NMR spectrum of compound 8 (CDCl₃, 600MHz)

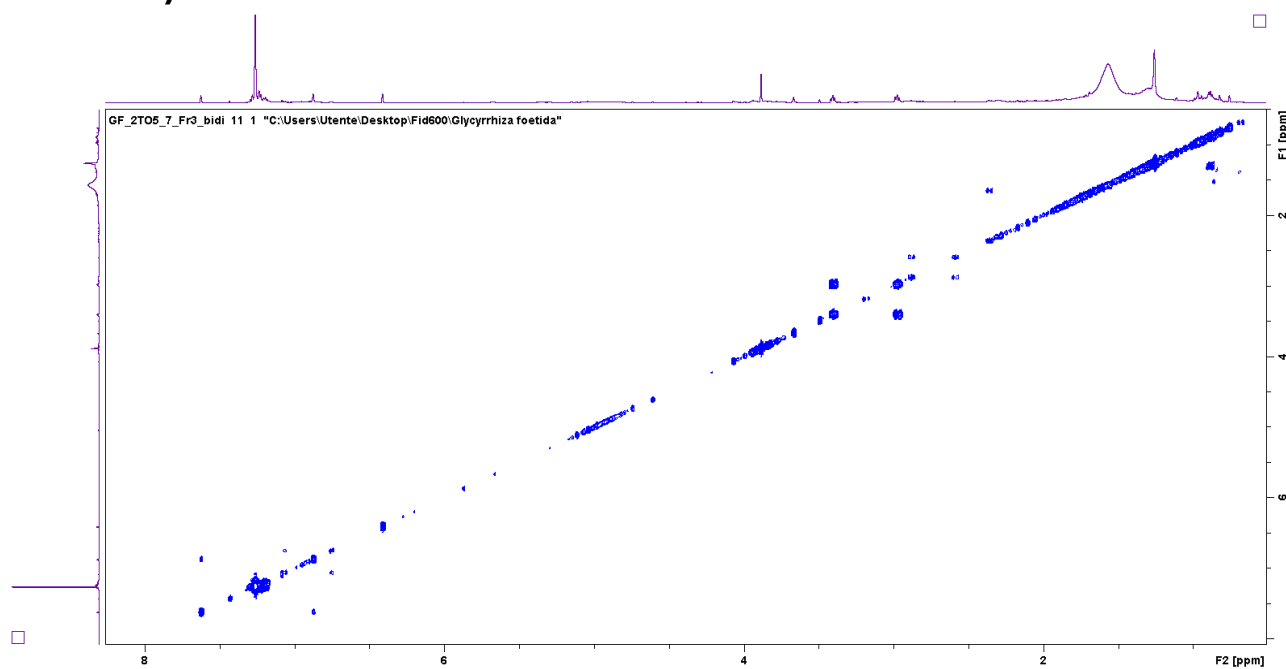


Figure S22: ¹H NMR spectrum of compound 9 (CDCl₃, 600MHz)

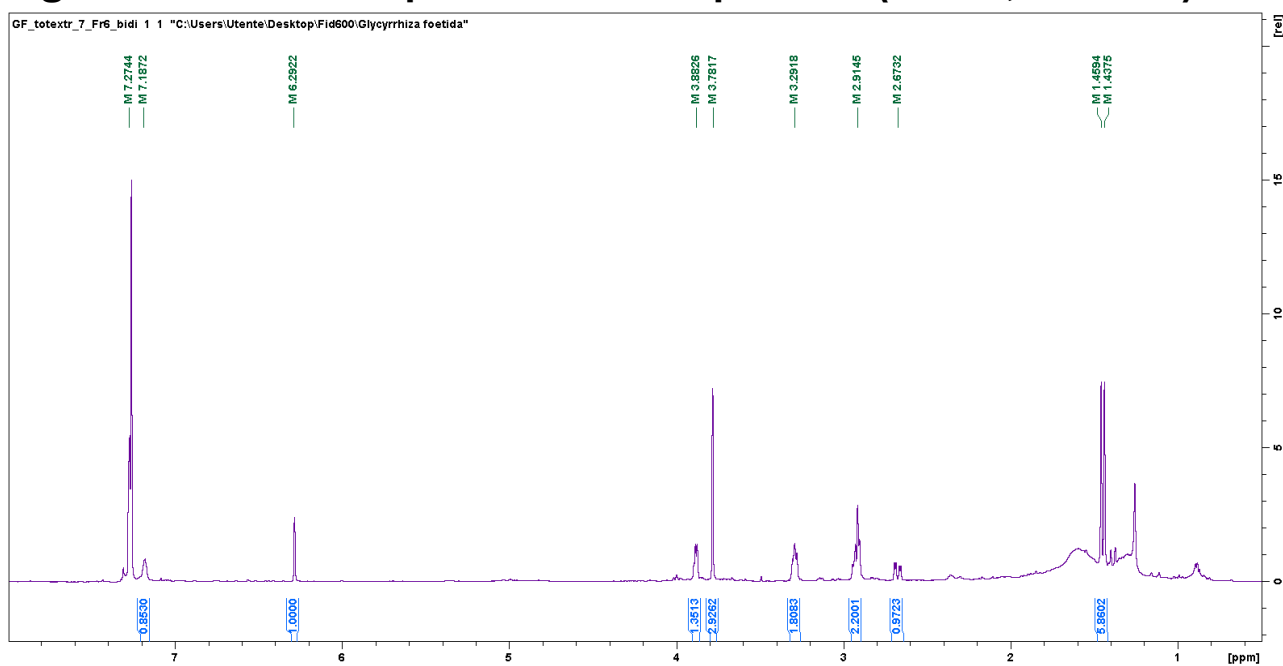


Figure S23: HSQC 2D NMR spectrum of compound 9 (CDCl₃, 600MHz)

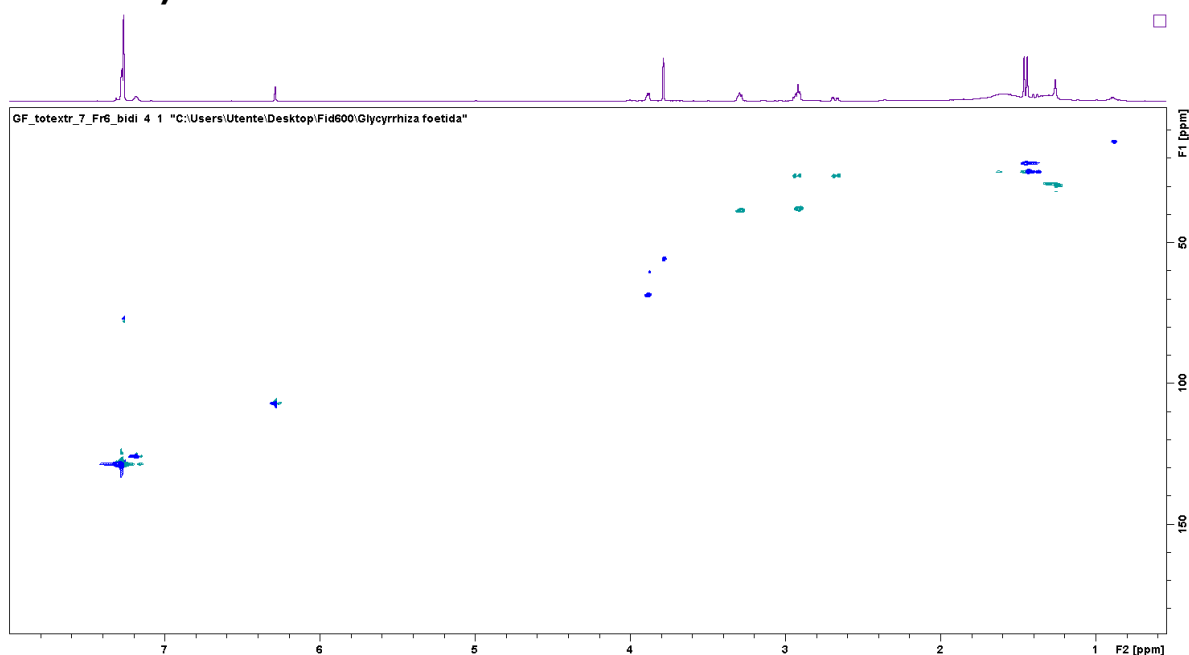


Figure S24: HMBC 2D NMR spectrum of compound 9 (CDCl₃, 700MHz)

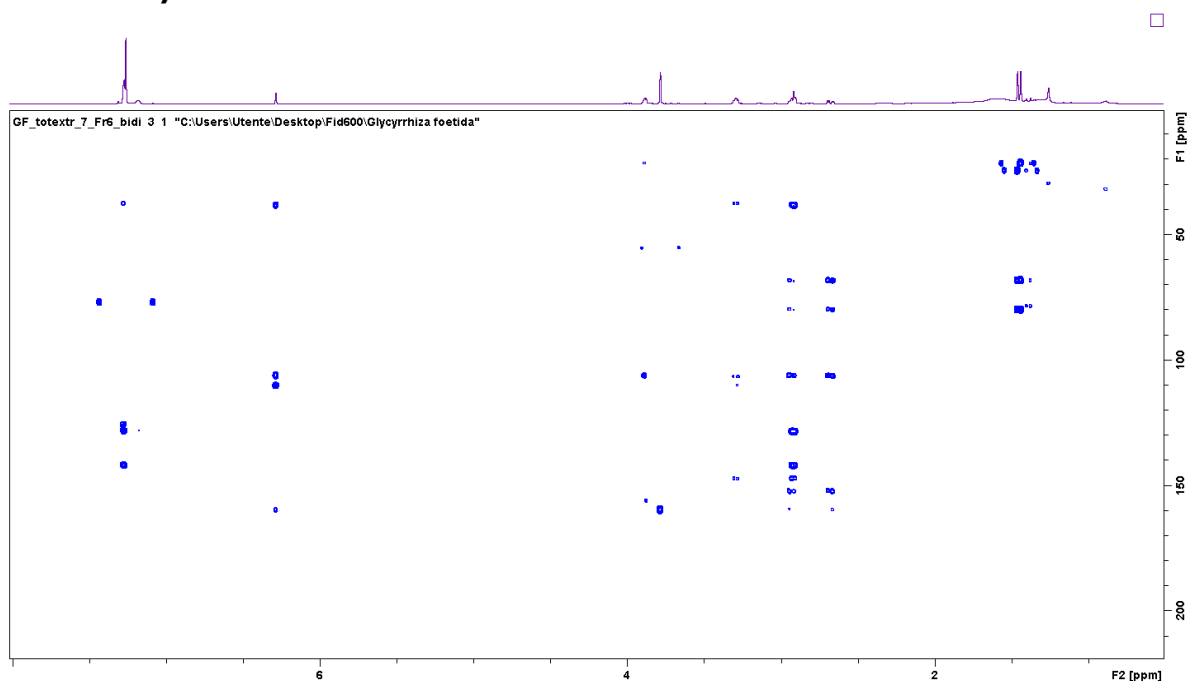


Figure S25: COSY 2D NMR spectrum of compound 9 (CDCl₃, 700MHz)

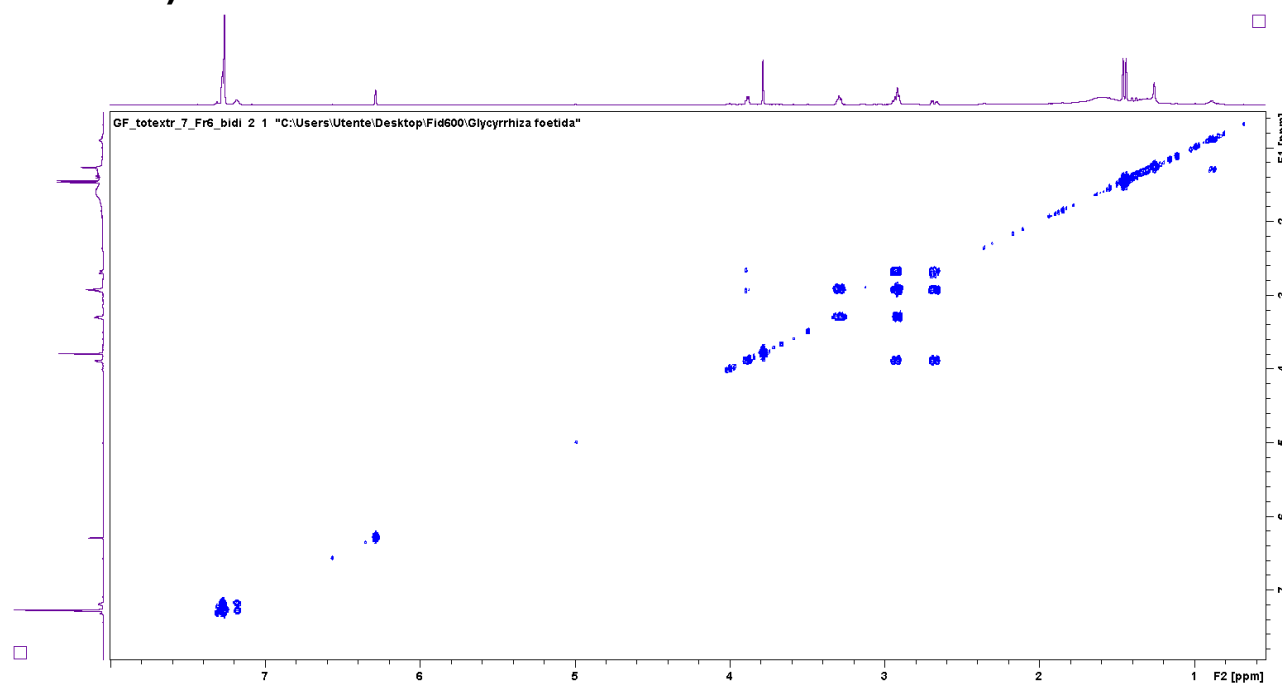


Figure S26: ¹H NMR spectrum of compound 10 (CDCl₃, 600MHz)

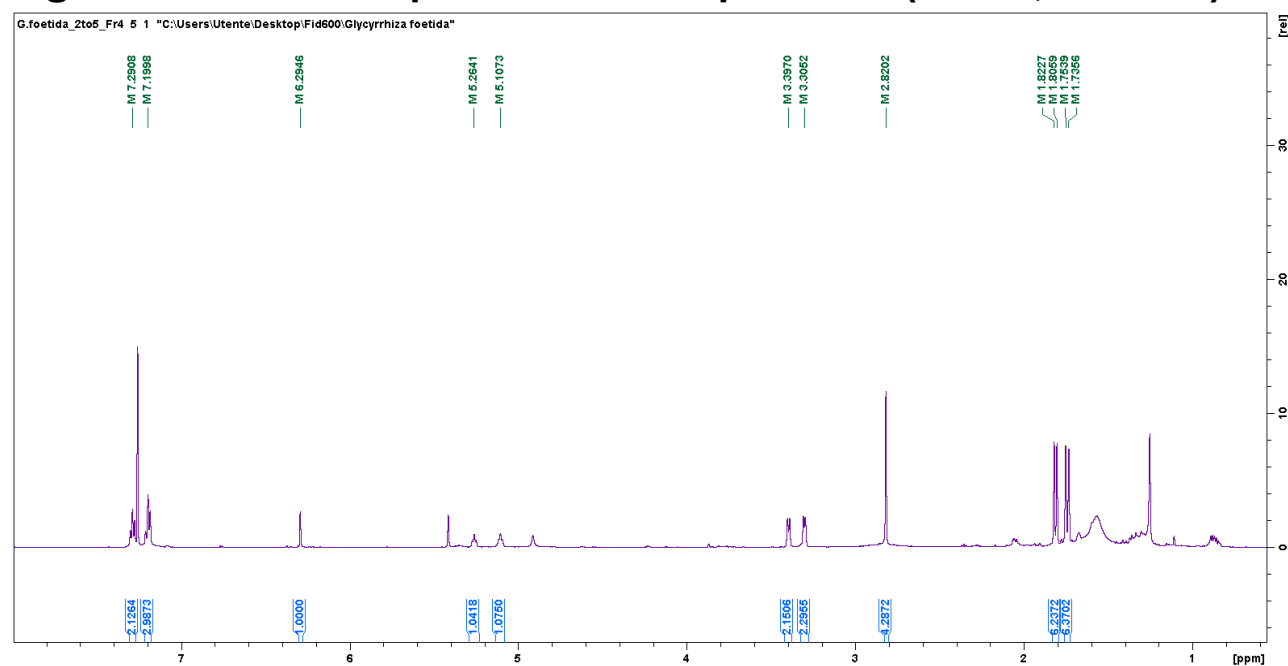


Figure S27: HSQC 2D NMR spectrum of compound 10 (CDCl₃, 600MHz)

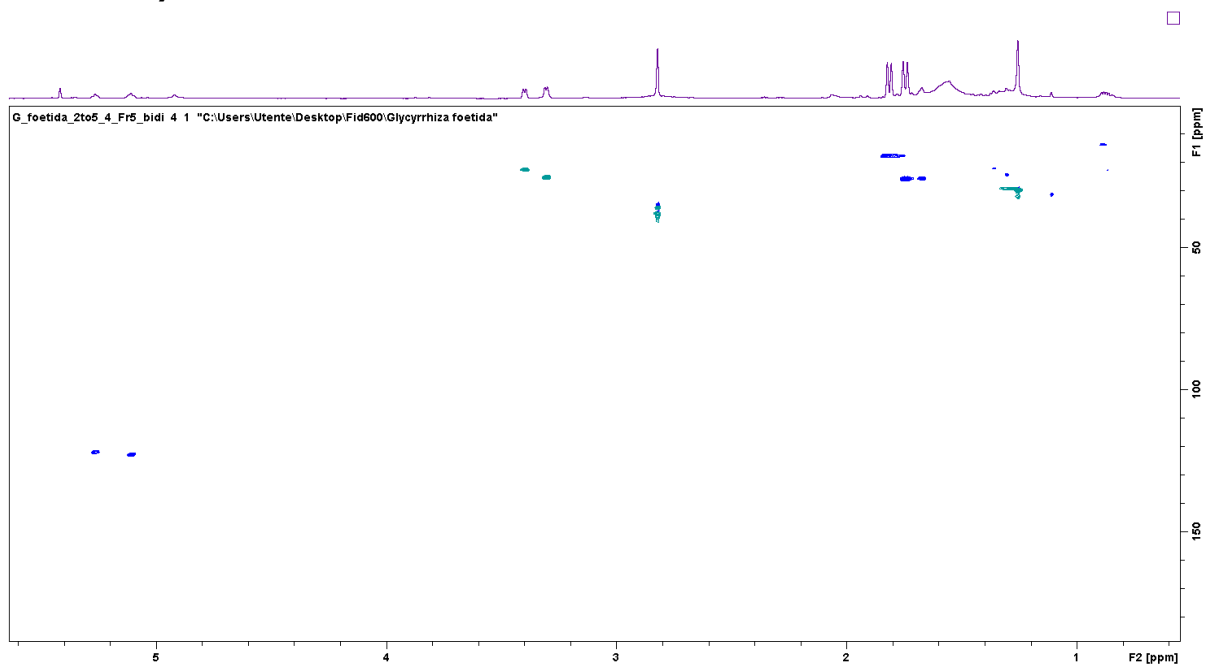


Figure S28: HMBC 2D NMR spectrum of compound 10 (CDCl₃, 600MHz)

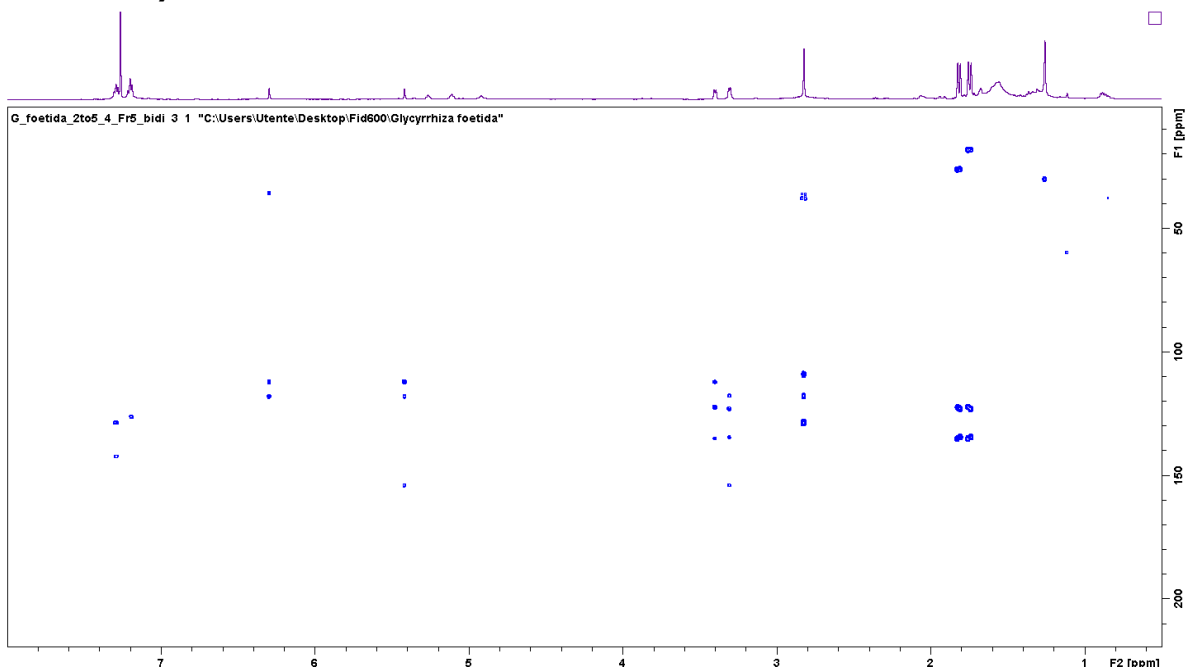


Figure S29: COSY 2D NMR spectrum of compound 10 (CDCl₃, 600MHz)

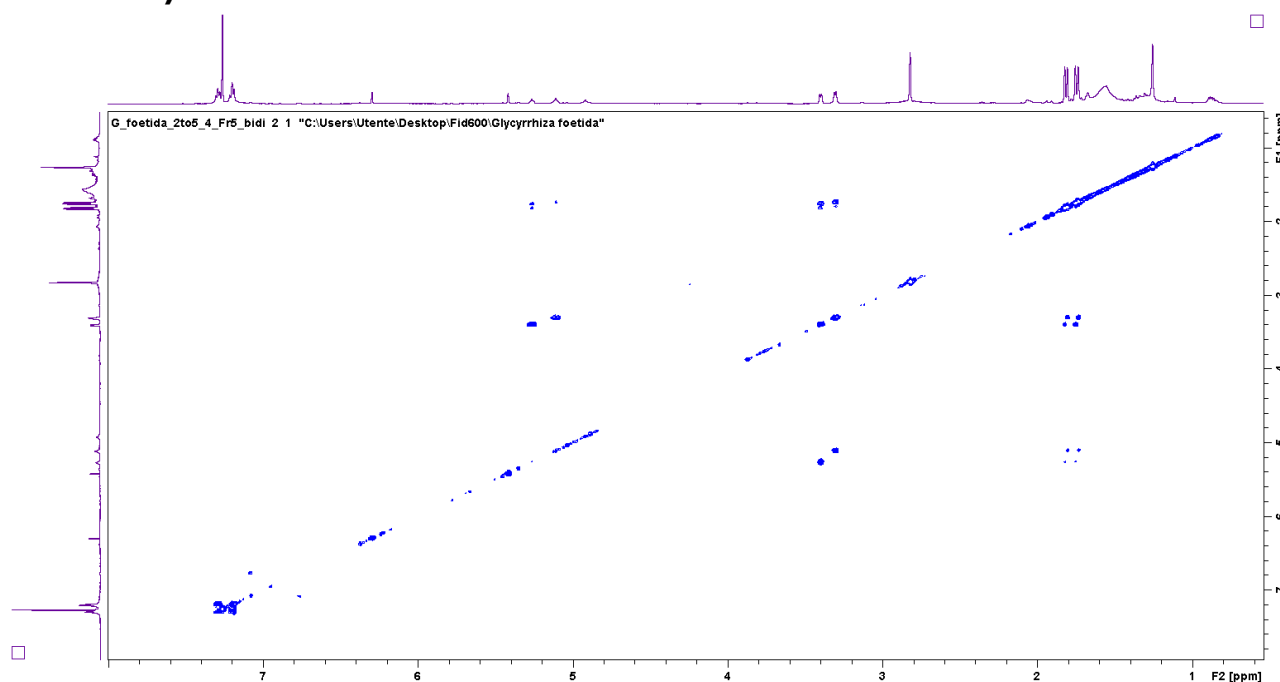


Figure S30: ¹H NMR spectrum of compound 14 (CDCl₃, 600MHz)

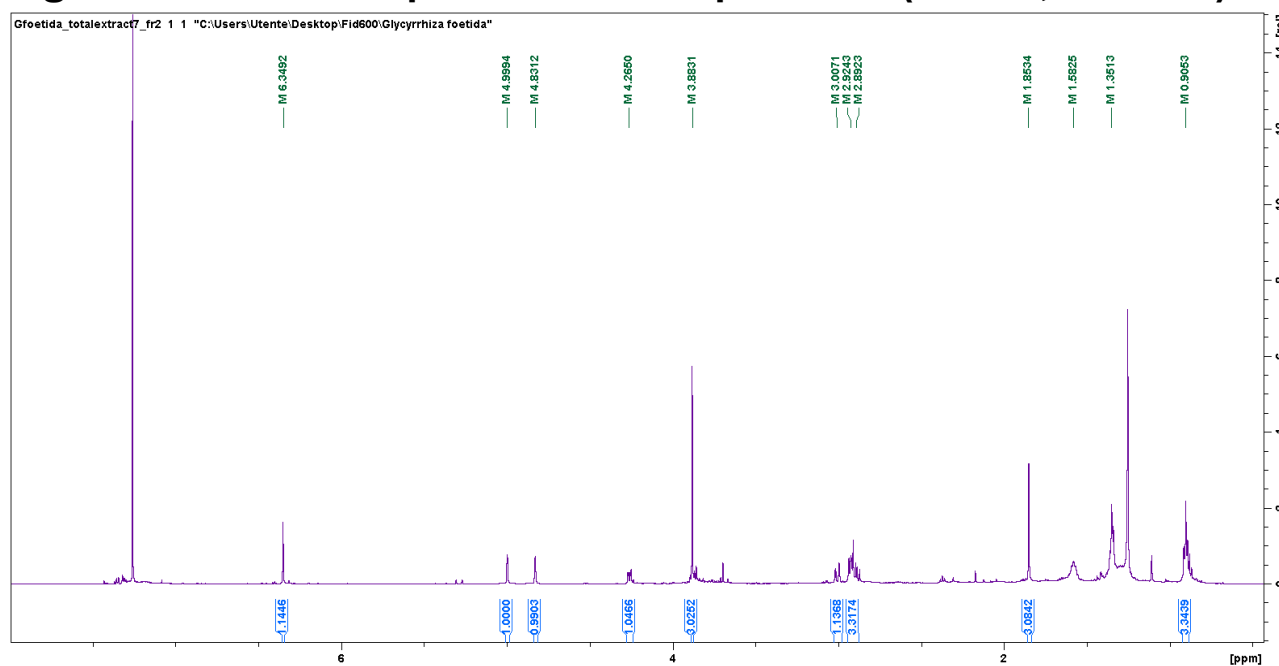


Figure S31: HSQC 2D NMR spectrum of compound 14 (CDCl₃, 600MHz)

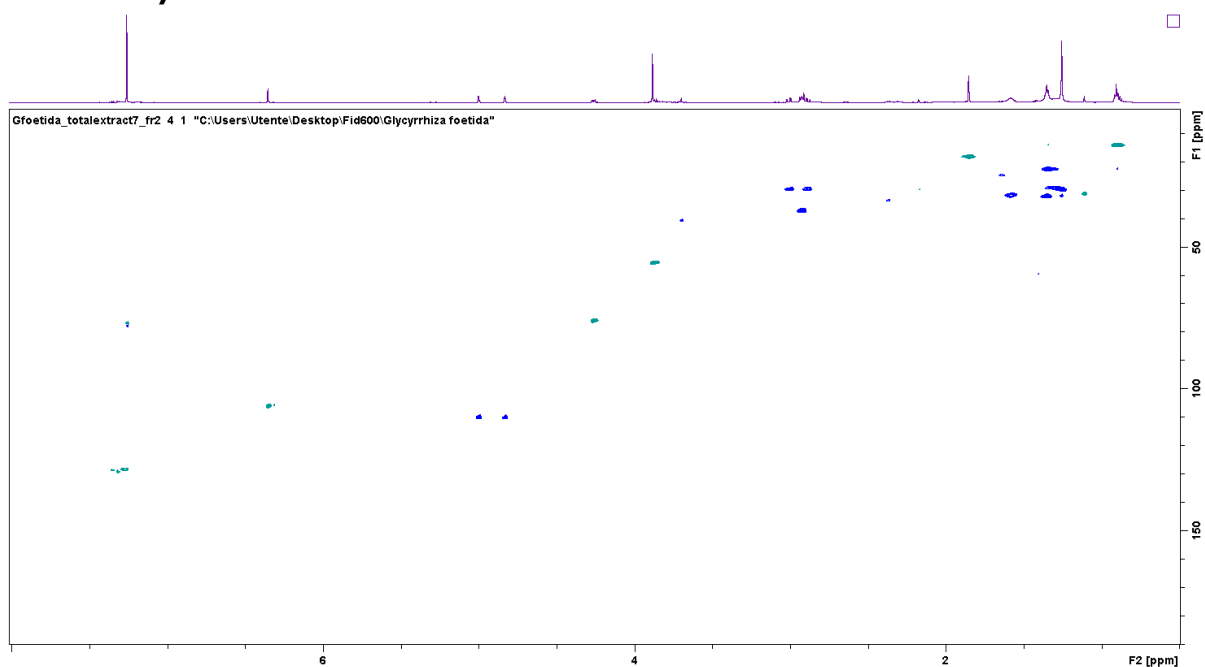


Figure S32: HMBC 2D NMR spectrum of compound 14 (CDCl₃, 600MHz)

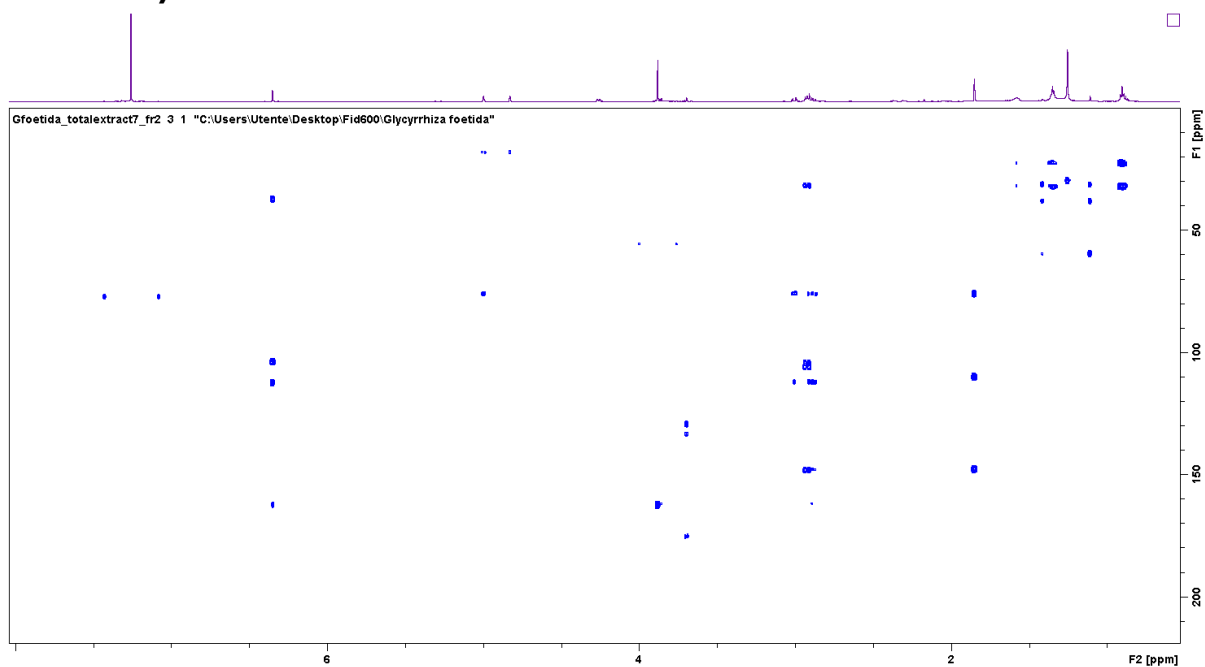


Figure S33: COSY 2D NMR spectrum of compound 14 (CDCl₃, 600MHz)

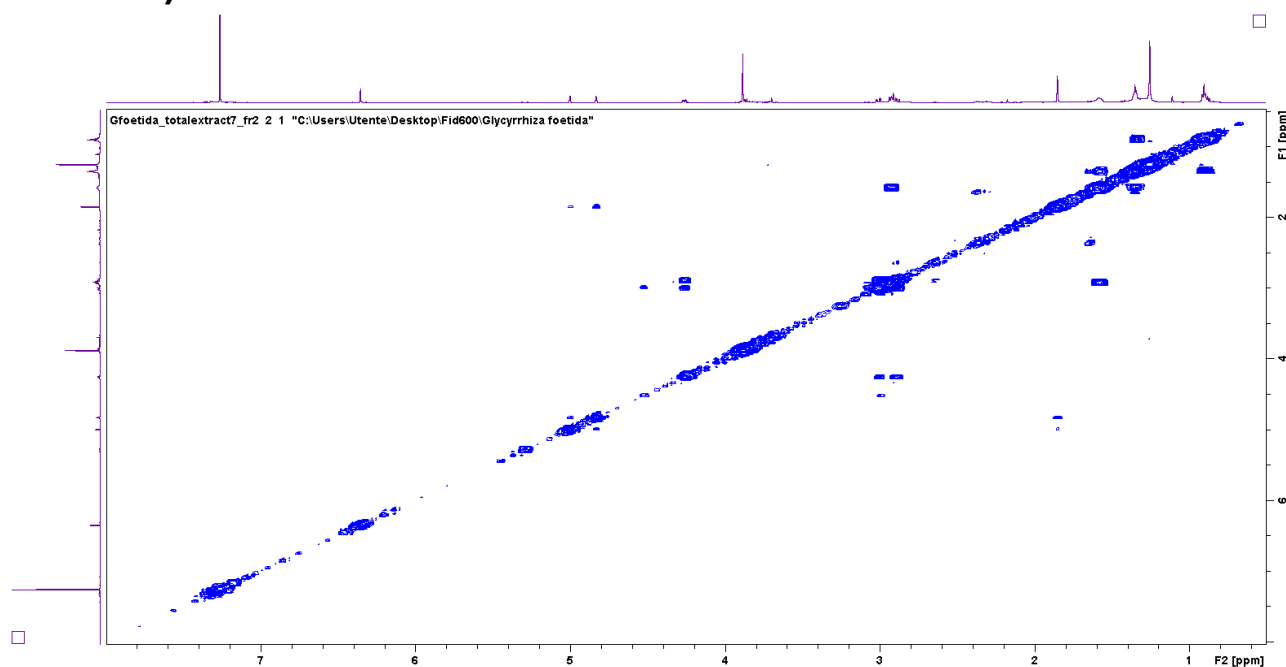


Figure S34: ¹H NMR spectrum of compound 15 (CDCl₃, 700MHz)

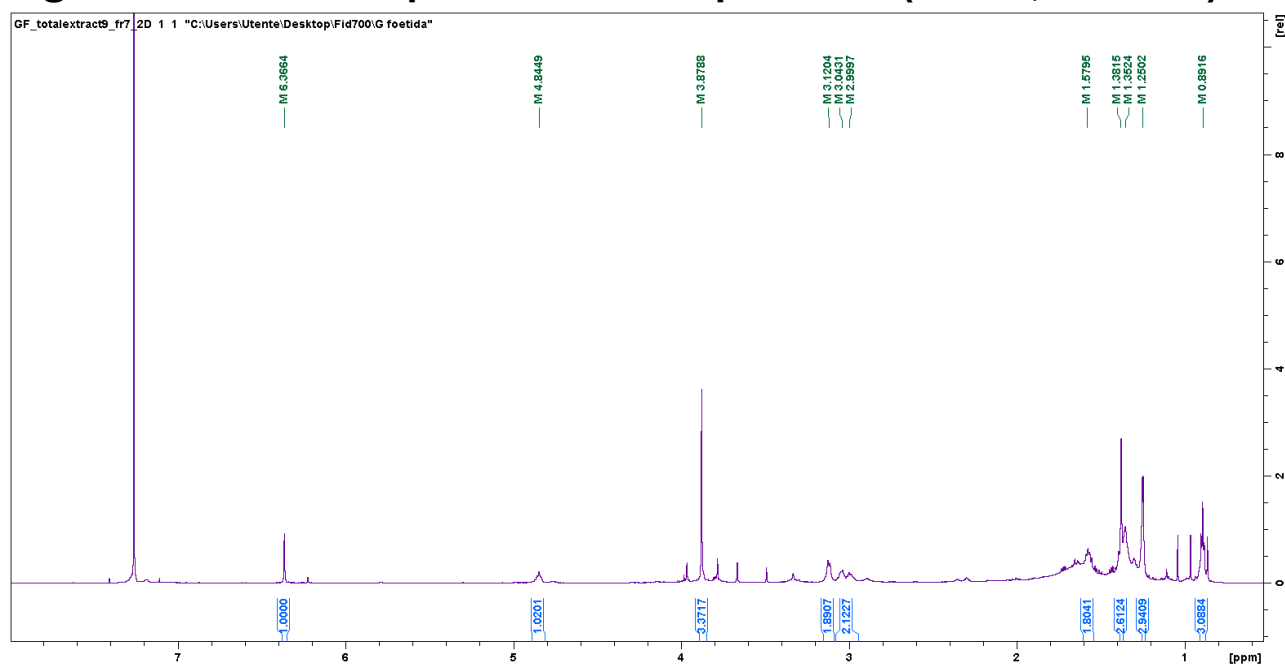


Figure S35: HSQC 2D NMR spectrum of compound 15 (CDCl₃, 700MHz)

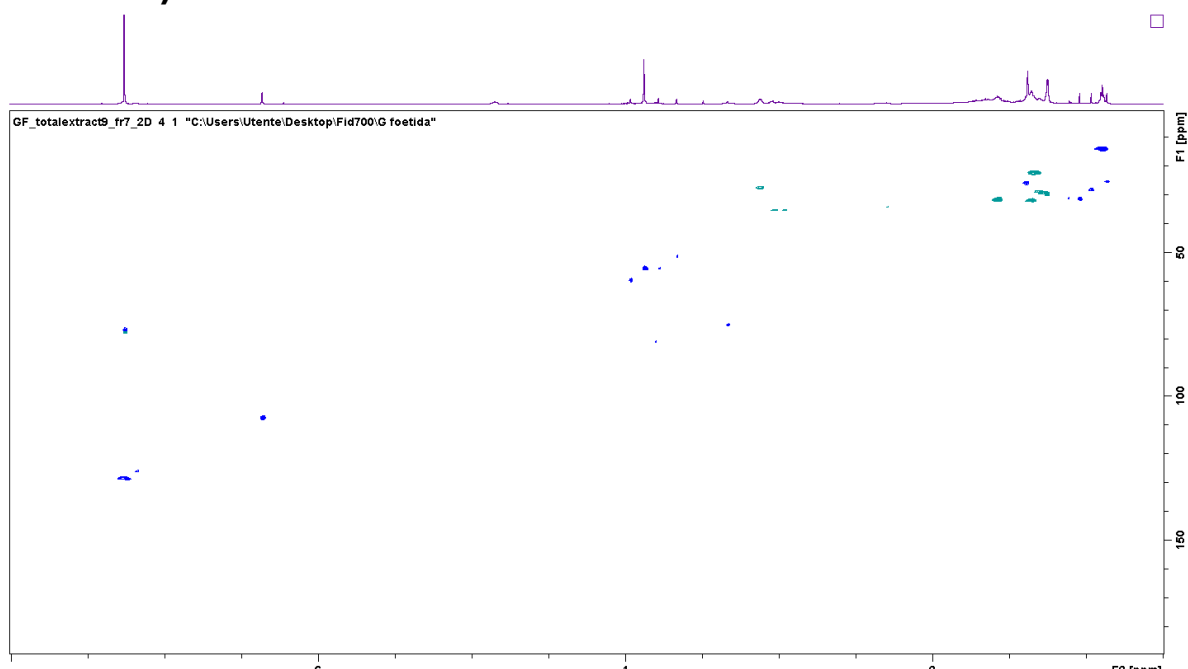


Figure S36: HMBC 2D NMR spectrum of compound 15 (CDCl₃, 700MHz)

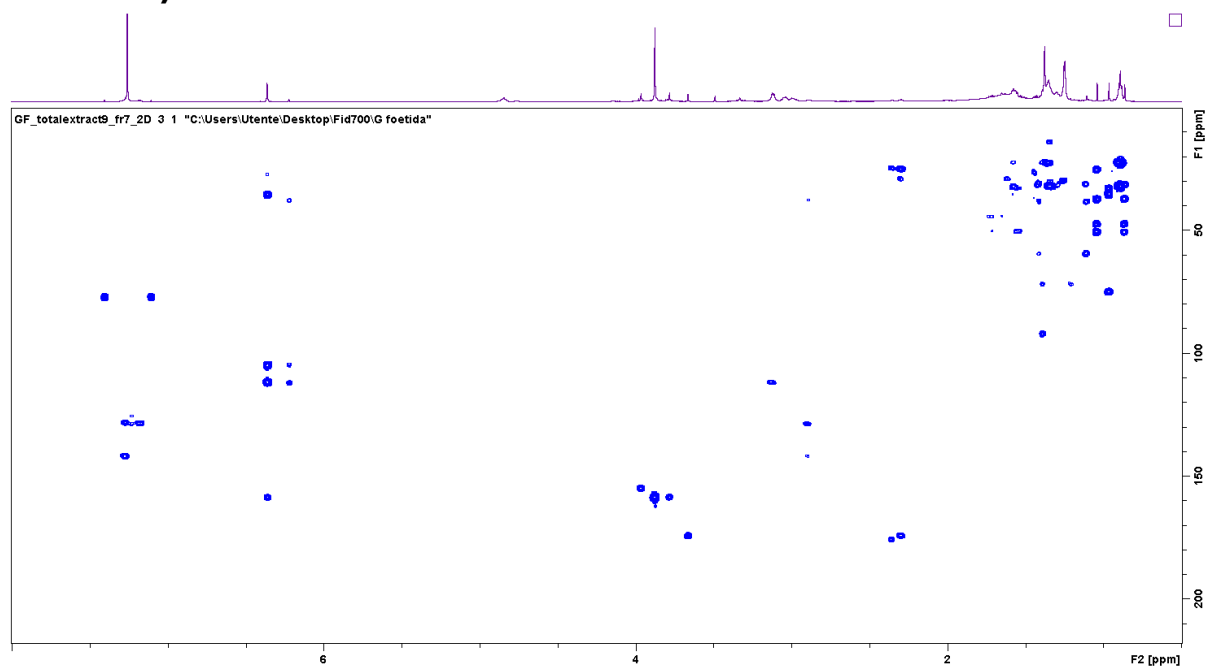


Figure S37: COSY 2D NMR spectrum of compound 15 (CDCl₃, 700MHz)

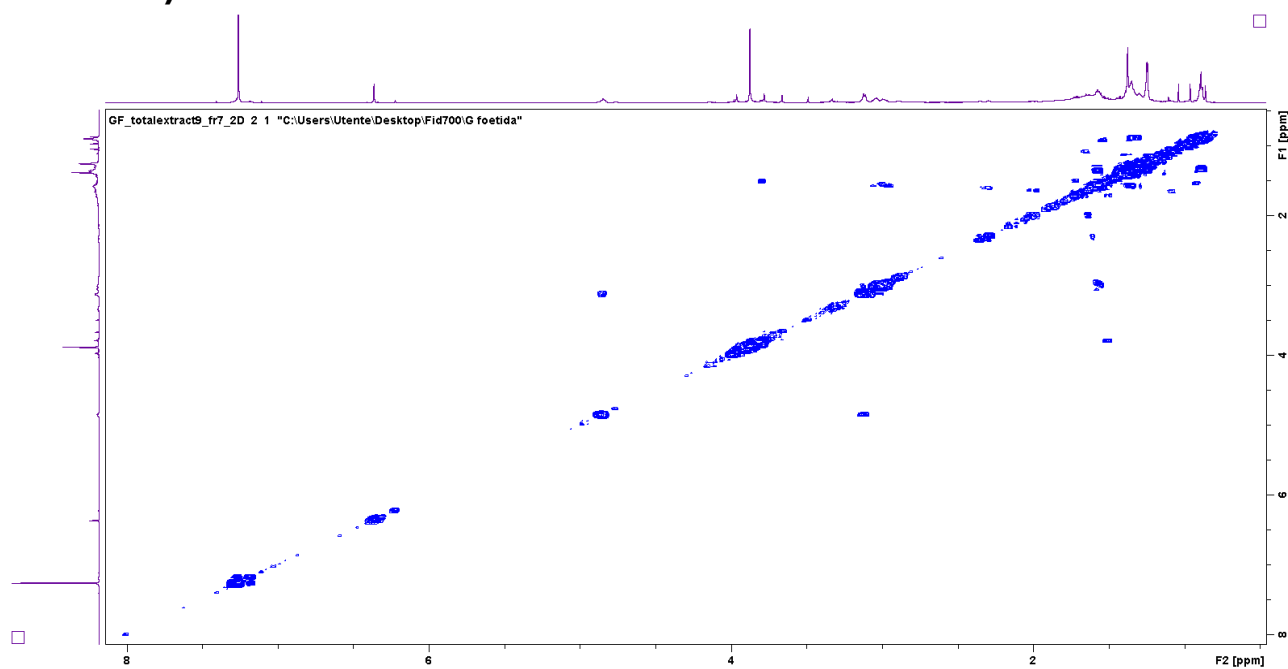


Figure S38: ¹H NMR spectrum of compound 16 (CDCl₃, 700MHz)

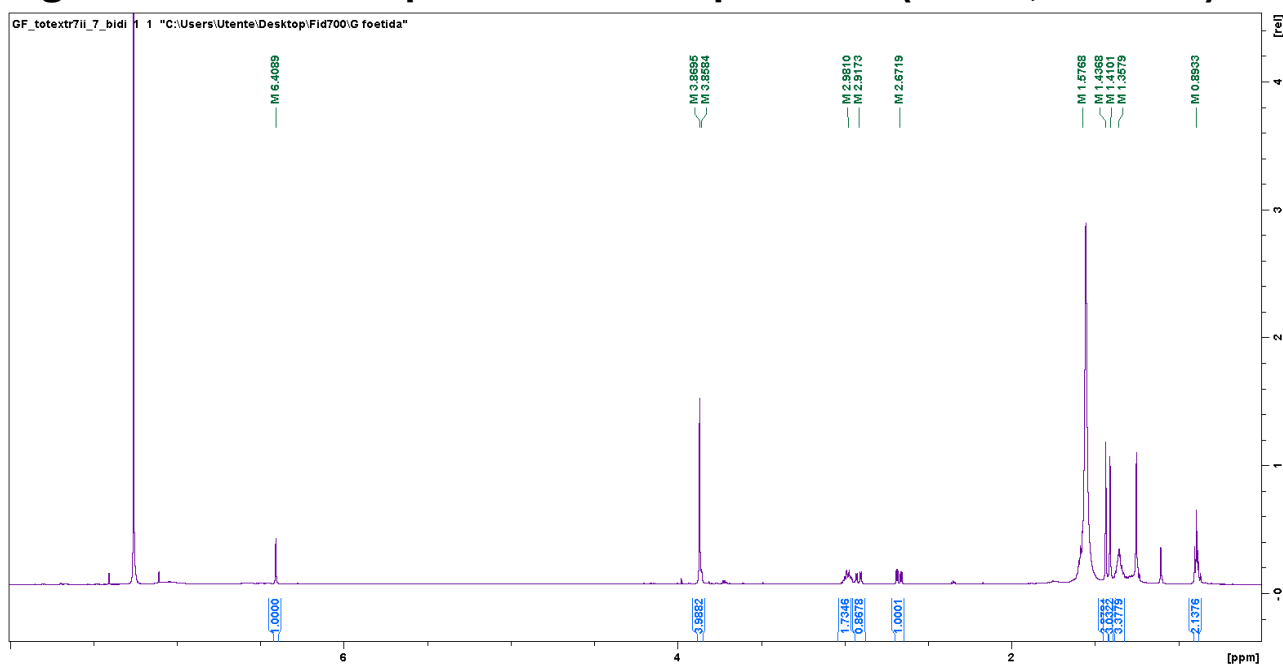


Figure S39: HSQC 2D NMR spectrum of compound 16 (CDCl₃, 700MHz)

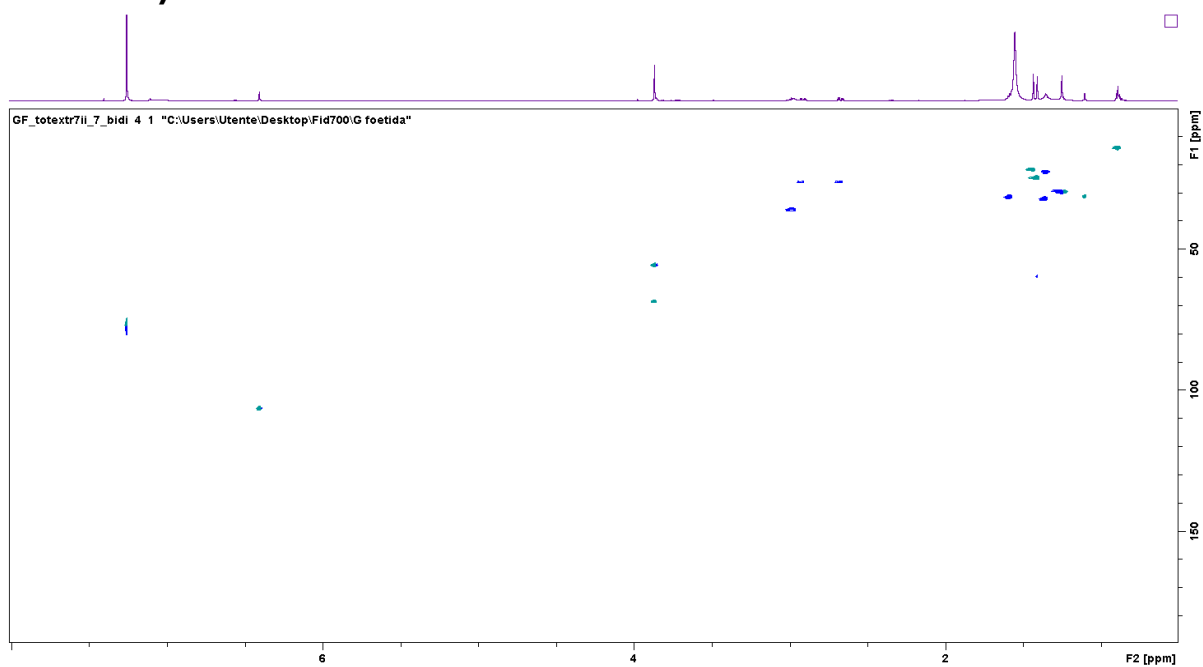


Figure S40: HMBC 2D NMR spectrum of compound 16 (CDCl₃, 700MHz)

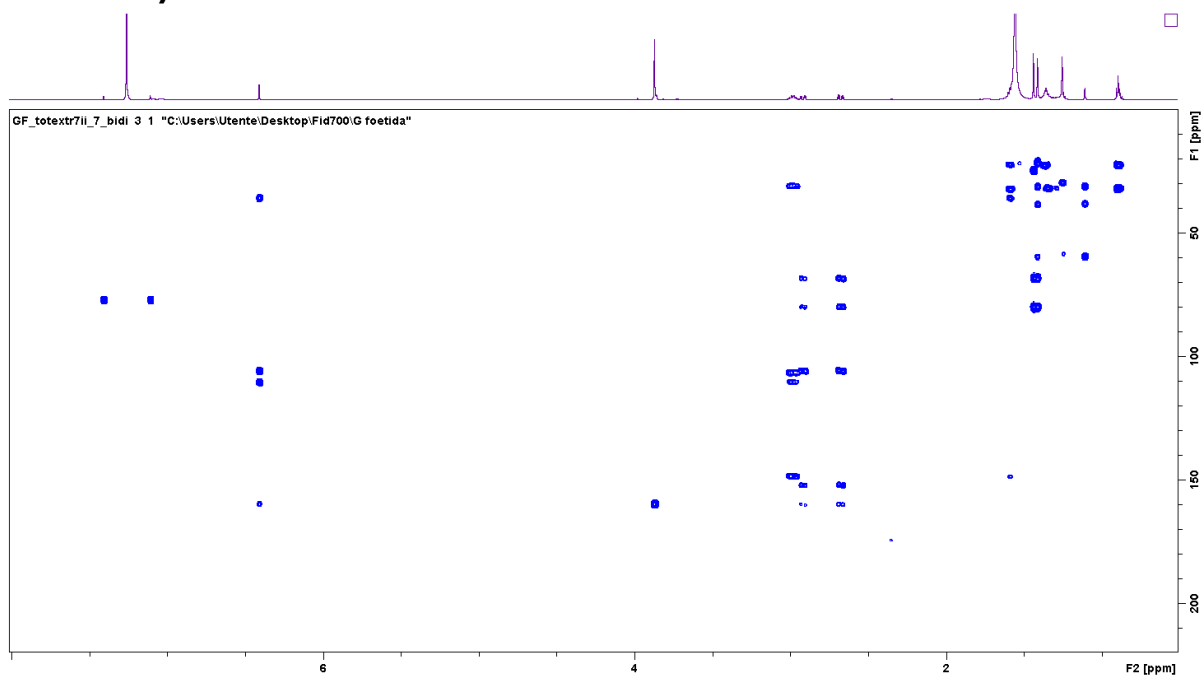


Figure S41: COSY 2D NMR spectrum of compound 16 (CDCl₃, 700MHz)

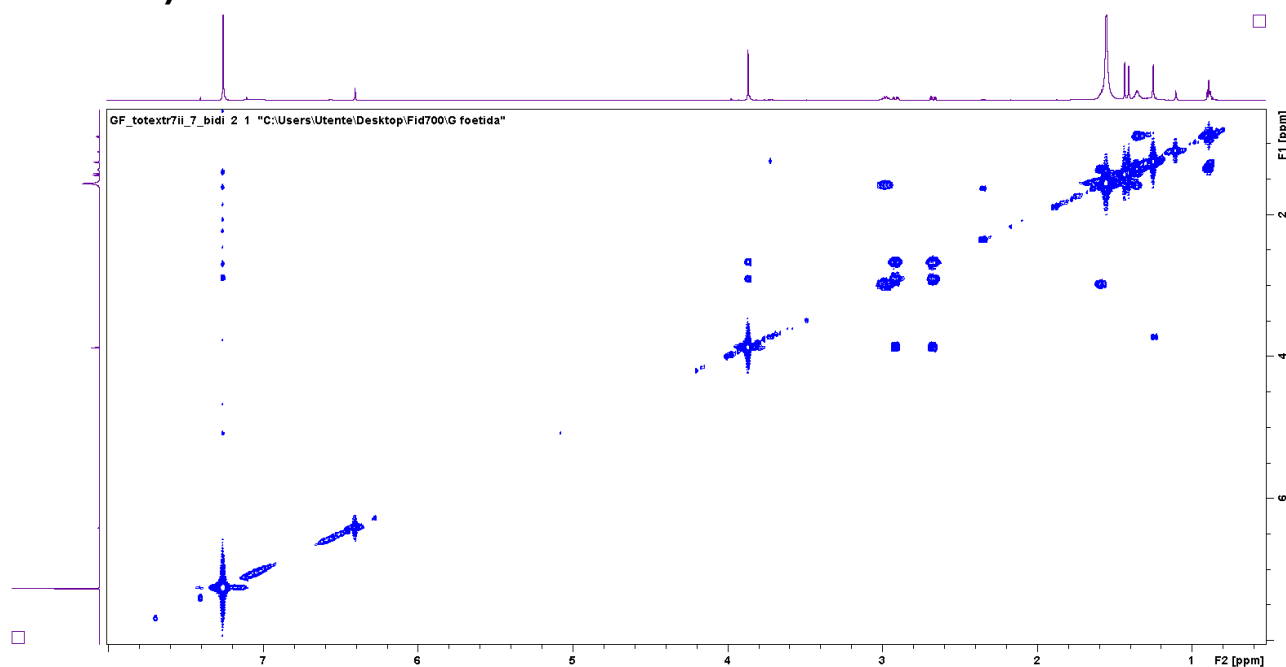


Figure S42: ¹H NMR spectrum of compound 17 (CDCl₃, 600MHz)

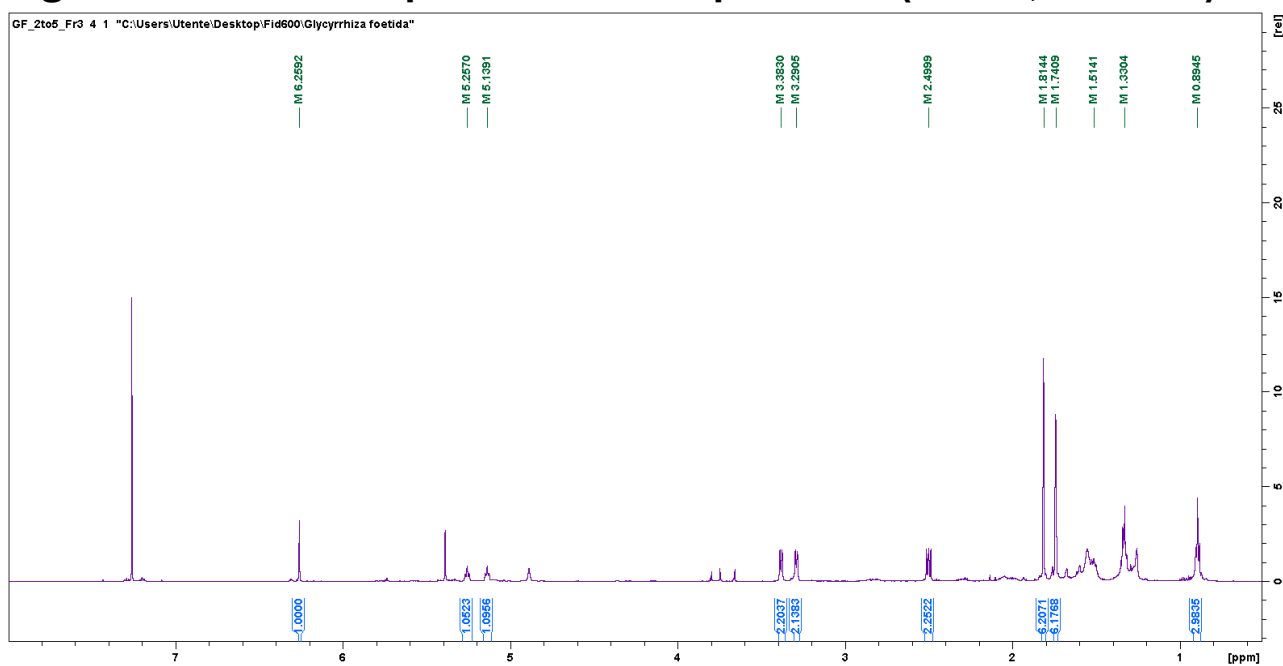


Figure S43: HSQC 2D NMR spectrum of compound 17 (CDCl₃, 600MHz)

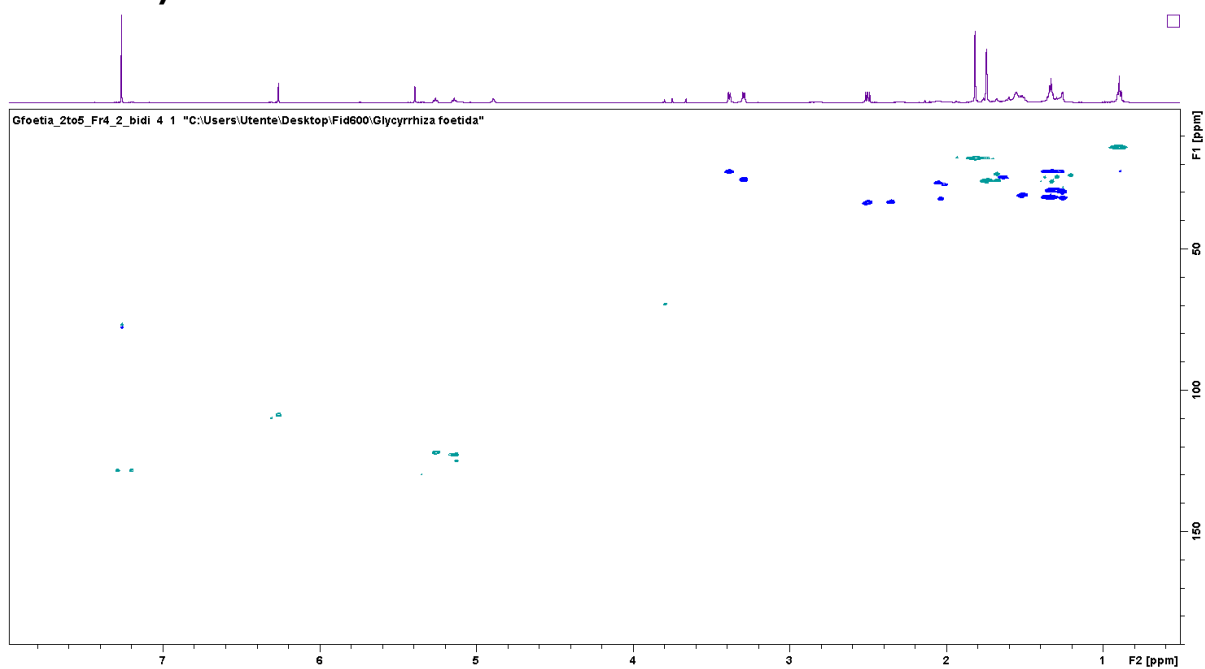


Figure S44: HMBC 2D NMR spectrum of compound 17 (CDCl₃, 600MHz)

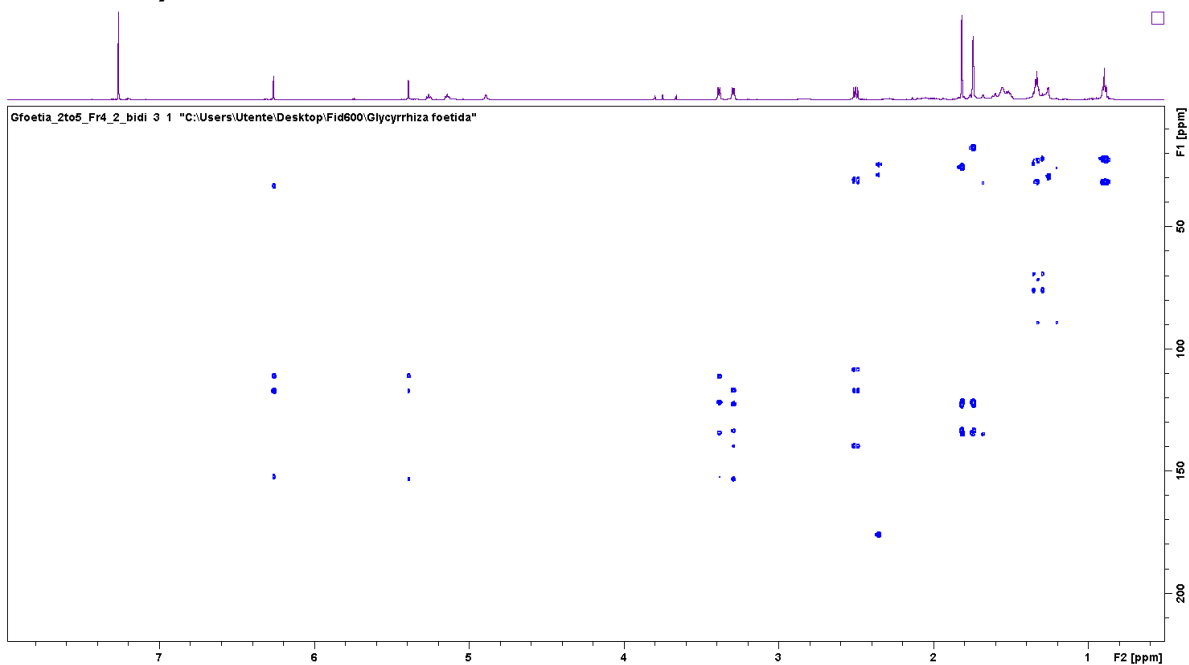


Figure S45: COSY 2D NMR spectrum of compound 17 (CDCl₃, 600MHz)

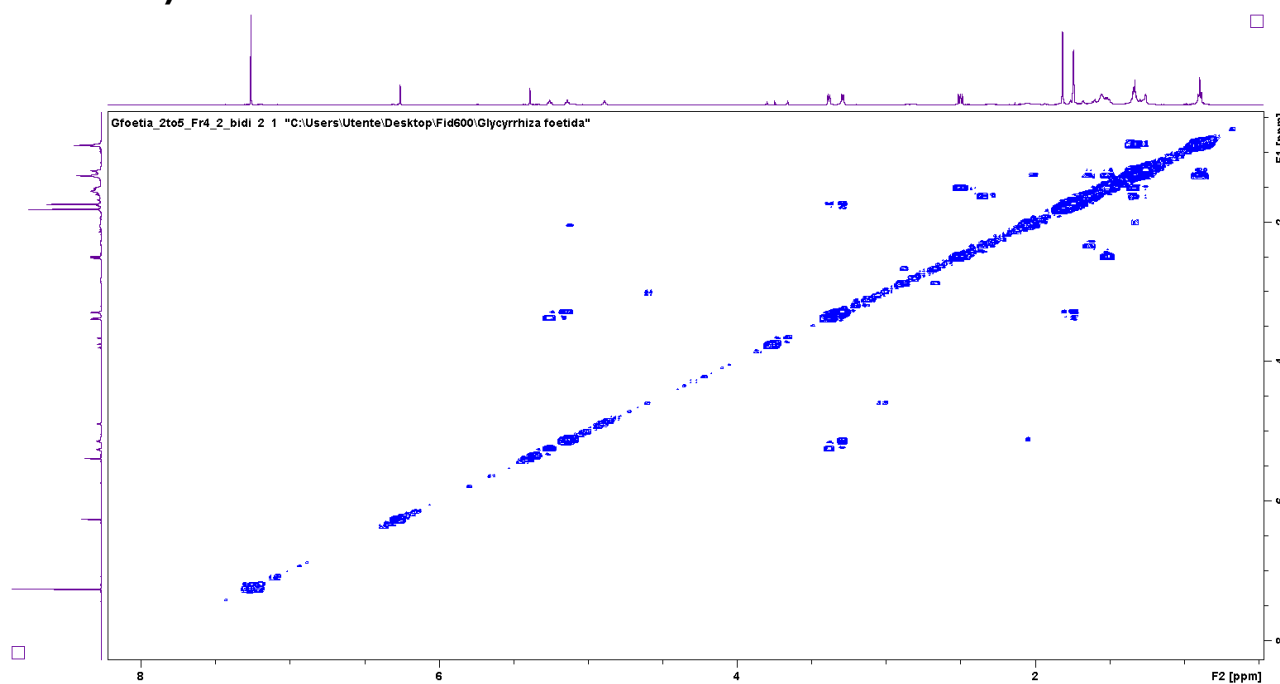


Figure S46: ¹H NMR spectrum of compound 22 (CDCl₃, 600MHz)

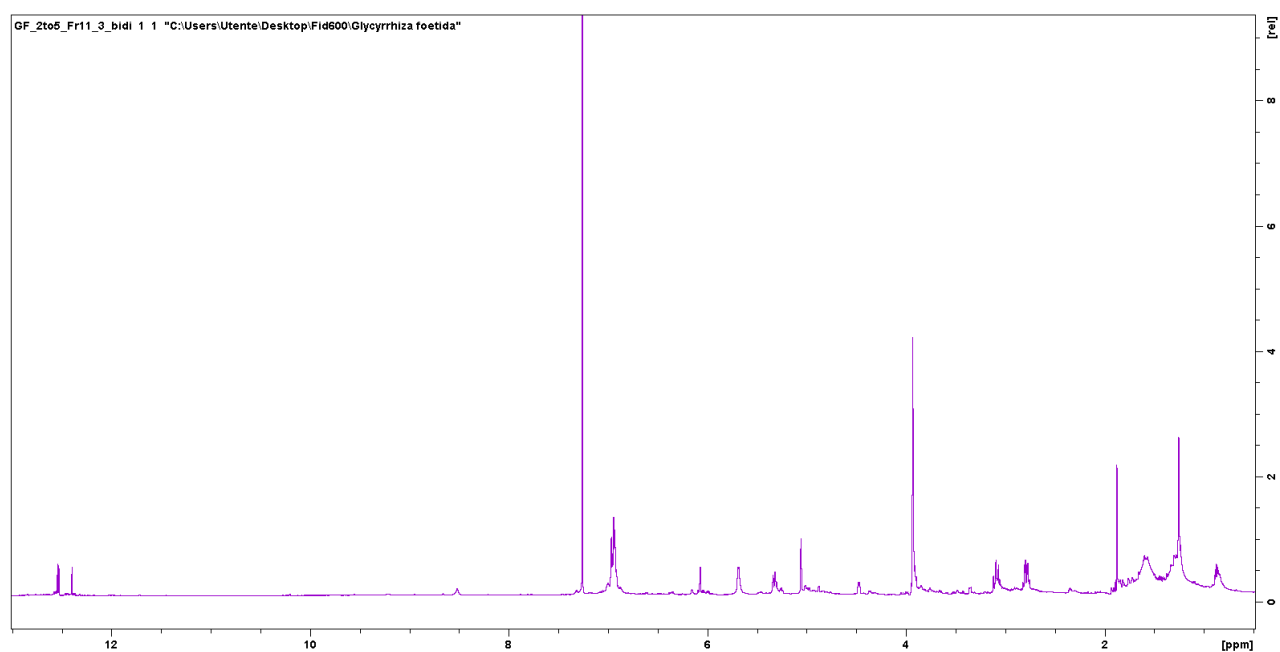


Figure S47: HSQC 2D NMR spectrum of compound 22 (CDCl₃, 600MHz)

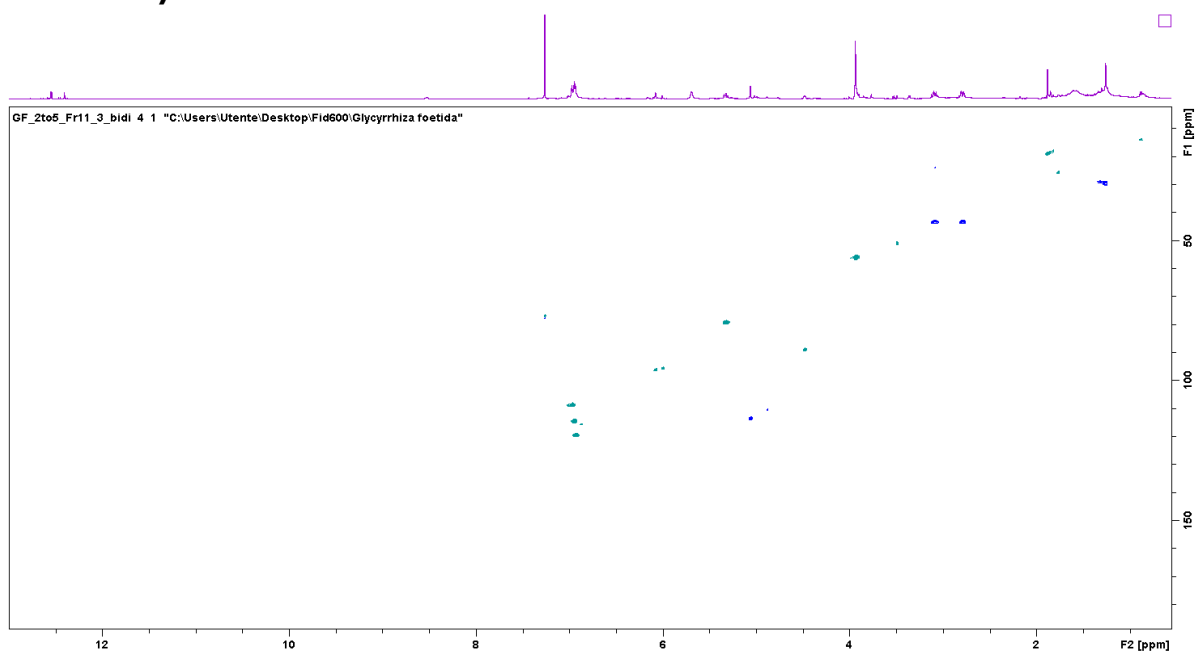


Figure S48: HMBC 2D NMR spectrum of compound 22 (CDCl₃, 600MHz)

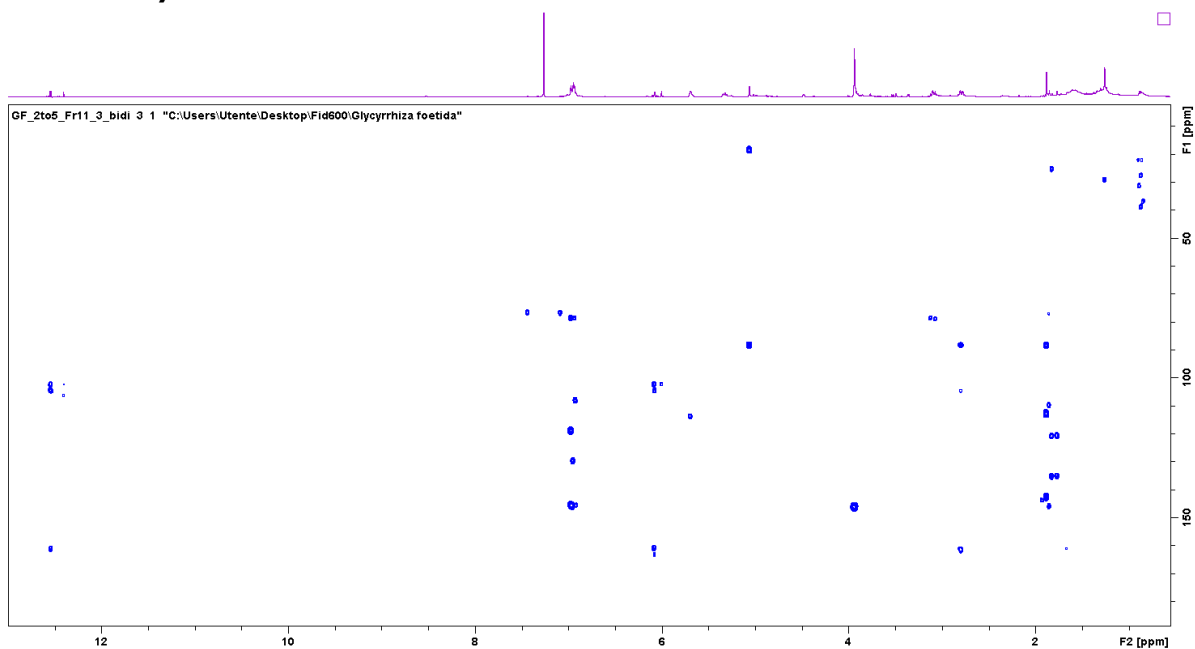


Figure S49: COSY 2D NMR spectrum of compound 22 (CDCl₃, 600MHz)

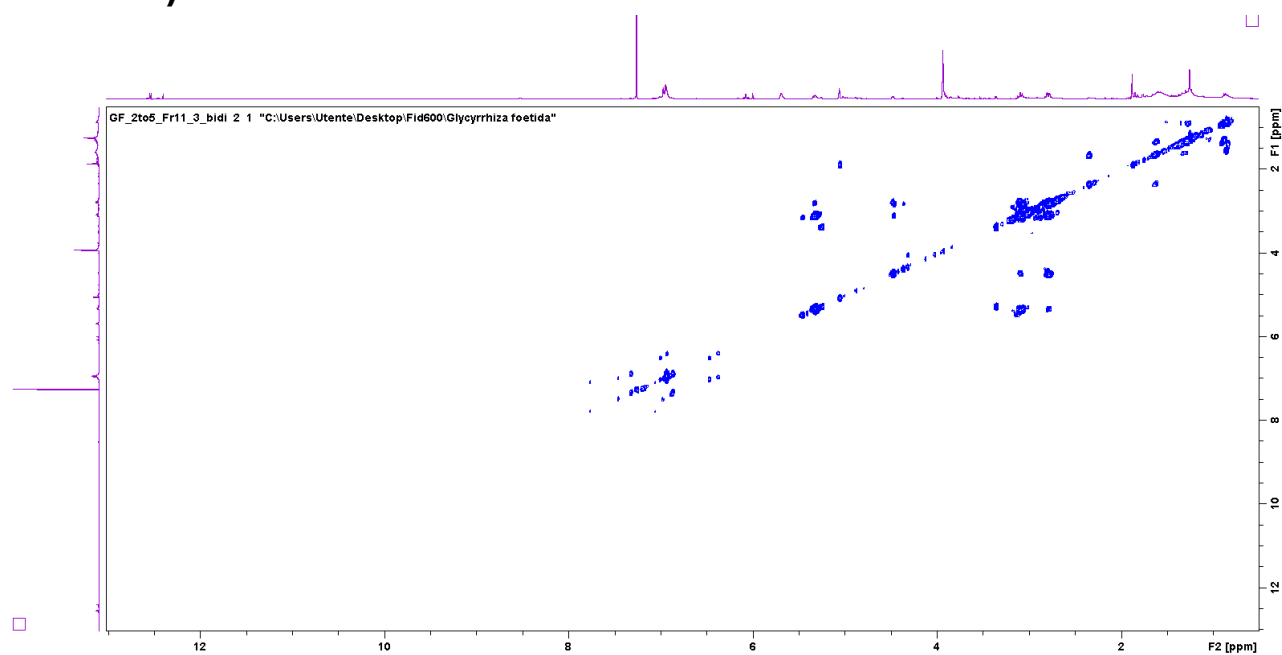
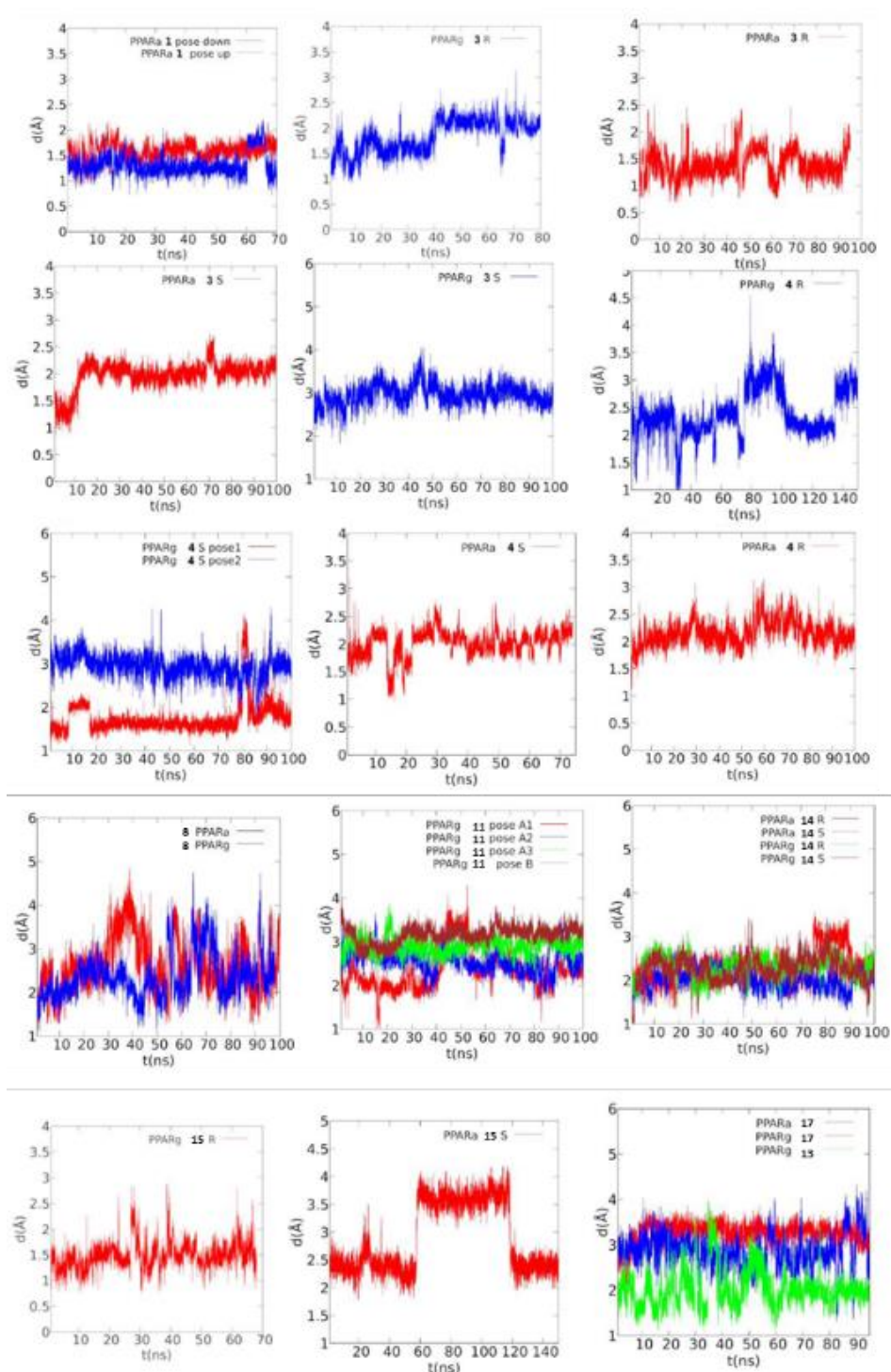


Figure S50: Root-mean-square deviation (RMSD) plots of amorfrutins in complex with PPAR α / γ over MD trajectories after best fitting of protein backbones. The ligand RMSD plot was smoothed with a five-point window running average.



CHAPTER 2: Oleanolic acid, a promising antidiabetic metabolite detected in Aglianico grape pomace

Grape pomace, a bulky component of winery waste, is a source of healthy compounds. So far, scientific research has mainly focused on its polyphenol content, but given the impressive number of bioactivities shown by grape pomace, it is not unlikely that, besides polyphenols, additional metabolites, so far undetected, may be involved. In order to verify such hypothesis, an in-depth chemical analysis of Aglianico (*Vitis vinifera*) grape pomace was conducted by NMR and LC-MS/MS. In addition to a number of polyphenols, a remarkable concentration of oleanolic acid (0.45 mg/g - fresh weight) was determined in the analyzed material. Oleanolic acid is a natural triterpenoid showing many bioactivities including antitumor, anti-inflammatory, antibiotic and antiviral properties. Also, it was proven a potential antidiabetic molecule in Type1 Diabetes rats. Hence, its influence on the mitochondrial and glucose uptake activities of C2C12 myoblast was here assessed, thus supporting oleanolic acid as a promising antidiabetic metabolite.

2.1 Potential bioactive secondary metabolites of Aglianico

Annually over 70 million tons of grapes are produced across the planet and approximately 70% of the total grape production is employed for winemaking (FAO, 2019; García-Lomillo & Gonzalez-SanJose, 2017). In major wine producing countries, impressive amounts of waste are generated by wineries, which cause environmental concerns such as air pollution, leaching and odors (Drevelegka & Goula 2020). It has been estimated that some 20–25% of the grape weight account for the main winery by-product usually referred to as grape or wine pomace (or marc). Grape pomace is constituted by residual pulp, pressed skins, seeds, pieces of stalks and yeast cells used in the fermentation phases of winemaking. This material has long been regarded as an unvaluable waste and a heavy economic and environmental burden to deal with. Therefore, over the past decades some possible reuses of pomaces have been proposed, in order to valorize it and derive economic benefits. Among major alternatives, distillation procedures have long been used to obtain distilled spirits, liquors and liqueurs (Gonzalez-SanJose, 2014). Alternatively, pomaces have been exploited to produce feed for farm animals or even fertilizing material to improve the nitrogen and mineral concentrations in vineyard soils

(Soceanu et al., 2021). However, in either case there are disadvantages, since grape pomaces also contain antinutritive molecules that are detrimental in terms of crop yields and animal growth (Dwyer et al., 2014). Another promising reuse of pomace capitalizes on the recovery of primary and secondary metabolites, including polysaccharides, phytochemicals, dietary fibers, along with other health-promoting nutrients (Iqbal et al., 2021). Polyphenols have been associated with interesting antioxidant, vasodilatory, antithrombotic, cardioprotective, anti-inflammatory, anticancer and antimicrobial bioactivities (Kato-Schwartz et al., 2020; Zhu et al., 2019; Peixoto et al., 2018). More recently, grape pomace polyphenols and their polysaccharide conjugates have been proven to exert antidiabetic effects (Campos et al., 2021). Thus, not surprisingly, the development of efficient and, at the same time, environmentally friendly, extraction methods for polyphenols from pomace has mainly attracted the interest of scientists. Among such techniques, microwaves, ultrasound, high hydrostatic pressure, pulsed electric fields, enzymes, and deep eutectic solvents seem to be quite promising (Calderon-Oliver & Ponce-Alquicira, 2021). The combination of the above techniques with other strategies, including micro- and nanoencapsulation, turned out to stabilize beneficial compounds (Calderon-Oliver & Ponce-Alquicira, 2021). However, even though some of the reported extraction technologies have already been scaled up to an industrial level in the frame of circular economy, their industrial use is still far from being imminent. Although anthocyanins and their derivatives certainly represent the major phenolics recovered from red grape pomace, other polyphenols such as flavanols, flavonols, condensed tannins, stilbenes and lignans have been detected in pomaces and reused in pharmaceutical and food industries (García-Lomillo & Gonzalez-SanJos, 2017). In this context, it is well documented that polyphenol-rich plant extracts have been employed as safe and clean alternatives to synthetic additives in order to delay food spoilage, retard microbial growth as well as fat deterioration, thus ultimately extending the shelf-life of products (Pateiro et al., 2021). Moreover, polyphenols are successfully used even in the cosmetic segment, as it is well-documented for phenols recovered from olive mill wastewater that are employed as UV boosters (Galanakis et al., 2018). However, it is now clear that the impressive potential of health related beneficial effects shown by grape pomace extracts, spanning from anti-inflammatory to antidiabetic (Campos et al., 2021) and antiviral activities (Huang et al., 2016), cannot be ascribed exclusively to the presence of polyphenols. The secondary metabolism of a single plant is rarely focused on a single class of compounds and, just to cite a well-known example, the multitude of bioactivities

associated to Lamiaceae plants (e.g. rosemary, sage, thyme, lemon balm or wild marjoram) (Sharma et al., 2020) is due to phenolic compounds, such as rosmarinic acid, but to a great extent also to triterpenoids such as oleanolic and ursolic acids (Lesellier et al., 2021). Accordingly, triterpenoids are emerging as important dietary supplements and they are experiencing a remarkable market growth, especially in the nutraceutical sector (Farzan et al., 2023). Oleanolic acid is a pentacyclic triterpenoid with a wide distribution in the plant kingdom, and it is a major component of the grape berry cuticle (Castellano et al., 2022) with antitumor, antibiotic and even antiviral properties (Lin et al., 2016). As demonstrated in some studies conducted during the COVID-19 pandemic, the molecule turned out to interfere with the interaction of the spike protein receptor binding domain with the carboxypeptidase domain of the Angiotensin Converting Enzyme (ACE2) (Carino et al., 2020). Furthermore, oleanolic acid can improve the structural integrity of mitochondria and endoplasmic reticulum in liver and pancreas of animals fed with HFF diet (high-fat diet plus fructose) (Wang et al., 2018) and showed both antioxidant and anti-inflammatory effects, reducing cartilage degeneration caused by hyperglycemia condition (Li et al., 2021). Equally interesting is the capacity of this triterpenoid of targeting many end points of anti-diabetic relevance by improving the insulin response, by preserving the functionality of pancreatic beta-cells and by generally protecting from diabetic complications (Castellano et al., 2022). Thus, oleanolic acid seems to be a promising natural small molecule which could be used as supplement in diets for patients affected by diabetic and metabolic syndromes. It is interesting to emphasize that oleanolic acid has been recently identified in Aglianico (*Vitis vinifera* L.) red wine but in concentrations well below those detected in grape samples of the same cultivar (Forino et al., 2019). Aglianico grapevine is largely grown across Southern Italy for the production of renowned red wines awarded the national highest wine classification, DOCG designation (Denominazione di Origine Controllata e Garantita); nonetheless, just a few reports are currently available on the chemical characterization of Aglianico grape varieties (Forino et al., 2019; Arapitsas et al., 2020) as well as on the bioactivities of the derived pomaces. In this regard, very recently some studies describing the antimicrobial, anti-proliferative and pro-apoptotic effects of Aglianico grape pomace extracts have been published (Caponio et al., 2022), but no information on the metabolites responsible for this activity was provided. Thus, with the purpose of contributing to fill in the above-mentioned knowledge gap, we decided to conduct a full chemical characterization of the bioactive secondary metabolites of the grape pomace deriving from the

vinification process of Aglianico with a specific focus on oleanolic acid. In fact, on account of its reduced solubility in hydroalcoholic solutions, such as wine, it was reasonable to assume that large amounts of oleanolic acid could be accumulated in red grape pomaces. The adopted analytical strategy was mainly based on a combined untargeted and targeted approach by means of NMR and Mass Spectrometry (MS). In particular, the untargeted analysis was aimed at identifying possible major bioactive compounds that would turn out to be worthy to recover and reutilize in the dietary, pharmaceutical and nutraceutical sectors, as described above (Annunziata et al., 2020; Badolati et al., 2020; Cecchi et al., 2019).

2.2 Chemical analysis of grape pomace

Grape pomace samples from Aglianico (*Vitis vinifera* L. cv) were collected in the autumn of 2021 from the Taurasi area (Avellino, Italy), renowned for the production of a high-quality red wine with the Controlled and Guaranteed Designation of Origin (Taurasi DOCG). Pomace was lyophilized soon after collection to minimize the degradation of organic molecules. An aliquot of the lyophilized grape pomace was subjected to different and subsequential extractions with solvents with increasing degree of polarity, thus obtaining three different fractions, named A, B and C, as detailed in Materials and methods. These fractions were all subjected to preliminary NMR-based analyses, a crucial step to identify potential phytochemicals worth to be recovered in the frame of circular economy. ^1H NMR spectrum of A contained typical chemical shifts of unsaturated fatty acids, and because of their estimated small recoverable amounts the fraction was no further analyzed. Fraction C turned out to be constituted mainly by glycerol, carbohydrates (signals attributable to glucose and fructose were detected), organic acids (tartaric, malic and citric acids) and amino acids, whose quali-quantitative content was determined by the HPLC-based method as reported below. Particularly interesting was fraction B. In its ^1H NMR spectrum diagnostic signals of triterpenoids, flavan-3-ols, anthocyanins, and polymeric pigments were detected. To support this preliminary assignment and identify such metabolites, fraction B was subjected to chromatographic separation through an 80 g-Silica Combiflash column, from which four different fractions were collected and analyzed by NMR spectroscopy. Once again, all of the four obtained fractions were analyzed by NMR spectroscopy. Fraction 1 turned out to be composed by a pure triterpenoid

compound, whose chemical identification was conducted by extensive mono- and bidimensional-700 MHz NMR analysis (^{13}C NMR, COSY, TOCSY, ROESY, HSQC and HMBC) without further purification (**Figure 2.1**). The comparison of the obtained NMR data with those previously reported (Forino et al., 2019) led us to unambiguously characterize the isolated triterpenoid as oleanolic acid.

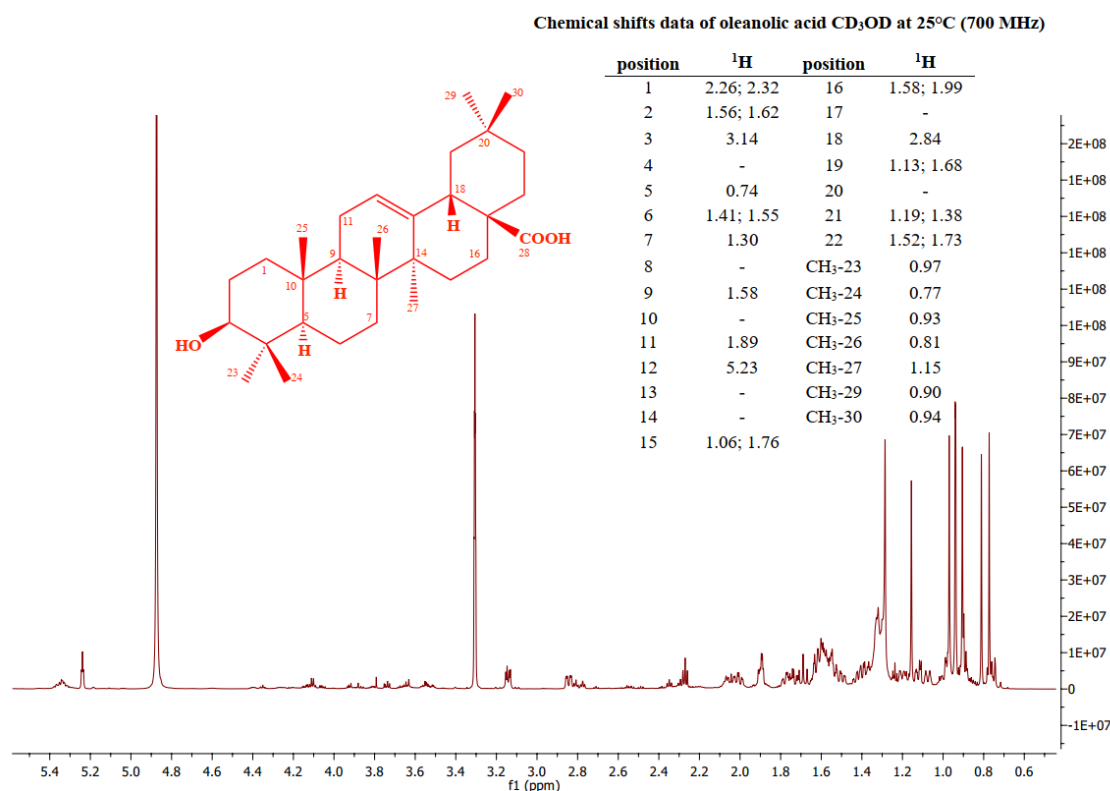


Figure 2.1 ^1H NMR spectrum (CD_3OD) of fraction 1 containing oleanolic acid as major component. In the Table ^1H chemical shifts of the molecule are reported

Our structural hypothesis was also supported by HR-ESIMS analysis that afforded a $[\text{M}+\text{H}]^+$ ion peak at m/z 457.3664 corresponding to $\text{C}_{30}\text{H}_{49}\text{O}_3$ (calculated for $\text{C}_{30}\text{H}_{49}\text{O}_3$ 457.3676). NMR spectroscopy also allowed an accurate quantitation of oleanolic acid, by using pyridine as an internal standard, a solvent selected because its proton resonances did not interfere with those of oleanolic acid. Thus, fraction 1 turned out to contain 160.1 ± 3.7 mg of oleanolic acid, a relevant piece of information to evaluate the potential recoverable amount of this metabolite. Hence, oleanolic acid accounted for about 3.2% of the dry weight of the EtOAc extract (4.96 g) and about 2% of the dry weight of the initial hydroethanolic extract (7.94 g) deriving from 100 g of dry grape pomace. The concentration of oleanolic acid in fresh Aglianico pomaces can be

estimated to hover around 0.45 mg/g, a value consistent with a reported estimation of 0.85 mg/g of oleanolic acid in Aglianico berry skins (Forino et al., 2019). Our NMR-based investigation did not highlight any other triterpenoids apart from oleanolic acid, at least at concentrations detectable by NMR. Hence, a targeted analysis was conducted by means of LC-MS in positive ion mode. It allowed us to detect just a $[M+H]^+$ single ion peak at 14.52 min at m/z 457.3664 attributable to oleanolic acid. This observation was instrumental in ruling out the occurrence of further oleanolic acid isomers, such as ursolic or betulinic acid. At trace levels, we detected a $[M+H]^+$ ion peak at m/z 471.3817 corresponding to $C_{31}H_{51}O_3$ (calculated for $C_{31}H_{51}O_3$ 471.3833) that was tentatively assigned to oleanolic acid methyl ester on the basis of previous reports (Pensec et al., 2014). Other triterpenoid derivatives, including aldehyde-containing analogues, detected in some grape cultivars (Pensec et al., 2014), were unsuccessfully searched for. In regards to fractions 2 and 3, preliminary NMR-based investigations brought to light the presence of typical resonances of flavan-3-ols, including (+)-catechin, (-)-epicatechin moieties in both fractions 2 and 3, and NMR signals diagnostic of anthocyanin units in fraction 3. To the aim of better defining the polyphenol content of either fraction, we resorted to LC-MS analysis, since it mainly served the purpose of an accurate targeted analysis of the compounds detected by NMR. Considering that the biological matrix under investigation was containing anthocyanin moieties, as also inferred by its red coloration, we resorted to MS experiments performed in positive ion mode suitable for the analysis of all classes of flavonoids. As detailed in **Table 2.1**, a number of polyphenols were detected and quantified. In fraction 2, phenolic acids and flavan-3-ol monomers were identified. More specifically and consistently with NMR data, the most abundant compounds turned out to be (+)-catechin and (-)-epicatechin, of which only a small fraction occurred as gallate esters. Additionally, among phenolic acids, caffeic acid was the most abundant. Conversely, A-type procyanidins were detected in very small quantities, consistently with previous reports (Unusan, 2020) that have demonstrated that in grapes B-type procyanidins are usually more abundant than their A-type analogs. Indeed, B-type procyanidins were identified in fraction 3, along with flavonols (quercetin and myricetin) and anthocyanins. Among the common anthocyanin monomers occurring in red wines, we could identify and quantify only malvidin-3-O-(6-*p*-coumaryl)-glucoside and peonidin-3-O-(6-*p*-coumaryl)-glucoside. In order to confirm the monomeric anthocyanin profile of the fraction, an HPLC-based analysis was conducted on the initial hydroethanolic extract of grape pomace according to the OIV method (OIV, 2020) for

the analysis of native anthocyanins in grapes and wine. This technique was selected on account of its enhanced selectivity and specificity for anthocyanins. Again, HPLC peaks relative to malvidin-3-*O*-(6-*p*-coumaryl)-glucoside and peonidin- 3-*O*-(6-*p*-coumaryl)-glucoside in the extract as major components were observed. HPLC peaks relative to malvidin-3-*O*-glucoside as well as to peonidin-3-*O*-glucoside, which are in the order the two most abundant anthocyanin monomers in the Aglianico red grapes, were detected but below their quantification limit. This datum is of a certain importance to enologists, as it could prompt studies on the extraction kinetics of anthocyanins from grapes into musts. This would suggest the best time to end the maceration process as to possibly avoid the release from crushed grapes of undesirable quantities of molecules, including tannins, that could affect the quality of the finished wines. Finally, by the NMR-based analysis it was ascertained that polymeric pigments had accumulated in fraction 4, as inferred from either the detection of chemical shifts typical of anthocyanin moieties and the drift of the ^1H -spectrum baseline. This was not surprising as the most abundant grape and wine polyphenols occur as polymeric compounds commonly ranging from 1000 to 5000 mg/L in red wines, and thus well above the levels of monomeric compounds such as resveratrol, whose concentrations hardly are found above 5 mg/L. This is the reason why the contribution of polymeric polyphenols in terms of health-related effects has necessarily to be taken into account (D'Onofrio et al., 2021). An impressive number of papers has been published on the bioactivities of grape and wine polyphenols either monomeric, oligomeric or polymeric (Li & Sun, 2019). As a consequence, fraction 4 containing essentially polymeric pigments was not further investigated.

Table 2.1 Compounds identified by LC-MS/MS in positive ion mode in fractions 2 and 3 deriving from the chromatographic separation of the EtOAc extract of Aglianico grape pomace

Compounds	Rt (min)	m/z	D(ppm)	m/z fragments	mg/g
Fraction 2					
<u>Phenolic Acids</u>					
Gallic acid	3.13	171.0285	-1.753	110.98, 127.01	0.009 ± 0.000
Caffeoyltartaric acid	7.88	313.0539	-4.818	150.81, 181.03	0.002 ± 0.000
Caffeic acid	12.08	181.0490	-2.791	136.90	0.045 ± 0.001
Syringic acid	12.12	199.0596	-2.411	140.05, 155.07	0.021 ± 0.001
<u>Flavan-3-ols</u>					
Catechin	7.78	291.0854	-3.177	127.04, 273.08	0.459 ± 0.001
Epicatechin	9.61	291.0851	-4.345	181.05, 273.08	0.526 ± 0.008
Catechin-3- <i>O</i> -gallate	10.83	443.0958	-3.392	273.08, 291.09	0.086 ± 0.001
Epicatechin-3- <i>O</i> -gallate	11.34	443.0960	-2.873	273.08, 291.09, 333.14	0.022 ± 0.000

PART 3

Procyanidin dimer A	12.65	577.1319	-3.626	425.09	0.005 ± 0.000
Procyanidin dimer A	13.7	577.1320	-3.504	425.09	0.003 ± 0.000
<i>Flavonols</i>					
Myricetin	13.29	319.0439	-2.801	301.03	0.054 ± 0.001
Quercetin	14.50	303.0488	-3.726	285.24	0.131 ± 0.001
<i>Stilbens</i>					
Resveratrol dimer	14.32	455.1475	-3.218	349.11, 361.11, 437.14	0.034 ± 0.001
Resveratrol trimer	14.79	681.2100	-2.744	587.17	0.010 ± 0.000
Fraction 3					
<i>Flavan-3-ols</i>					
Procyanidin dimer B	6.69	579.1481	-2.819	289.08, 409.09, 453.12, 561.09	0.064 ± 0.001
Procyanidin dimer B	6.99	579.1478	-3.268	289.08, 409.09, 453.12, 561.09	0.070 ± 0.000
Procyanidin dimer B	7.28	579.1474	-4.045	289.08, 409.09, 453.12, 561.09	0.078 ± 0.003
Catechin	7.82	291.0851	-4.035	127.04, 273.08	0.081 ± 0.002
Procyanidin dimer B	7.98	579.1476	-3.596	289.08, 409.09, 453.12, 561.09	0.104 ± 0.001
Procyanidin dimer B	8.73	579.1477	-3.389	289.08, 409.09, 453.12, 561.09	0.163 ± 0.002
Procyanidin trimer B	9.14	867.2118	-1.534	289.07, 579.11	0.011 ± 0.000
Procyanidin trimer B	9.49	867.2109	-2.561	289.07, 579.11	0.016 ± 0.000
PC2-G isomer 1	9.50	731.1579	-3.722	409.09, 427.10, 579.11, 605.10	0.089 ± 0.001
Epicatechin	9.58	291.0856	-2.352	181.05, 273.08	0.154 ± 0.001
Procyanidin trimer B	9.68	867.2120	-1.258	289.07, 579.11	0.007 ± 0.000
Catechin-3-O-gallate	10.83	443.057	-3.483	273.08, 291.09	0.061 ± 0.002
Procyanidin dimer B	11.19	579.1478	-3.285	289.08, 409.09, 453.12, 561.09	0.043 ± 0.002
Epicatechin-3-O-gallate	11.34	443.0960	-2.873	273.08, 291.09, 333.14	0.012 ± 0.000
PC2-G isomer 2	11.92	731.1577	-4.050	409.09, 427.10, 579.11, 605.10	0.016 ± 0.002
<i>Anthocyanins</i>					
Pn-3-p-coumglc <i>cis</i>	13.60	609.1584	-3.131	301.07	0.012 ± 0.001
Mv-3-p-coumglc <i>cis</i>	13.72	639.1689	-2.944	331.08	0.156 ± 0.007
Pn-3-p-coumglc <i>trans</i>	13.83	609.1587	-2.491	301.07	0.027 ± 0.002
Mv-3-p-coumglc <i>trans</i>	13.95	639.1691	-2.757	331.08	0.259 ± 0.008

PCn: n units of (epi)catechin; G: galloyl unit. Pn-3-p-coumglc: Peonidin-3-O-(6-O-p-coum aryl)glucoside; Mv-3-p-coumglc: Malvidin-3-O-(6-O-p-coumaryl)glucoside. All the data are expressed as means ± standard deviation, (p < 0.05).

2.2.1 HPLC-DAD analyses of amino acids

The amino acid content and composition, along with polyphenols, is often regarded as a chemotaxonomic tool to differentiate grape varieties. In addition to variety, amino acid

composition of grapes is strictly dependent upon pedo-climatic conditions and cultivation practices. We considered a key issue providing insights into the profile of amino acids of the analyzed grape pomace to evaluate their specific residual amounts and assess as to whether it would be advisable a potential recovery. The Aglianico-derived pomace is a fermented one and thus the yeasts play a critical role in determining the amino acidic content. In fact, through the yeast activities, amino acids are converted into secondary metabolites, such as esters, alcohols, aldehydes and ketones that ultimately influence the aroma of wines. The qualitative analysis of amino acids was carried out according to the DEEMM (diethyl ethoxymethylenemalonate)-based method reported by Gomez-Alonso et al. (2007). **Table 2.3** displays the identified amino acids along with their quantitation, conducted as described in Materials and methods. Proline and tyrosine turned out to be the two most abundant amino acids detected in the pomace, in which they occur in comparable and remarkable amounts. In the frame of a possible reuse and valorization of Aglianico grape pomace, it is worth bearing in mind that, over the past decades, tyrosine has been proven to increase dopamine availability, a key metabolite able to enhance cognitive processes (Kühn et al., 2019). This has suggested tyrosine as a food supplement. Also, tyrosine is an interesting compound as a precursor of tyrosol, a powerful bioactive compound (Karkovic Markovic et al., 2019) produced by yeasts during alcoholic fermentation through the Ehrlich pathway. Tyrosol exerts antioxidant activity and it is able to bind human LDL (low-density lipoproteins) (Covas et al., 2003), and thus it can prevent lipid peroxidation and atherosclerotic processes. In regards to proline, it is not surprising that this amino acid was quite abundant in pomaces, as it cannot be assimilated by yeasts during fermentation. The high content of proline in grape pomace supports its potential as food and feed sources in either human and animal diets. It is interesting to underline that proline is generally synthesized by plants as a protection against water stress caused by drought and salinity. Given that the pomace analyzed in this study was obtained from grapes collected after a quite warm and dry summer, it would be worth to study the possible concentration variation of proline as a function of different summer climatic characteristics.

Table 2.3 Amino acids detected in Aglianico grape pomace. Concentrations are expressed as mg/gr of dry pomace.

	<i>Retention time</i>	<i>mg/g</i>
Glutamic acid	27.02; 26.84	0.010 ± 0.003
Serine	31.38	*

PART 3

Asparagine	32.35	*
HO-proline	33.82	0.013 ± 0.000
Glutamine	34.47	*
Histidine	35.01	0.024 ± 0.001
Glycine	35.54	0.030 ± 0.000
Threonine	36.28	0.006 ± 0.001
β-alanine	36.66	0.007 ± 0.000
Arginine	37.21	0.010 ± 0.001
α-alanine	38.62	0.041 ± 0.001
GABA	39.17	0.005 ± 0.001
Proline	39.78	0.536 ± 0.006
Tyrosine	42.23	0.592 ± 0.006
Valine	46.21	*
Methionine	47.40	*
Cysteine	50.58; 56.56	0.007 ± 0.005
Isoleucine	51.92	*
Tryptophan	52.07	*
Leucine	53.04	0.012 ± 0.005
Phenylalanine	53.71	*
Ornithine	58.80	0.008 ± 0.004
Lysine	59.73	0.009 ± 0.001

All the data are expressed as means ± standard deviation, ($p < 0.05$). * These amino acids were occurring below the estimated quantitation limit.

2.4 Effect of oleanolic acid enriched fractions on mitochondrial activity of in-vitro cultured C2C12 myoblast

The isolation of remarkable amounts of oleanolic acid offered us the opportunity to carry out some biological assays to deepen our knowledge on the bioactivity of this interesting triterpenoid. As mentioned above, one of the most promising properties of oleanolic acid is its antidiabetic activity. Indeed, obese diabetic mice treated with oleanolic acid showed a decreased fasting blood glucose, improved glucose and insulin tolerance, enhanced insulin signaling and inhibited gluconeogenesis (Wang et al., 2013). Antidiabetic effects of oleanolic acid have been also demonstrated in Type 1 Diabetes (T1DB) rats, in which treatments with this molecule significantly reduced the fasting blood glucose levels. (Wang et al., 2011). Accordingly, it has been also shown that oleanolic acid reduces the inflammatory status and dictates a specific macrophage polarization in adipose tissue of obese mice, thus improving insulin resistance, in part through the inhibition of mitochondrial voltage channel VDAC and, in

addition, by reducing mitochondrial ROS. Taken together these data suggest that oleanolic acid prevents the activation of NLRP3 inflammasome and MAPK signaling pathway (Li et al., 2021). In the wake of the above studies and observations, we decided to focus our attention on the mitochondrial and glucose uptake activities of oleanolic acid. It has been demonstrated that the introduction of lipophilic cations, such as triphenylphosphonium, could target and accumulate the parent compound to the mitochondria of tumor cells enhancing cytotoxicity and selectivity (Zielonka et al., 2017). Recently, triphenylphosphonium cation moiety has been described to improve the selectivity and cytotoxicity of natural pentacyclic triterpenoids by targeting the mitochondria of human cancer cells (Li et al., 2022). In order to verify the pro-metabolic effect of fractions enriched in oleanolic acid, we investigated their ability to promote mitochondrial activity in C2C12 myoblasts. Mitochondrial activity was measured by using the mitochondrial selective probe MitoTracker CMXRos. The probe accumulates in the mitochondrial intermembrane space and emits fluorescence with an intensity that is correlated to the difference in potential between the mitochondrial matrix and the mitochondrial intermembrane space. This electrochemical potential originates from the activity of the electron transport chain and of the H⁺ pumps, providing a measure of mitochondrial oxidative respiration activity. After treatment of C2C12 cells with the EtOAc extract and fractions 1, 2, and 3 at the dosage of 3 µg/mL for 48 h, we observed a statistically significant increase in the mitochondrial activity of C2C12 cells treated with fraction 1 and EtOAc extract compared to those treated with vehicles (DMSO 0,1%) (**Figure 2.2**, panel A). We did not observe any significant variations of C2C12 mitochondrial activity upon treatments with fractions 2 and 3 (**Figure 2.2**, panel A). Thus, fraction 1 turned out to be able to induce mitochondrial activity at the dosage of 1.8 µg/mL, with a dose-dependent potency (**Figure 2.2**, panel B). Since, as reported before, fraction 1 almost exclusively contains oleanolic acid, this bioactivity must be ascribed to this triterpenoid. An imbalance between energy production and its utilization may lead to alteration in the cell metabolism, finally resulting in metabolic syndrome; therefore, the demonstrated effect of oleanolic acid on mitochondrial activity may have a relevance in the prevention/ treatment of this condition.

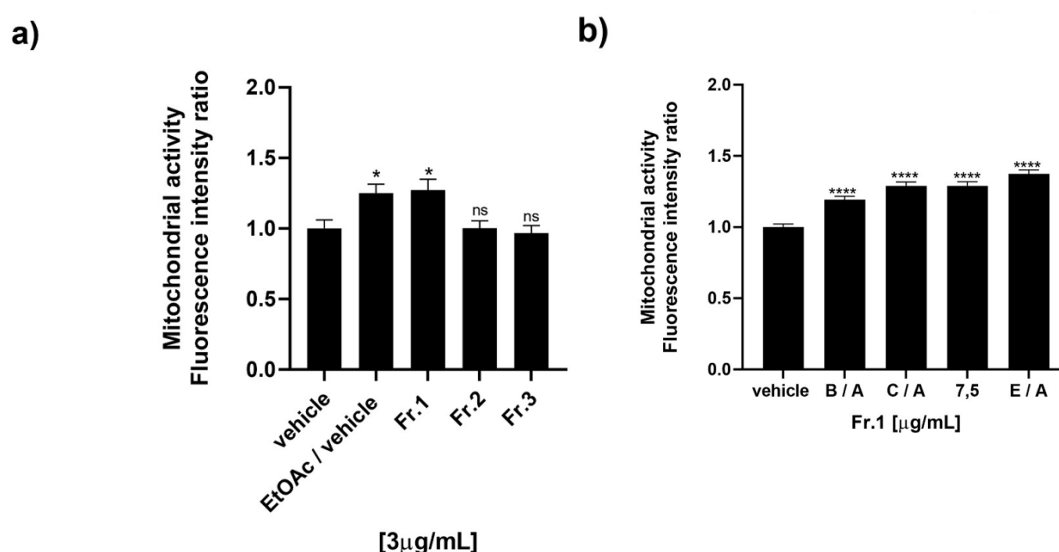


Figure 2. 2 Oleanolic acid enriched fractions promote mitochondrial metabolism in C2C12 cells. **a)** Mito-tracker Normalized Fluorescence (see methods for details) measured upon incubation of C2C12 cells with crude extract (crude) or the indicated fractions at the concentration of 3 µg/mL for 48 h. **b)** Mito-tracker Normalized Fluorescence measured upon incubation of C2C12 cells with Fr.1 tested at the indicated concentrations for 48 h. Values are expressed as mean ± S.D. of three independent replicates and were considered statistically different when p value were * < 0.01, *** < 0.001, **** < 0.0001; and considered not statistically (ns) different when p value was > 0.05.

2.5 Effect of oleanolic acid enriched fractions on glucose uptake activity of in-vitro cultured C2C12 myoblast

Once assessed the influence of oleanolic acid on the mitochondrial activities of in-vitro cultured C2C12 myoblast, we moved to test whether crude extract of Aglianico pomace and fractions 1, 2 and 3 were able to promote glucose uptake via membrane glucose transporters (GLUT). Dysmetabolic disorders, which are frequently characterized by insulin resistance and hyperglycemia, have impaired control over the handling of circulating glucose levels. In this context, the GLUT class of membrane transporters plays an essential role to remove glucose from the bloodstream. The majority of these carrier proteins are controlled by insulin and react to insulin stimulation via PI3K signaling. Insulin stimulation can increase the number of GLUT receptors (mostly GLUT4) on the plasma membrane of cells, boost glycolysis, and ultimately enhance the activity of passive transporters (GLUT1) by reducing intracellular glucose levels. Insulin resistance impairs GLUT transporter function, and pharmacological stimulation of these transporters enhances glucose tolerance and reduces hyperglycaemia. The ability of crude extract and fractions to modulate glucose uptake was measured by monitoring the

uptake of NBDG, a fluorescent derivative of deoxyglucose covalently bound to the fluorescent chemical nitro blue tetrazolium (NBT). The increase in fluorescence of C2C12 cells incubated with NBDG correlates with the activity of GLUT transporters (**Figure 2.3**). C2C12 cells were treated with 3 $\mu\text{g/mL}$ of the EtOAc extract and fractions 1, 2 and 3 for 48 h. At the end of the incubation, NBDG was added to the culture medium. Compared to vehicle (DMSO 0.1%), cells treated with the EtOAc extract and the oleanolic acid rich-fraction 1 showed an increase in NBDG uptake. Thus, modulation of glucose uptake through stimulation of GLUT transporters can be a parallel mechanism that contributes to the beneficial effects of oleanolic acid on diabetes and metabolic syndrome.

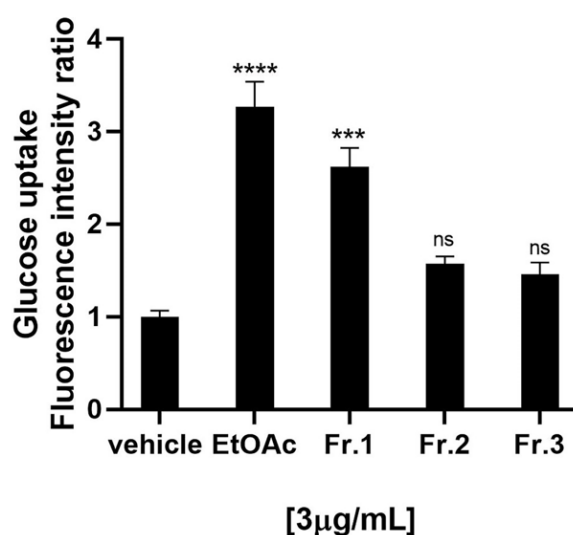


Figure 2. 3 Oleanolic acid enriched fractions promote glucose uptake in C2C12 cells. 2-NBDG Normalized Fluorescence (see methods for details) measured upon incubation of C2C12 cells with crude extract (crude) or the indicated fractions at the concentration of 3 $\mu\text{g/mL}$.

2.6 Conclusions

For the first time, an in-depth chemical analysis of Aglianico grape pomace was conducted by means of different chromatographic (column chromatography, UV–Vis HPLC) spectrometric (HRMS, LC-MS/MS) and spectroscopic (NMR) techniques. This investigation disclosed the presence of a number of bioactive compounds, mainly polyphenols. Among the several anthocyanins commonly occurring in red grapes and wines, only the coumaroyl glucoside derivatives of malvidin and peonidin were found, probably due to their higher hydrophilicity compared to the abundant parent compounds. However, the most relevant outcome of the

reported analysis was the isolation in Aglianico pomace of a remarkable amount of oleanolic acid, in the order of 0.45 mg per gram of pomace (fresh weight). This allowed us to verify our initial hypothesis that additional major bioactive metabolites along with typical grape polyphenols are responsible for the reported bioactivities of grape pomace extracts. In fact, oleanolic acid is endowed with a wide selection of biological properties and lately it has been attracting interest for its antidiabetic effects. Our results on the mitochondrial and glucose uptake activities of oleanolic acid on in-vitro cultured C2C12 myoblasts further supported the promotion of this triterpenoid as a promising antidiabetic molecule. More generally, our results clearly demonstrated the utility of an untargeted chemical analysis of food materials and of grape pomace in particular. Indeed, it can bring to light the occurrence of possibly undetected bioactive compounds occurring even in remarkable quantities, such as oleanolic acid that would have been overlooked in by a targeted approach focused on polyphenols. In perspective, the presented data are expected to prompt further studies aimed at the optimization of extraction and recovery of healthy molecules, and in particular of triterpenoids, from grape pomace by means of sustainable protocols in the frame of circular economy.

2.7 Experimental Section

2.7.1 Chemicals and reagents

All solvents used in this study were of HPLC grade or higher. Malvidin-3-O-glucoside, (+)-catechin, (-)-epicatechin, caffeic acid, quercetin, *trans*-resveratrol, gallic acid, formic acid, acetonitrile, ethanol (EtOH), methanol (MeOH), *n*-hexane, dichloromethane, butanol, ethyl acetate (EtOAc), boric acid, amino acids standards, sodium azide, and diethylethoxymethylenemalonate (DEEMM) were purchased from Sigma-Aldrich (Milan, Italy). Aqueous solutions were prepared with Milli-Q water from Millipore (Bedford, MA, USA).

2.7.2 Grape pomace collection, extraction and chromatographic separation

Aglianico (*Vitis vinifera* L. cv) grape pomace was obtained from the Taurasi area - Avellino, Italy in late November of 2021. Fresh pomace (2000 g), constituted by skins, seeds, stalks and residual pulp, was collected after the pressing stage to obtain a homogenous sample. 1480 g of fresh pomace were immediately lyophilized, thus affording 418 g of dry biological material. From this quantity of dry pomace, 99.48 g were subjected to extraction as described below.

The lyophilized material was grounded by a laboratory blender and extracted twice at room temperature (to avoid degradation of the extracted compounds) with 1.0 L of H₂O:EtOH 2:8 (v/v) overnight. The use of acidified solvents was avoided, since in literature it is reported that either mineral or organic acids should not be used in order to reduce the risk of degrading anthocyanin derivatives or producing anthocyanidins from flavanols and proanthocyanidins (Revilla et al., 1998). After extractions, the dry weight was 91.54 g. The hydroethanolic extract was concentrated to remove residual water and partitioned twice in the order against *n*-hexane 200 mL, EtOAc 300 mL, and *n*-butanol 300 mL. This order of partitioning solvents allowed the extraction of compounds with increasing polarity. Typically, hexane extracts fatty acids (if present) and apolar terpenoids, EtOAc extracts compounds with medium polarity and butanol extracts polar organic compounds. More specifically, three different extracts termed fractions A (*n*-hexane phase), fraction B (EtOAc phase) and fraction C (*n*-butanol phase), respectively, were obtained. All of the obtained phases were fully evaporated under reduced pressure at 30 °C using a rotary evaporator. The final dry weight of the *n*-hexane, EtOAc and *n*-butanol extracts were 0.296 g, 4.95 g, and 2.496 g, respectively. The EtOAc extract was further separated through a 80-g silica Combiflash column (3.0 mL/min flow) connected to a Teledyne Isco CombiFlash Rf flash chromatography system, eluted with the following gradient elution: dichloromethane:EtOAc 1:1 (200 mL; fraction 1), 100% EtOAc (200 mL; fraction 2), EtOAc:MeOH 9:1 (200 mL; fraction 3), EtOAc:MeOH 7:3 (400 mL; fraction 4).

2.7.3 LC-MS/MS analyses

LC-HR MS experiments were performed on a hybrid linear ion trap LTQ-Orbitrap XLTM Fourier transform mass spectrometer equipped with an ESI ION MAXTM source (Thermo Fisher, San José, USA) coupled to an Agilent 1100 LC binary system (Palo Alto, USA). Chromatographic separations were achieved using a Kinetex Polar C18 column (100 × 3.0 mm, 100 Å, 2.6 µm). The injection volume was 0.5 mL/min and the mobile phase consisted of a combination of A (0.1% formic acid in water, v/v) and B (MeOH). For fraction 2, the chromatographic method lasted 25 min and the gradient design was as follows: 50% B (3 min), from 3 to 15 min B reached 95%, and hold at 95% for 5 min. For fractions 3 and 4, the chromatographic method lasted 25 min and the gradient design was as follows: an initial 3 min at 10% B, then from 3 to 20 min B reached 95%, which was hold for 5 min. The analysis was carried out in positive ion mode (ESI +). Source conditions were: spray voltage: 3.5 kV (positive mode) and 2.9 kV (negative mode);

capillary voltage: 25 V; source temperature: 320°C; normalized collision energy: 25. The acquisition range was m/z 150–1500. Quantitation of analytes was carried out by analyzing extracted ion chromatograms (XICs) relative to each compound to be quantified with 5 ppm mass tolerance. As standards, (+)-catechin, (-)-epicatechin, caffeic acid, quercetin, *trans*-resveratrol, gallic acid were used. For each of these standards, a calibration curve was plotted on the basis of peak areas (triplicate injections) obtained by using six different concentrations (0.1, 0.5, 1.0, 5.0, 10.0 and 20.0 mg/L). The following calibration curve equations were obtained:

(+)-catechin: $y = 2.36e-6x - 0.0032$ ($R^2 = 0.99$; LOD = 0.025 $\mu\text{g}/\mu\text{L}$; LOQ = 0.0850 $\mu\text{g}/\mu\text{L}$);

(-)-epicatechin: $y = 2.02e-6x - 0.0004$ ($R^2 = 0.99$; LOD = 0.022 $\mu\text{g}/\mu\text{L}$; LOQ = 0.0076 $\mu\text{g}/\mu\text{L}$);

caffeic acid: $y = 0.00344x - 0.1970$ ($R^2 = 0.99$; LOD = 0.0007 $\mu\text{g}/\mu\text{L}$; LOQ = 0.0025 $\mu\text{g}/\mu\text{L}$);

quercetin: $y = 9.41e-5x - 0.0076$ ($R^2 = 0.99$; LOD = 0.0023 $\mu\text{g}/\mu\text{L}$; LOQ = 0.0076 $\mu\text{g}/\mu\text{L}$);

trans-resveratrol: $y = 3.73e-5x + 0.0026$ ($R^2 = 0.99$; LOD = 0.0199 $\mu\text{g}/\mu\text{L}$; LOQ = 0.0663 $\mu\text{g}/\mu\text{L}$);

gallic acid: $y = 0.0015x - 0.0087$ ($R^2 = 0.99$; LOD = 0.0012 $\mu\text{g}/\mu\text{L}$; LOQ = 0.0054 $\mu\text{g}/\mu\text{L}$).

The remaining analytes for which a standard was not used, on the basis of chemical similarity, it was assumed they shared the same molar response as the standard representative of the chemical class. Specifically, (+)-catechin was used for quantifying catechin-3-*O*-gallate and procyanidin dimers and trimers; (-)-epicatechin was used for quantifying epicatechin-3-*O*-gallate; quercetin was used for quantifying myricetin; caffeic acid was used for quantifying caffeoyltartaric acid; gallic acid was used for quantifying syringic acid; and *trans*-resveratrol was used for quantifying resveratrol dimer and trimer.

2.7.4 NMR analysis

^1H (600 and 700 MHz) NMR spectra were measured on a Bruker spectrometer by using a Norell Select Series 5 mm NMR tubes. Chemical shifts were referenced to the residual solvent signal (CDCl_3 : δ_{H} 7.26, δ_{C} 77.0; CD_3OD : δ_{H} 3.31, δ_{C} 49.3 ppm). To quantitation purposes 5 μL of pyridine as a standard were added to the NMR tube containing the compound to quantitate. ^1H NMR spectrum of the obtained solution was acquired with a d1 value of 7.0 sec. Representative NMR signals were chosen on the basis of resolution and their area was measured by integration, divided by the number of the protons generating them, and finally converted into

the relative number of moles by comparing the obtained area with that of the pyridine protons, for which the number of moles was known.

2.7.5 High-Performance liquid chromatography analyses of anthocyanins

The separation of the monomeric anthocyanins, contained in the initial hydroethanolic extract of grape pomace, was performed according to the OIV method of analysis (OIV, 2020) by using an HPLC Agilent 1260 infinity II LC apparatus (Santa Clara, USA) consisting of a binary pump, two-channel degasser unit, a diode array detector-DAD (G7114A), and a manual injector equipped with a 20- μ L loop. A column heating device set at 40°C, with a C18 column, Synergi 4 μ m Fusion – RP 80 (250 $\times$ 3.0 mm, 4 μ m particles diameter) with a pre-column was used. All the samples were filtered through 0.45- μ m filters (Durapore membrane filters, Millipore-Ireland) into glass vials and immediately injected into the HPLC system. Detection was performed by monitoring the absorbance signals at 518 nm. Data collection and analyses were performed through the software OpenLAB CDS ChemStation Edition (Agilent Technologies, Santa Clara, USA). Elution was carried out by using a flow rate of 0.80 mL/min. Eluents were as follows: A consisted of water milli-Q (Sigma Aldrich)/formic acid (Sigma Aldrich $\geq$ 95%)/acetonitrile (Sigma Aldrich $\geq$ 99.9%) (96:1:3) v/v/v, and B consisted of water/formic acid/acetonitrile (49:1:50) v/v/v. The following gradient was used: initial conditions 94% A and 6% B, after 15 min the pumps were adjusted to 70% A and 30% B, at 30 min to 50% A and 50% B, at 35 min to 40% A and 60% B, at 41 min through the end of analysis, to 94% A and 6% B (Forino et al., 2022). For calibration, the external standard method was used: the calibration curve was plotted for the malvidin-3- O-glucoside on the basis of peak area. Concentrations were expressed as mg/L of malvidin-3-O-glucoside. All experiments were carried out in duplicate and two analytical replicas were performed. The following calibration curve equation was obtained: $y = 5E-06x$; $R^2 = 0.999$; LOQ = 2 mg/L; LOD = 0.6 mg/L.

2.7.6 High-Performance liquid chromatography analyses of amino acids

Amino acids were determined by the methods described by Gomez- Alonso et al. (2007) and by Ortega-Heras et al. (2014) with some modifications as reported below. 10-gram dry pomace was extracted with 130.0 mL of a mixture of acetonitrile/MeOH/ultra-pure water (35:35:30, v/v/v) under a 30-min stirring. The obtained extract was filtered through filter paper and derivatized according to the reported method. In more detail, aminoenone derivatives of amino

acids were obtained by reaction of 1.75 mL of a borate 1 M buffer solution, 750 μ L of MeOH, 1 mL of the grape pomace extract, and 30 μ L of DEEMM in a screw-cap test tube. The mixture was homogenized and placed in an ultrasound bath for 30 min. The sample was then heated at 70°C for 2 h to eliminate the excess DEEMM and other reagent products. Before the injection, the samples were filtered through 0.45 μ m filters (Durapore membrane filters, Millipore-Ireland) into a glass vial and injected into the HPLC system. Analyses were performed with the HPLC apparatus previously described. The chromatographic separation was carried out on a reverse-phase Agilent Infinity Lab Poroshell 120 EC-C18 column (3.0 $\times$ 150 mm- 2.7 μ m). The solvents were (A) 25 mM acetate buffer (pH 5.8) with 0.02% sodium azide and (B) acetonitrile/methanol (80:20). The gradient used was: 100% solvent A, after 20 min the pumps were adjusted to 95% A and 5% B, at 30.5 min 90% A and 5 % B, at 33.5 min 83% A and 17% B, at 65 min 60% A and 40% B, at 73 min 28% A and 72% B, at 78 min 18% A and 82% B, at 80 min until the end of analysis 100% B. The flow rate was 0.5 mL/min. After a 10 min equilibrium period, the next sample was injected. For calibration, the external standard method was used: the calibration curve was plotted for L-2-aminoadipic acid on the basis of peak area (Gomez-Alonso et al., 2007). The following calibration curve equation was obtained: $y = 0.0045x + 0.7612$ ($R^2 = 0.99$; LOD = 0.00013 μ g/ μ L; LOQ = 0.00039 μ g/ μ L). Concentrations were expressed as mg/L. All experiments were carried out in duplicate and two analytical replicas were performed, in order to assess the variability of results.

2.7.7 Measurement of mitochondrial activity

Mitochondrial activity of the EtOAc extract and of the fractions deriving from its purification was measured in C2C12 cells using the MitoTracker® Red CMXRos probe (Thermo Fisher Scientific), as already described (Riccio et al., 2018). Briefly, the probe was diluted in DMEM to a final dye working solution of 100 nM. C2C12 were washed twice in PBS before being incubated with the probe for 60 min in a cell incubator set to 37°C and 5% CO₂. Upon incubation, cells were washed three times in DMEM and once in PBS to be then fixed in 3.7% formaldehyde for 30 min. MitoTracker fluorescence was measured using a Perkin Elmer Envision 2105 Multiplate Reader (Perkin Elmer) with the following parameters: λ excitation 579 nm, λ emission 599 nm. Cells were then permeabilized in 0.1% Triton X-100 in PBS, and stained with the nuclear dye DAPI. DAPI fluorescence was measured with λ excitation 351 nm, and emission 450 nm. MitoTracker

fluorescence was normalized with DAPI fluorescence in order to correlate mitochondrial activity to total number of cells.

2.7.8 Measurement of glucose uptake

The glucose uptake activity was evaluated for the EtOAc extract and for the fractions deriving from its purification. C2C12 cells were seeded (5 × 10³/well) on a 96-well black bottom microtiter plate with (Perkin Elmer, Waltham, USA) in a final volume of 100 µL/well of growth media. When cells achieved 80–90% of confluency, the culture media was carefully removed and replaced with 100 µL of HBSS containing 100 µM, 2-Deoxy-Glucose (2-DG), 0.4 g/L BSA, and 1.27 mM CaCl₂ (without any growth factors or FBS) and when indicated the compounds to be tested, as already described (Lapi et al., 2020). Cells were incubated for one hour at 37°C, the reported temperature to achieve the better vitality of cells. The same HBSS was then further supplemented with 64 µM the fluorescent probe 2-NDBG. After 45 min of incubation, plates were washed twice in PBS. Uptake of 2-NDBG was determined using Perkin Elmer Envision 2105 Multiplate Reader (Perkin Elmer) with the following parameters: λ excitation at 471 nm, λ emission at 529 nm, and monochromator cut off at 360 nm. Following the determination of 2-NDBG, cells were fixed in 3.7% paraformaldehyde for 30 min, permeabilized in 0.1% Triton X-100 in PBS, and stained with the nuclear dye DAPI (100 µM). DAPI fluorescence was measured with the following parameters: λ excitation 351 nm and λ emission 450 nm. Normalized refers to the ratio between intracellular 2-NDBG fluorescence and DAPI fluorescence.

2.7.9 Statistical analysis

For pharmacological studies, comparisons and differences were analyzed for statistical significance by one-way ANOVA and Bonferroni multiple comparisons test with a single pooled variance. Statistical analysis was performed using GraphPad Prism (GraphPad Software 7.03, San Diego, USA). Values were considered statistically different when p value were * < 0.01, *** < 0.001, **** < 0.0001; and considered not statistically (ns) different when p value was > 0.05.

References

- Annunziata, G., Sanduzzi Zamparelli, M., Santoro, C., Ciampaglia, R., Stornaiuolo, M., Tenore, G.C., Sanduzzi, A. and Novellino, E., **2020**. May polyphenols have a role against coronavirus infection? An overview of in vitro evidence. *Frontiers in medicine*, 7, p.240.
- Arapitsas, P., Ugliano, M., Marangon, M., Piombino, P., Rolle, L., Gerbi, V., Versari, A. and Mattivi, F., **2020**. Use of untargeted liquid chromatography–mass spectrometry metabolome to discriminate Italian monovarietal red wines, produced in their different terroirs. *Journal of Agricultural and Food Chemistry*, 68(47), pp.13353-13366.
- Badolati, N., Masselli, R., Sommella, E., Sagliocchi, S., Di Minno, A., Salviati, E., Campiglia, P., Dentice, M., Tenore, G.C., Stornaiuolo, M. and Novellino, E., **2020**. The hepatoprotective effect of taurisolo, a nutraceutical enriched in resveratrol and polyphenols, involves activation of mitochondrial metabolism in mice liver. *Antioxidants*, 9(5), p.410.
- Calderón-Oliver, M. and Ponce-Alquicira, E., **2021**. Environmentally friendly techniques and their comparison in the extraction of natural antioxidants from green tea, rosemary, clove, and oregano. *Molecules*, 26(7), p.1869.
- Campos, F., Peixoto, A.F., Fernandes, P.A., Coimbra, M.A., Mateus, N., de Freitas, V., Fernandes, I. and Fernandes, A., **2021**. The antidiabetic effect of grape pomace polysaccharide-polyphenol complexes. *Nutrients*, 13(12), p.4495.
- Caponio, G.R., Cofano, M., Lippolis, T., Gigante, I., De Nunzio, V., Difonzo, G., Noviello, M., Tarricone, L., Gambacorta, G., Giannelli, G. and De Angelis, M., **2022**. Anti-proliferative and pro-apoptotic effects of digested aglianico grape pomace extract in human colorectal cancer cells. *Molecules*, 27(20), p.6791.
- Carino, A., Moraca, F., Fiorillo, B., Marchianò, S., Sepe, V., Biagioli, M., Finamore, C., Bozza, S., Francisci, D., Distrutti, E. and Catalanotti, B., **2020**. Hijacking SARS-CoV-2/ACE2 receptor interaction by natural and semi-synthetic steroidal agents acting on functional pockets on the receptor binding domain. *Frontiers in Chemistry*, 8, p.572885.
- Castellano, J.M., Ramos-Romero, S. and Perona, J.S., **2022**. Oleanolic acid: extraction, characterization and biological activity. *Nutrients*, 14(3), p.623.
- Cecchi, L., Schuster, N., Flynn, D., Bechtel, R., Bellumori, M., Innocenti, M., Mulinacci, N. and Guinard, J.X., **2019**. Sensory profiling and consumer acceptance of pasta, bread, and

- granola bar fortified with dried olive pomace (pâté): A byproduct from virgin olive oil production. *Journal of food science*, 84(10), pp.2995-3008.
- Covas, M.I., Miró-Casas, E., Fitó, M., Farré-Albadalejo, M., Gimeno, E., Marrugat, J. and De La Torre, R., **2003**. Bioavailability of tyrosol, an antioxidant phenolic compound present in wine and olive oil, in humans. *Drugs under experimental and clinical research*, 29(5-6), pp.203-206.
- D'Onofrio, N., Martino, E., Chianese, G., Coppola, F., Picariello, L., Moio, L., Balestrieri, M.L., Gambuti, A. and Forino, M., **2021**. Phenolic profiles of red wine relate to vascular endothelial benefits mediated by SIRT1 and SIRT6. *International Journal of Molecular Sciences*, 22(11), p.5677.
- Drevelegka, I. and Goula, A.M., **2020**. Recovery of grape pomace phenolic compounds through optimized extraction and adsorption processes. *Chemical Engineering and Processing-Process Intensification*, 149, p.107845.
- Dwyer, K., Hosseinian, F. and Rod, M., **2014**. The market potential of grape waste alternatives. *Journal of Food Research*, 3(2), pp.91-106.
- Food and Agriculture Organization of the United Nations. **2019**. The state of food and agriculture. *2019, Moving forward on food loss and waste reduction*.
- Forino, M., Gambuti, A. and Moio, L., **2019**. NMR-based systematic analysis of bioactive phytochemicals in red wine. First determination of xanthurenic and oleanic acids. *Food chemistry*, 278, pp.497-501.
- Forino, M., Cassiano, C., Gambuti, A., Picariello, L., Aversano, R., Villano, C., Basile, B., Moio, L. and Frusciante, L., **2022**. Aging Behavior of Two Red Wines from the PIWI Pathogen-Resistant Grapevines "Cabernet Eidos "and "Merlot Khorus ". *ACS Food Science & Technology*, 2(4), pp.638-646.
- Farzan, M., Farzan, M., Shahrani, M., Navabi, S.P., Vardanjani, H.R., Amini-Khoei, H. and Shabani, S., **2024**. Neuroprotective properties of Betulin, Betulinic acid, and Ursolic acid as triterpenoids derivatives: A comprehensive review of mechanistic studies. *Nutritional neuroscience*, 27(3), pp.223-240.
- Galanakis, C.M., Tsatalas, P. and Galanakis, I.M., **2018**. Implementation of phenols recovered from olive mill wastewater as UV booster in cosmetics. *Industrial Crops and Products*, 111, pp.30-37.

- García-Lomillo, J. and González-SanJosé, M.L., **2017**. Applications of wine pomace in the food industry: Approaches and functions. *Comprehensive reviews in food science and food safety*, 16(1), pp.3-22.
- Gómez-Alonso, S., Hermosín-Gutiérrez, I. and García-Romero, E., **2007**. Simultaneous HPLC analysis of biogenic amines, amino acids, and ammonium ion as aminoenone derivatives in wine and beer samples. *Journal of agricultural and food chemistry*, 55(3), pp.608-613.
- Gonzalez-SanJose, M. L. **2014**. *The Oxford handbook of food fermentations* (pp. 304–342). Oxford Univ. Press.
- Huang, Q., Chen, H., Ren, Y., Wang, Z., Zeng, P., Li, X., Wang, J. and Zheng, X., **2016**. Anti-hepatocellular carcinoma activity and mechanism of chemopreventive compounds: Ursolic acid derivatives. *Pharmaceutical Biology*, 54(12), pp.3189-3196.
- Iqbal, A., Schulz, P. and Rizvi, S.S., **2021**. Valorization of bioactive compounds in fruit pomace from agro-fruit industries: Present Insights and future challenges. *Food Bioscience*, 44, p.101384.
- Karković Marković, A., Torić, J., Barbarić, M. and Jakobušić Brala, C., **2019**. Hydroxytyrosol, tyrosol and derivatives and their potential effects on human health. *Molecules*, 24(10), p.2001.
- Kato-Schwartz, C.G., Corrêa, R.C.G., de Souza Lima, D., de Sá-Nakanishi, A.B., de Almeida Gonçalves, G., Seixas, F.A.V., Haminiuk, C.W., Barros, L., Ferreira, I.C., Bracht, A. and Peralta, R.M., **2020**. Potential anti-diabetic properties of Merlot grape pomace extract: An in vitro, in silico and in vivo study of α -amylase and α -glucosidase inhibition. *Food Research International*, 137, p.109462.
- Kühn, S., Düzel, S., Colzato, L., Norman, K., Gallinat, J., Brandmaier, A.M., Lindenberger, U. and Widaman, K.F., **2019**. Food for thought: association between dietary tyrosine and cognitive performance in younger and older adults. *Psychological research*, 83, pp.1097-1106.
- Lapi, D., Stornaiuolo, M., Sabatino, L., Sommella, E., Tenore, G., Daglia, M., Scuri, R., Di Maro, M., Colantuoni, A. and Novellino, E., **2020**. The pomace extract taurisolo protects rat brain from ischemia-reperfusion injury. *Frontiers in Cellular Neuroscience*, 14, p.3.

- Lesellier, E., Lefebvre, T. and Destandau, E., **2021**. Recent developments for the analysis and the extraction of bioactive compounds from *Rosmarinus officinalis* and medicinal plants of the Lamiaceae family. *TrAC Trends in Analytical Chemistry*, 135, p.116158.
- Li, L. and Sun, B., **2019**. Grape and wine polymeric polyphenols: Their importance in enology. *Critical reviews in food science and nutrition*, 59(4), pp.563-579.
- Li, W., Zeng, H., Xu, M., Huang, C., Tao, L., Li, J., Zhang, T., Chen, H., Xia, J., Li, C. and Li, X., **2021**. Oleanolic acid improves obesity-related inflammation and insulin resistance by regulating macrophages activation. *Frontiers in Pharmacology*, 12, p.697483.
- Li, Y., Zeng, Q., Wang, R., Wang, B., Chen, R., Wang, N., Lu, Y., Shi, F., Dehaen, W. and Huai, Q., **2022**. Synthesis and discovery of mitochondria-targeting oleanolic acid derivatives for potential PI3K inhibition. *Fitoterapia*, 162, p.105291.
- Lin, C., Wen, X. and Sun, H., **2016**. Oleanolic acid derivatives for pharmaceutical use: A patent review. *Expert opinion on therapeutic patents*, 26(6), pp.643-655.
- OIV Organisation Internationale de la Vigne et du Vin. Compendium of International Methods of wine And Must Analysis. Available online: <https://www.oiv.int> (accessed on 23 November **2022**).
- Ortega-Heras, M., Pérez-Magariño, S., Del-Villar-Garrachón, V., González-Huerta, C., Moro Gonzalez, L.C., Guadarrama Rodríguez, A., Villanueva Sanchez, S., Gallo González, R. and Martín de la Helguera, S., **2014**. Study of the effect of vintage, maturity degree, and irrigation on the amino acid and biogenic amine content of a white wine from the Verdejo variety. *Journal of the Science of Food and Agriculture*, 94(10), pp.2073-2082.
- Pateiro, M., Gómez-Salazar, J.A., Jaime-Patlán, M., Sosa-Morales, M.E. and Lorenzo, J.M., **2021**. Plant extracts obtained with green solvents as natural antioxidants in fresh meat products. *Antioxidants*, 10(2), p.181.
- Peixoto, C.M., Dias, M.I., Alves, M.J., Calhelha, R.C., Barros, L., Pinho, S.P. and Ferreira, I.C., **2018**. Grape pomace as a source of phenolic compounds and diverse bioactive properties. *Food chemistry*, 253, pp.132-138.
- Pensec, F., Paczkowski, C., Grabarczyk, M., Woźniak, A., Bénard-Gellon, M., Bertsch, C., Chong, J. and Szakiel, A., **2014**. Changes in the triterpenoid content of cuticular waxes during fruit ripening of eight grape (*Vitis vinifera*) cultivars grown in the Upper Rhine Valley. *Journal of agricultural and food chemistry*, 62(32), pp.7998-8007.

- Revilla, E., Ryan, J.M. and Martin-Ortega, G., **1998**. Comparison of several procedures used for the extraction of anthocyanins from red grapes. *Journal of Agricultural and Food Chemistry*, 46(11), pp.4592-4597.
- Riccio, G., Sommella, E., Badolati, N., Salviati, E., Bottone, S., Campiglia, P., Dentice, M., Tenore, G.C., Stornaiuolo, M. and Novellino, E., **2018**. Annurca apple polyphenols protect murine hair follicles from taxane induced dystrophy and hijacks polyunsaturated fatty acid metabolism toward β -oxidation. *Nutrients*, 10(11), p.1808.
- Sharma, Y., Velamuri, R., Fagan, J. and Schaefer, J., **2020**. Full-spectrum analysis of bioactive compounds in rosemary (*Rosmarinus officinalis* L.) as influenced by different extraction methods. *Molecules*, 25(20), p.4599.
- Soceanu, A., Dobrinas, S., Sirbu, A., Manea, N. and Popescu, V., **2021**. Economic aspects of waste recovery in the wine industry. A multidisciplinary approach. *Science of the Total Environment*, 759, p.143543.
- Unusan, N., **2020**. Proanthocyanidins in grape seeds: An updated review of their health benefits and potential uses in the food industry. *Journal of Functional Foods*, 67, p.103861.
- Wang, S., Du, L.B., Jin, L., Wang, Z., Peng, J., Liao, N., Zhao, Y.Y., Zhang, J.L., Pauluhn, J., Hai, C.X. and Wang, X., **2018**. Nano-oleanolic acid alleviates metabolic dysfunctions in rats with high fat and fructose diet. *Biomedicine & Pharmacotherapy*, 108, pp.1181-1187.
- Wang, X., Bai, H., Zhang, X., Liu, J., Cao, P., Liao, N., Zhang, W., Wang, Z. and Hai, C., **2013**. Inhibitory effect of oleanolic acid on hepatocellular carcinoma via ERK-p53-mediated cell cycle arrest and mitochondrial-dependent apoptosis. *Carcinogenesis*, 34(6), pp.1323-1330.
- Zhu, M., Huang, Y., Wang, Y., Shi, T., Zhang, L., Chen, Y. and Xie, M., **2019**. Comparison of (poly) phenolic compounds and antioxidant properties of pomace extracts from kiwi and grape juice. *Food Chemistry*, 271, pp.425-432.
- Zielonka, J., Joseph, J., Sikora, A., Hardy, M., Ouari, O., Vasquez-Vivar, J., Cheng, G., Lopez, M. and Kalyanaraman, B., **2017**. Mitochondria-targeted triphenylphosphonium-based compounds: syntheses, mechanisms of action, and therapeutic and diagnostic applications. *Chemical reviews*, 117(15), pp.10043-10120.

PART 4: Chemical profiling of some antidiabetic dietary plants

This part of the thesis describes two tropical food plants, both with anti-diabetic potential, that have been analyzed for their chemical composition, by using a combination of analytical methods.

The first chapter presents the results obtained of *Piper borbonense*, a nearly uninvestigated species of the genus *Piper*, whose chemical profile has been disclosed through the application of untargeted metabolomic approach. This project was carried out in collaboration with Professor Giovanni Appendino at the University of Eastern Piedmont and at the École Polytechnique at “*Laboratoire de Chimie Moléculaire*” (LCM) during my visiting period, where I had the opportunity to deepen the MZmine and MetGem workflows and to perform HS-GCMS analysis.

The second chapter describes the preliminary results of my research at the Centre National de la Recherche Scientifique in at “*Institut de Chimie des Substances Naturelles*” (CNRS-ICSN) in Gif-sur-Yvette, where I spent six months under the supervision of Professor Véronique Eparvier, in collaboration with Professor David Touboul from the École Polytechnique. There, thanks to the collaboration of 'Herbier IRD de Guyane', I worked on the development of a chemotaxonomic classification procedure for the different species of the genus *Annona* using statistical analysis. I then optimised an LC-MS/MS method for the annotation of annonaceous acetogenins from *Annona* plants by performing post-column infusions with different salts to improve the annotation of this class of compounds.

CHAPTER 1: Untargeted phytochemical analysis of *Piper borbonense*

1.1 *Piper* species and their antidiabetic potential

Piper genus includes more than 700 species, being one of the largest in Piperaceae family. These plants commonly grow in tropical and subtropical regions, particularly in South and Central America, as well as in Southeast Asia. *Piper* plants exhibit a wide range of growth forms, including shrubs, small trees, and climbing vines. They typically have simple, alternate leaves, and their inflorescences are characterized by spikes with small flowers, while the fruits are usually small drupes as shown in **Figure 1.1** (Mgbeahuruike et al., 2017). Members of the genus *Piper* have been used extensively not only in culinary recipes, but also as perfumes and in traditional medicine and even rituals across different cultures. *Piper* plants are well known for their economic importance, particularly *Piper nigrum* which produces black pepper, one of the most widely used spices in the world, but they are also rich in bioactive compounds such as alkaloids, flavonoids and essential oils that have been studied for their pharmacological properties (Salehi et al., 2019).



Figure 1.1 Fruits from *Piper* species. A) *P. nigrum*; B) *P. longum*; C) *P. borbonense*.

For example, *P. cubeba* and *P. longum* have been evaluated for their antiproliferative effects (Graidist et al., 2015; Pandey et al., 2016), and in particular the role of the (pseudo)alkaloids piperine and piperlongumine in the definition of this effect has been examined in different cell lines (Bezerra et al., 2005). The alkaloid-rich extract of *Piper guineense* has shown antibacterial potential (Mgbeahuruike et al., 2018), and several species have been studied for their anti-

inflammatory and antioxidant activity (Lister et al., 2020; Seo et al., 2022). In line with our research interests, we were attracted by the literature reports on the antidiabetic properties of various *Piper* plants. For example, the methanolic extract of *P. sarmentosum* was evaluated in fructose-induced obese rats and showed the ability to suppress lipid accumulation and inhibit HMG-CoA, thereby reducing obesity and hyperlipidaemia (Kumar et al., 2021), while PipeNig®-FL, a liquid extract of black pepper, was patented for its ability to improve insulin sensitivity, promote glucose uptake and inhibit 3T3-L1 adipocyte differentiation (Geddo et al., 2019). The bioactive compounds of *Piper* species and their molecular targets have been analysed, revealing that (pseudo)alkaloids in particular exert anti-diabetic and anti-obesity effects. For example, retrofractamide analogues isolated from different *Piper* species, such as *P. chaba*, have shown anti-diabetic potential by acting on the expression of PPAR γ , GLUT4 and increasing the levels of adiponectin, a peptide commonly referred to as an obesity biomarker (Lu et al., 2008). In particular, although retrofractamide A has a higher EC₅₀ compared to glitazones, it can modulate PPAR γ with a different expression profile than the synthetic drugs and could therefore be a potential advantageous candidate as a PPAR γ agonist in the treatment of type 2 diabetes mellitus (Zhang et al., 2008). Interestingly, a similar bioactive profile has also been observed with piperlonguminine (Yamaguchi et al., 2014). Furthermore, trans-pellitorine can modulate preadipocyte differentiation and reduce lipid accumulation at low micromolar concentrations. The complex biomolecular mechanism behind this may involve TRPV1 and TRPA1 channels (Lieder et al., 2017), highlighting that many other alkaloids acting on these targets can attenuate metabolic syndrome (Okumura et al., 2010).

1.2 *Piper borbonense*

The traditional name for *P. borbonense* is Voatsipérifery, a word derived from the Malagasy word *voa*, meaning fruit, and *tsiperifery*, the local name for the plant on which it grows. This wild pepper is native to the islands of Réunion, Mauritius and Madagascar, where it is traditionally gathered. In fact, the plant is harvested entirely by hand by Malagasy communities whose villagers have to climb high vines to collect the fresh fruit. The rarity of this species and the traditional manual labour involved in its production is reflected in its high market price, \$200 per kilogram, around 10 times higher than that of the domesticated *P. nigrum* (Weil et al., 2020). Furthermore, *P. borbonense* is used in gourmet cooking in Europe, where it is exported, due to

its distinctive flavour (Razafimandimby et al., 2017). However, *P. borbonense* can be confused with other genetically related Malagasy wild peppers, such as *P. malgassicum* and *P. tsarasotrae*, which belong to the same clade (Palchetti et al., 2018). Species identification and authentication remain challenging, but the combination of botanical characteristics with essential oil and piperine content can help distinguish between them (Weil et al., 2021). *P. borbonense* has also been used in folk medicine in its native area. In fact, the leaves of this plant have been used to treat diarrhoea and for their diuretic effect in La Réunion (Fortin et al., 2002), while in the Comoros the branches are used to treat dental diseases, whereas the leaves are used for their anti-diabetic and anti-malarial properties, as well as to treat intestinal worms and diarrhoea (Soidrou et al., 2014). The traditional use of this plant has led to recent scientific evaluation of its medicinal properties. For example, phenolic extracts of *P. borbonense* were evaluated for immunomodulatory activity (Soidrou et al., 2014) and sesquiterpene-rich essential oils for fungicidal activity (Soidrou et al., 2013). Recently, its essential oil composition has been studied in greater detail, providing some insights into its peculiar aroma composition, characterised by low pungency and high aroma values compared to those of *P. nigrum* (Weil et al., 2017; Weil et al., 2020), while the profile of its major alkaloids has been briefly compared with that of other species (Luca et al., 2021). However, a detailed analysis of the secondary metabolites occurring in this species is still lacking. Thus, given our interest in chemically profiling plant extracts with anti-diabetic potential, we carried out a phytochemical characterization of *P. borbonense* using untargeted metabolomic approach.

1.3 Metabolomic profiling of *Piper borbonense* using LC-MS/MS-FBMN and LC-MSMS guided purification

The metabolomic fingerprint of *P. borbonense* was first achieved through a preliminary analysis of the crude acetone extract by LC coupled to HR-MS and feature-based molecular network (FB-MN) analysis (Nothias et al, 2020; Wang et al, 2016) using MZmine (Schmid et al, 2023) and MetGem open-source software (Olivon et al, 2018). The first software is used to pre-process the raw files, while MetGem compares MS/MS data, immediately visualizing the similarities among the spectra, which consequently cluster together according to their fragmentation pattern analogies. One of the nodes obtained in the cluster (**Figure 1.2**)

corresponded to piperine (**6**), parent ion at m/z 286.1433, which appear to be the main component of the mixture from the FB-MN and LC-UV-MS² analysis (**Figure 1.3**).

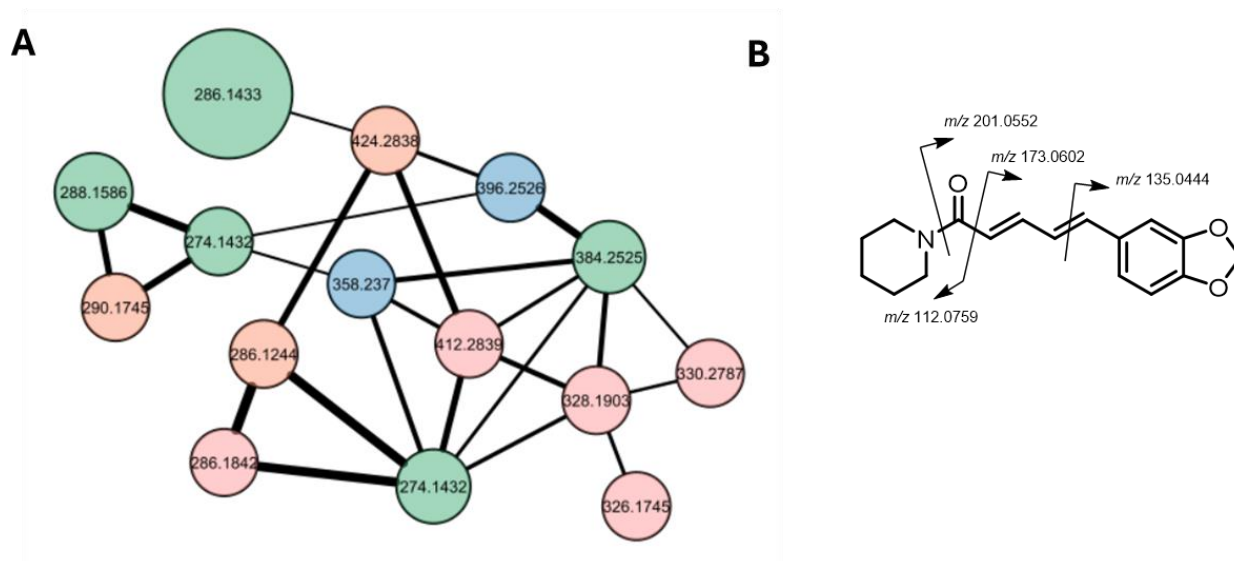


Figure 1.2 A) Molecular networking cluster of piperamides. Nodes are labeled with the parent m/z ratio and sized according to the peak area; edge thickness is related to the cosine similarity score. In green: the isolated metabolites; in blue: the annotated metabolites; in pink: not annotated. B) Piperine structure and its diagnostic detected fragments.

The detected fragment ions correspond to the loss of the piperidinyl moiety (m/z 201.0552), the alpha cleavage to the amide group (m/z 173.0602) and the related fragment with the opposite charge retention (m/z 112.0759), and finally the loss of the benzodioxole moiety (m/z 135.0444), in agreement with the data reported in the literature (Luca et al., 2021). Considering the recurrence of m/z 135.0444 fragment in the cluster, we assumed that also the other nodes belong to benzodioxole piperamides. Consequently, we assigned the other nodes to guineesine (**16**) (m/z 384.2525), clustered with m/z 396.2526, tentatively identified as 13-(1,3-benzodioxol-5-yl)-1-(1-piperidinyl) and to piperchambamide D (m/z 358.2371). This latter clustered together with two isomers with different retention time and size (nodes at m/z 274.1432), one assigned to 4,5-dihydropiperyline (**1**), which closely clustered with piperanine (**3**) (m/z 288.1586), and the other confidently assigned to piperlonguminin (**5**) given the different fragmentation pattern.

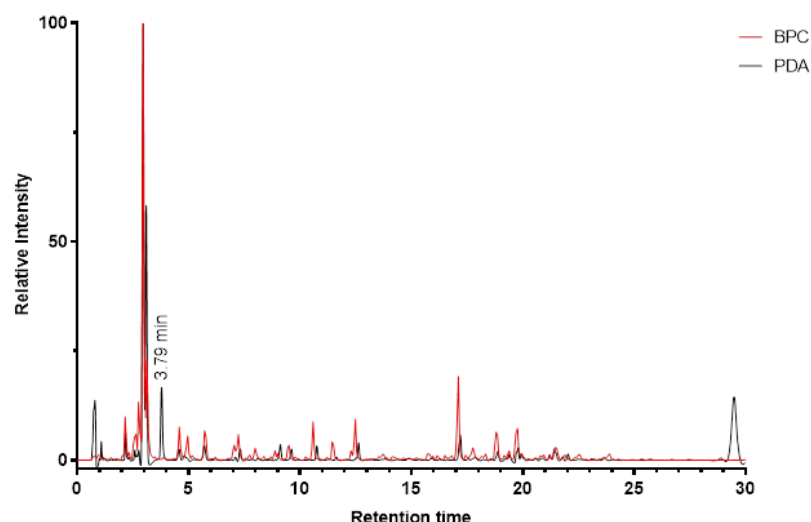


Figure 1. 3 Overlay BPC (red) and PDA (red) chromatograms of *P. borbonense* acetonic extract.

This preliminary information guided a detailed manual analysis of the LC-HRMSMS, resulting in the putative annotation of 36 piperamides (**Table 1.1**) thanks to their molecular formula and fragmentation pattern analysis that matched those reported in literature (Luca et al., 2021). 19 of them were later isolated and unambiguously identified with ESI-HRMS and ^1H NMR analysis.

Table 1.1 Annotated compounds from LC-HRMS/MS; compounds are listed in order of elution; t.id. stands for tentative identification

Compound name	Molecular formula	t_R (min)	Precursor ion (m/z)	Δ (ppm)	Fragments (m/z)	Identification criteria
4,5-Dihydropiperylin e (1)	$\text{C}_{16}\text{H}_{19}\text{NO}_3$	2.03	274.1430	- 2.663	203.0707; 175.0758; 161.0601; 135.0443	MN -isolation
Piperyline (4)	$\text{C}_{16}\text{H}_{17}\text{NO}_3$	2.14	272.1276	- 2.058	201.0551; 173.0602; 143.0494; 135.0443; 115.0503	MN -isolation
Dihydropiperlonguminine (2)	$\text{C}_{16}\text{H}_{21}\text{NO}_3$	2.58	276.1590	- 1.630	203.0704; 175.0757; 161.0599; 135.0441	isolation
Piperlonguminin (5)	$\text{C}_{16}\text{H}_{19}\text{NO}_3$	2.69	274.1432	- 1.897	201.0553; 175.0760; 145.0651; 135.0444; 115.0541	MN -isolation
Piperanine (3)	$\text{C}_{17}\text{H}_{21}\text{NO}_3$	2.81	288.1586	- 2.950	203.0709; 175.0761; 161.0602; 135.0444	isolation
Piperine (6)	$\text{C}_{17}\text{H}_{19}\text{NO}_3$	3.03	286.1427	- 3.739	201.0552; 173.0602; 135.0444; 112.0759	MN - isolation
7-(1,3-Benzodioxol-5-yl)-1-(1-	$\text{C}_{18}\text{H}_{21}\text{O}_3\text{N}$	3.20	300.1590	- 1.499	229.0863, 201.0914, 135.04442	MN - t.id.

PART 4

pyrrolidinyl)-2,4-heptadien-1-one						
Piperdardine (8)	C ₁₉ H ₂₃ NO ₃	4.75	314.1742	- 2.706	229.0867; 201.0915; 135.0443	MN -isolation
Sarmentine	C ₁₄ H ₂₃ NO	4.98	222.1849	- 1.624	151.1120; 133.1014; 124.0759; 95.0492	<i>t.id.</i>
Marginatine	C ₁₅ H ₂₀ O ₂	5.10	233.1533	- 1.357	215.1438, 187.1487; 177.0916; 133.1015	MN - <i>t.id.</i>
Pellitorine (18)	C ₁₄ H ₂₅ NO	6.00	224.2004	- 2.235	168.1358; 151.1119; 133.1013; 109.0648	Isolation
Neopellitorine B (19)	C ₁₅ H ₂₅ NO	7.23	236.2004	- 2.121	151.1120; 133.1013; 123.1171; 112.0757	Isolation
Dehydropipernonaline (12)	C ₂₁ H ₂₅ NO ₃	7.46	340.1905	- 0.588	255.1027; 227.1076; 179.1311; 161.0602; 112.0759	MN -isolation
Pipernonaline (10)	C ₂₁ H ₂₇ NO ₃	8.21	342.2056	- 2.251	255.1390; 229.1226; 220.1698; 201.0354	Isolation
1-(Pyrrolidinyl)-11-(3',4'-methylenedioxyphenyl)-2,4,10-undecatrien-1-one (13)	C ₂₂ H ₂₇ NO ₃	9.02	354.2059	- 1.328	283.1345; 255.1391; 232.1705; 215.1073; 135.0443	MN - isolation
Piperacide (14)	C ₂₂ H ₂₉ NO ₃	9.40	356.2218	- 0.703	283.1339; 255.1389; 234.1859; 215.1074; 135.1171	Isolation
Piperchabamide D	C ₂₂ H ₃₁ O ₃ N	10.50	358.2370	- 1.955	285.1393; 235.1862; 135.1170	MN - isolation
Piperundecalidine (15)	C ₂₃ H ₂₉ NO ₃	10.85	368.2213	- 1.848	283.1332; 255.1383; 135.0441	MN - Isolation
Piperchabamide B (11)	C ₂₃ H ₃₁ NO ₃	11.74	370.2372	- 1.405	285.1492, 257.1545; 161.0601; 135.0442	Isolation
Brachyamide A	C ₂₄ H ₃₁ NO ₃	12.24	382.2372	- 1.361	313.1805; 283.1699; 135.0442	<i>t.id.</i>
Guineensine (16)	C ₂₄ H ₃₃ NO ₃	12.73	384.2525	- 2.057	311.1646; 283.1698; 161.0598; 135.0442	MN -isolation
13-(1,3-Benzodioxol-5-yl)-1-(1-piperidinyl)-2,4,12-tridecatrien-1-one	C ₂₅ H ₃₃ NO ₃	13.98	396.2526	- 1.768	311.1646; 283.1697; 135.0442	MN - <i>t.id.</i>
N-Isobutyl-2,4,10-hexadecatrienamide	C ₂₀ H ₃₅ NO	14.09	306.2786	- 1.800	233.1911; 205.1961	<i>t.id.</i>
1-(1-Piperidinyl)-2,4,10-hexadecatrien-1-one	C ₂₁ H ₃₅ ON	15.79	318.2784	- 2.203	233.1910; 112.0758	<i>t.id.</i>
Piperolein B	C ₂₁ H ₂₉ NO ₃	16.71	344.2219	- 0.466	259.2063; 231.2115; 112.0757;	<i>t.id.</i>

PART 4

N-Isobutyl-hexadeca-2,4-dienamide	C ₂₀ H ₃₇ ON	16.78	308.2943	- 1.756	265.4167; 235.2061; 182.1539	<i>t.id.</i>
N-Isobutyl-2,4,12-octadecatrienamide	C ₂₂ H ₃₉ NO	17.42	334.3100	- 1.231	261.2218; 233.2266	<i>t.id.</i>
N-Isobutyl-2,4,10,12-Octadecatetraenamide	C ₂₂ H ₃₇ NO	18.10	332.2942	- 1.810	261.2218; 233.2277	<i>t.id.</i>
1-(1-Piperidiny)-2,4,12-octadecatrien-1-one	C ₂₃ H ₃₉ NO	18.94	346.3097	- 2.083	261.2221; 233.2267; 112.0759	<i>t.id.</i>
Pipericine	C ₂₂ H ₄₁ NO	19.68	336.3256	- 1.580	263.2369; 234.0333; 157.7024	<i>t.id.</i>
1-(Piperidiny)-2,4-octadecadien-1-one	C ₂₃ H ₄₁ NO	19.85	348.3257	- 1.095	264.2330; 236.2014; 112.0758	<i>t.id.</i>
N-Isobutyl-2,4,14-eicosatrienamide	C ₂₄ H ₄₃ NO	20.09	362.3407	- 2.957	289.2535; 261.2585	<i>t.id.</i>
1-(1-Pyrrolidiny)-2,4,15-eicosatrien-1-one (22)	C ₂₄ H ₄₁ NO	20.94	360.3257	- 1.142	289.2533, 262.2183,	isolation
Pipercitine	C ₂₃ H ₄₃ NO	21.41	350.3412	- 1.403	264.2398; 212.1963	<i>t.id.</i>
1-(Piperidiny)-2,4,14-eicosatrien-1-one (21)	C ₂₅ H ₄₃ NO	21.50	374.3412	- 1.473	318.28; 289.2535; 261.2588	isolation
1-(Piperidiny)-2,4-eicosadien-1-one	C ₂₅ H ₄₅ NO	22.50	376.3568	- 1.545	291.2700; 112.0760	<i>t.id.</i>

t.id. = tentatively identified

Piperamides were the most representative class of compounds present in the extract. Two main different subclasses can be identified within this class, basing on the terminal part of the fatty acid chain, one bearing a benzodioxole moiety, as deduced from the FBMN analysis, while the other does not contain an aromatic moiety and simply ends with a methyl group. The positive ion mode used for the LC-MS² experiment was well suited for the analysis of piperamides, but the LC-UV analysis showed an interesting pattern (**Figure 1.3**). In particular, in the PDA chromatogram, some peaks were clearly missing from BPC, suggesting the

presence of a non-ionising metabolite. When LC-UV-MS² analysis was performed in negative mode, the BPC resulted flat, suggesting that the non-ionising metabolite was not detected in either positive or negative mode. Thus, in order to carry out a comprehensive and detailed analysis of the secondary metabolites present in the acetonic extract of *P. borbonense* fruits, we perform an MPLC purification guided by the LC-MSMS analysis.

Approximately 5.0 g of *P. borbonense* fruit acetone extract were purified by MPLC as described in detail in the Experimental section. As expected from the LC-MSMS analysis, piperamides were the most representative class of specialized metabolites present in the extract, but from our analysis, also terpenes and lignans constituted representative metabolites of the extract. In fact, from this first purification step 73.1 mg of 5-p-menthene-1,2-diol (**23**) (İşcan et al., 2012), 495.6 mg of piperin (**6**) (Wu et al., 2004) and 221.8 mg of sesamin (**34**) (Nan-Jun et al., 1987) were isolated in pure form, while the other fractions had to be further purified.

1.4 Isolation and characterization of secondary metabolites

Among piperamides, besides piperine, the main compound (**6**) (Wu et al., 2004), other sixteen compounds were isolated, including 4,5-dihydropiperlyne (**1**) (Li et al., 2001), dehydropiperlonguminin (**2**), piperanine (**3**) (Wu et al., 2004), piperiline (**4**) (Venkatasamy et al., 2004), piperlonguminin (**5**) (Wu et al., 2004), isochavine (**7**) (Ternes et al., 2002), piperdardine (**8**) (Chen et al., 2002), chingchengenamide A (**9**) (Li et al., 2004), pipernonaline (**10**) (Wu et al., 2004), piperchabamide B (**11**) (Morikawa et al., 2004), dehydropipernonaline (**12**) (Lee et al., 2006), (2E,4E,10E)-11-(1,3-benzodioxol-5-yl)-1-(1-pyrrolidinyl)-2,4,10-undecatrien-1-one (**13**) (Tang et al., 2011), piperacide (**14**) (Lee, 2005), piperundecalide (**15**) (Tabuneng et al., 1983) and guineesine (**16**) (Wu et al., 2004). Finally, also piperodione (**17**) (Kubo et al., 2013) was isolated. On the other hand, pellitorine (**18**) (Ley et al., 2004), neopellitorine B (**19**) (Li et al., 2004), N-isobutyl-2E,4E,14Z-eicosatrienamide (**20**) (Wu et al., 2004), 1-(piperidinyl)-2,4,14-eicosatrien-1-one (**21**) and 1-(1-pyrrolidinyl)-2,4,15-eicosatrien-1-one (**22**) (David et al., 2012) were isolated as representative of conjugates of N-linked long-chain fatty acids (**Figure 1.4**).

PART 4

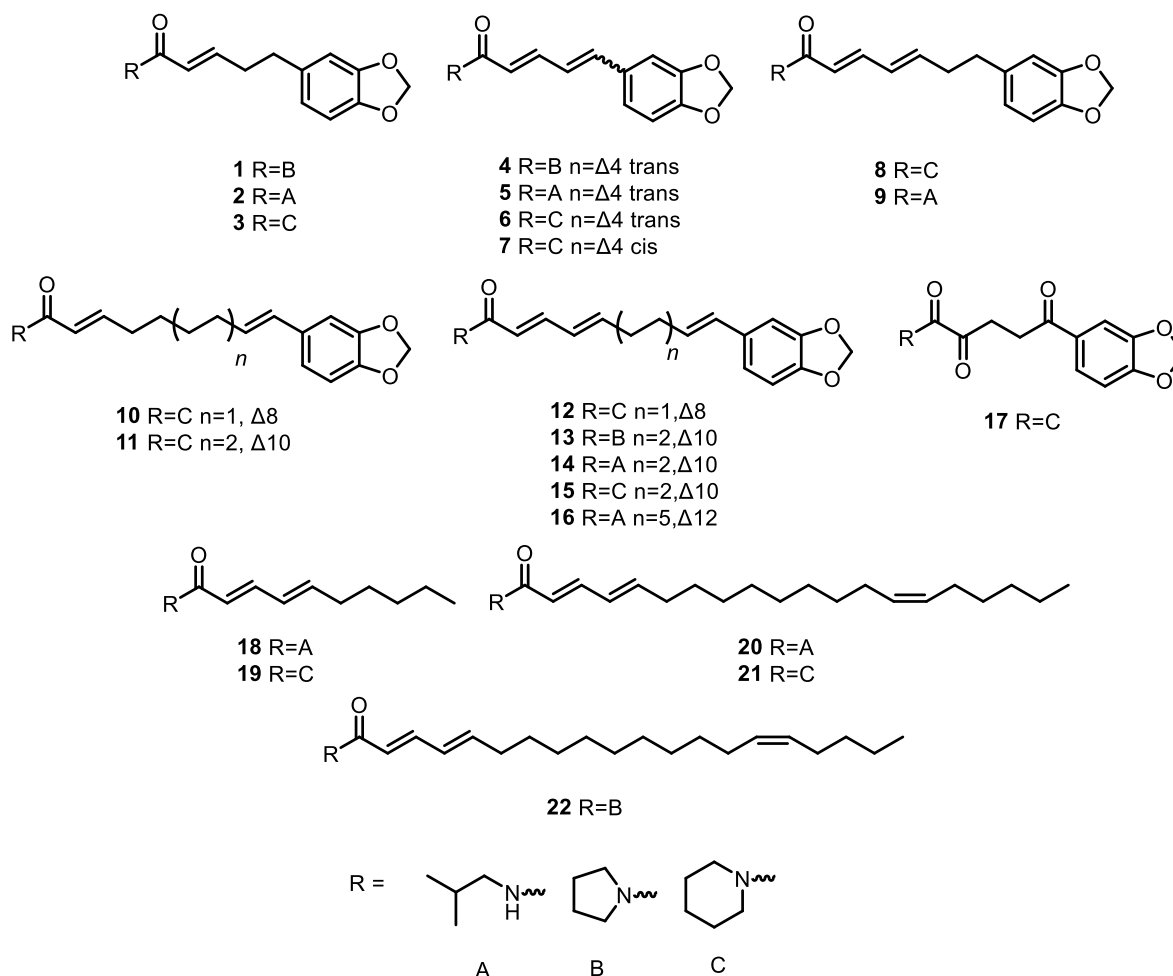


Figure 1.4 Piperamides isolated from *P. borbonense*

The first fractions eluted by MPLC showed a rich content of small terpenic metabolites, including, besides the above mentioned **23**, the monoterpene 5β-hydroxy-2-oxo-p-menth-6(1)-ene (**24**) (Liu et al., 2005), and the sesquiterpenoids multistalactone E (**25**) (Liu et al., 2013), which was never reported in genus *Piper* before, multistalactone D (**26**) (Liu et al., 2013), neolitacumone B (**27**) (Chang et al., 2002), also unreported before in genus *Piper*, and the previously undescribed 4-hydroxyisogelehomanolide (**29**), a closely related analogue of the germacrene gelehomanloide (**28**) (Zhang et al., 2015) co-occurring in the plant (**Figure 1.5**).

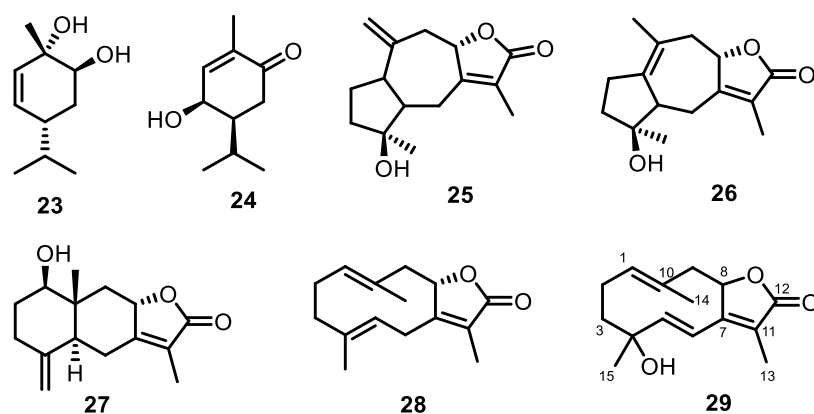


Figure 1.5 Terpenes isolated from *P. borbonense*

Besides sesamin, other lignans were in the extract, namely 5-methoxysesamin (**35**) (Tan et al., 1998), (+)-acuminatolide (**30**) (Saladino et al., 2007), which was never reported in genus *Piper* before, 4-[(1S,3aR,4S,6aR)-4-(1,3-benzodioxol-5-yl)-tetrahydro-1H,3H-furo[3,4-c]-furan-1-yl]-2,6-dimethoxyphenol (**31**) (Li et al., 2007), piperitol (**32**), pluviatiol (**33**) and verrucosin (**36**) (Hattori et al., 1987), also unreported before in the genus *Piper* before. Furthermore, two monoterpenyl hydroxycinnamic esters derivatives were isolated, including the known pressafonin A (**38**) (Cheng et al., 2010) and the unprecedented **37** (**Figure 1.6**).

CHART A

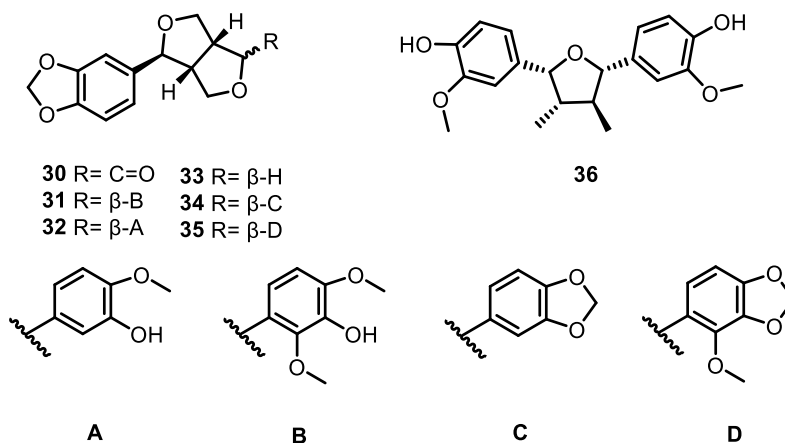


CHART B

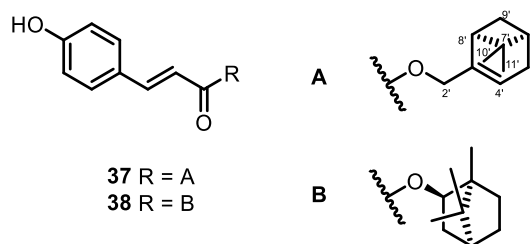


Figure 1.6 Phenols isolated from *P. borbonense*. Chart A: lignans; Chart B: hydroxycinnamic esters.

4-Hydroxyisogelehomanolide (**29**) was isolated as colorless amorphous solid with HR-ESIMS m/z 249.1485, in agreement with the molecular formula of $C_{15}H_{20}O_3$, implying six degrees of unsaturation. The ^1H NMR spectrum of **29** compound was very similar to that of **28**, although the former compound includes one additional oxygen atom compared to the second. A close comparison of the ^1H NMR spectra revealed that the most remarkable difference was a pair of mutually coupled olefinic protons resonating at δ_H 6.10 (1H, d, J 17.0 Hz) and 5.90 (1H, d, J 17.0 Hz) in place of the doublet at δ_H 4.85 (1H, d, J 12.0 Hz, H-5). This assignment was further supported by the inspection of ^{13}C and HSQC 2D NMR spectra allowing the identification of 15 carbon signals including two allylic methyls (δ_C 16.9, C-14 and 23.9, C-2), one aliphatic methyl (8.8, C-13), three methylenes (δ_C 23.9, C-2; 46.3, C-9 and 49.3, C-3), three olefinic methines (δ_C 116.3, C-5; 139.6, C-1; 152.9, C-6), three olefinic quaternary carbons (δ_C 121.6, C-11; 127.1, C-10; 159.2, C-7), three oxygenated carbons, including one methine resonating at δ_C 77.8 (C-8), one quaternary sp^3 carbon (δ_C 75.5, C-4) and one quaternary sp^2 carbon (δ_C 174.9, C-12). Analysis of the 2D NMR COSY experiments led to the definition of two spin systems, spanning from C-1 to C-3 and C-5 to C-6, and the assembly, by the 2D NMR HMBC cross-peaks, of the ten-membered ring of germacrene skeleton, elucidating the structure of **29** as a 4-hydroxylated isomer of **28**. In particular, the diagnostic HMBC correlations between H_{3-15} with C-3, C-4, C-5 and of H-6 with C-4, C-7, C-8, C-11 unambiguously assigned the position of the Δ^5 double bond, while the cross peak between H-8 and C-12 confirmed the ring closure of the lactone moiety (**Figure 1.7**). While the E geometry of Δ^5 double bond could be easily deduced by the J_{H-H} coupling constant at 17 Hz, the determination of the absolute configuration at the two stereogenic centers C-4 and C-8. Given the biogenetic relationship with the known **28**, we postulated that C-8 share the same absolute configuration. But the configuration of C-4 still remained to be determined. We faced this problem by conducting ECD analysis.

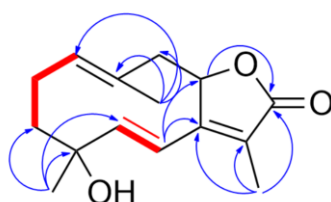


Figure 1.7 COSY (red bold) and HMBC (blue arrows) of **29**.

A configurational analysis was conducted on compounds **29**, combining electronic circular dichroism (ECD) measurements with quantum chemical ECD *ab initio* calculations. An initial conformational analysis of the R stereoisomers of **29-a** was performed using a simulated annealing procedure in the INSIGHT II software package. The resulting conformers were ranked by conformational energy and grouped into families, revealing two energy minima for **29-a1** and **29-a2** (**Figure 1.9**). Conformers were then optimized with Gaussian 16W software using DFT at the M06-2X functional and the 6-311G(2d, 2p) basis set. The ECD spectrum was simulated for all conformers of **29-a**. Excitation energies, as well as the oscillator and rotational strengths of electronic transitions, were calculated using TDDFT at the M06-2X/6-311G(2d, 2p) level. The final ECD spectrum for **29-a** was obtained by averaging the simulated spectra of the two lowest-energy conformers according to the Boltzmann distribution. The theoretical ECD curve obtained for **29-a** was then compared to the experimental spectrum of compound **29**, recorded in MeOH. The theoretical ECD curve of the **29-a** model was found to be fully superimposable with the experimental spectrum, thus confirming the absolute configuration as S at carbon 4 in compound **29** (**Figure 1.10**).

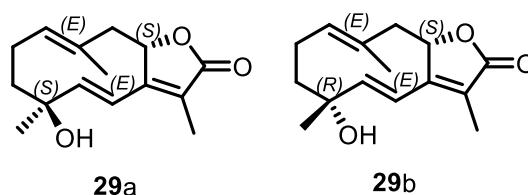
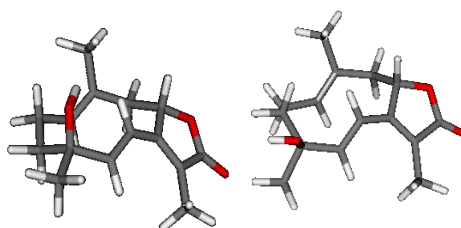


Figure 1.8 The two possible diastereoisomers of compound **29**.



	29-a1	29-a2	
Conformers	Kcal/mol	ΔG KJoule/mol	%pop.
29a-1	0.00	0.00	98.8
29a-2	2.62	10.09	1.2

Figure 1.9 The most populated conformers of **29a**.

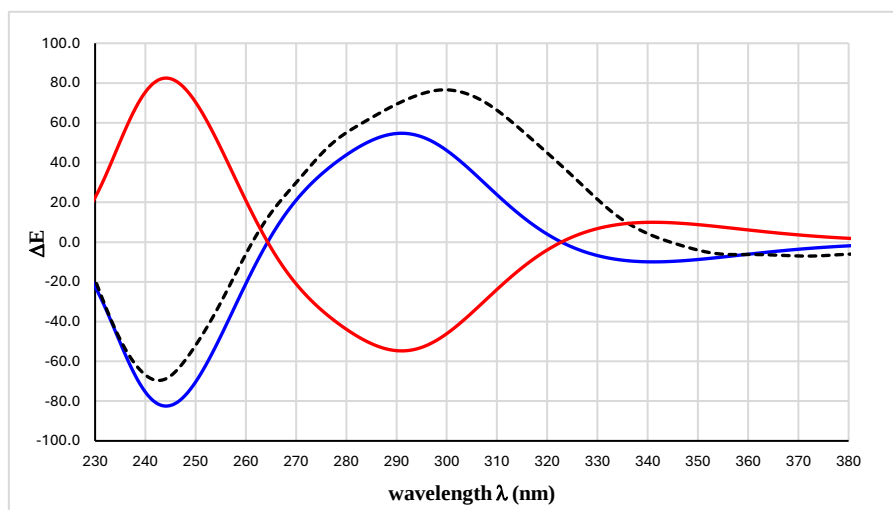


Figure 1.10 Experimental ECD spectrum of **29** (black); TDDFT-calculated curves of model **29-a** (SS) (blue) and its enantiomer (RS) (red).

Compound **37** showed an ESIMS peak at $[M-H]^-$ at m/z 297.1490, suggesting $C_{19}H_{22}O_3$ as molecular formula. The ^1H NMR spectrum of **37** showed close similarities with that of **38**, especially in the low-field region, suggesting that the two compounds share a hydroxycinnamyl moiety, as indicated by the two coupled doublets at δ_H 7.43 (2H, d, J 8.6 Hz, H-5, H-9) and 6.83 (2H, d, J 8.6 Hz, H-6, H-8), belonging to the p-disubstituted phenyl ring, and δ_H 7.64 (1H, d, J 15.9 Hz, H-2) and 6.30 (1H, d, J 15.9 Hz, H-2), belonging to the conjugated trans olefinic group in the conjugated system. In the mid-field region, a signal resonating at δ_H 5.58 and two mutually coupled doublets at δ_H 4.59 and 4.56 (both 1H, d, J 12.7, 1.3 Hz) were detected, while the high-field region appeared crowded with partially overlapping multiplets, along with two methyl singlets at δ_H 1.30 and 0.85. Once each proton has been assigned to its directly bonded carbon by HSQC, a combined analysis of the COSY and HMBC 2D NMR spectra allowed the structural elucidation of **37** (**Figure 1.8**). In particular, the COSY correlations grouped the high-field signals from H-7 to H₂-4 to the olefinic proton H-3 (1H, m, H-3'). The HMBC cross peaks and through a detailed interpretation of the HMBC correlations we constructed the bicyclic monoterpene scaffold of myrtenol. The HMBC cross-peaks H₃-10' and H₃-11' with C-5', C-7' and C-8' located the unprotonated carbon at δ_C 37.7, bearing the two methyl groups, on the four membered-ring. The HMBC correlations of the isolated methylene (δ_H 4.55, 4.59) with C-2', C-3' and C-7', placed the allylic oxymethylene at C-1', thus fully defining a myrtenol unit. Finally, the HMBC cross peaks H₂-1'/C-1 allowed us to join the hydroxycinnamic acid unit and

the myrtenol moieties, confirming that **37** is an unprecedented hydroxycinnamic acid ester derivative. Due to the positive $[\alpha]_D$ value and the absence of other chiral centres in the hydroxycinnamic acid residue, the myrtenol moiety was assigned as (+)-myrtenol.

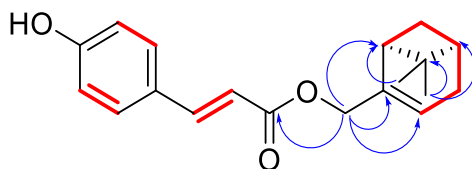


Figure 1.8 COSY (red bold) and key HMBC correlations (blue arrow) of compound **37**

1.5 Quantification of sesamin

Sesamin was the compound responsible for the absorption in the PDA chromatogram at R_t 3.79 min, giving the similar MPLC and LC-UV-MS² profiles and the UV spectrum achieved ($\lambda_{\max}$ at 236 nm and 286 nm). The unexpected relative abundance of sesamin stimulated our interest to measure its quantitative content in the extract. The quantitative profiling of sesamin and the main components, piperine isomers, in the extracts of *P. borbonense* was performed via GC-MS analysis. Sesamin and a mixture of piperin isomers were used as external standards, each at concentrations of 12.5, 25, 50, 100, and 200 ppm. The calibration curve demonstrated a high coefficient of determination ($R^2 = 0.9986$ – 0.9993), indicating the reliability of the method. Due to their structural similarity, piperine was quantified as the sum of all stereoisomers, which eluted at retention times of 24.89, 25.04, and 26.88 minutes. The results indicated comparable contents of piperine isomers (0.233 mg/g) and sesamin (0.152 mg/g) in the extract. The piperine content quantified in *P. borbonense* was consistent with previous reports (Weil et al., 2020). On the other hand, this is the first quantitative report of sesamin in *P. borbonense*, a lignan that had been detected in the common black pepper just in a few papers (Cai et al., 2011; Ramesh et al., 2015). The relatively high concentration of sesamin had been probably underestimated in earlier studies due to the widespread use of LC-ESI-MS/MS for the quantitative determination of *Piper* spp. chemical constituents. Furthermore, according to recent studies conducted on sesame oil, where this lignan is found in high concentrations, sesamin can affect aroma due to its strong antioxidant properties (Mostashari et al., 2024). In fact, lignans can influence flavour formation and improve sensory perception of sesame oil by interfering with lipid oxidation (Yin et al., 2004). It has been reported in the literature that lignan content

depends on growing conditions and location of the plant and is strongly related to seed coat colour, with black sesame seeds having the highest levels (Wei et al., 2022). Moreover, sesamin content can influence the processing steps of the biomass, but to our knowledge it was not taken into account in the previous studies concerning the processing of *P. borbonense* (Weil et al., 2017).

1.6 Analysis of the volatile compounds of *P. borbonense* fruits

The essential oil composition of *P. borbonense* has been evaluated in previous studies, highlighting its low pungency and high aroma potential (Weil et al., 2020). However, to our knowledge, headspace gas chromatography-mass spectrometry (HS-GC-MS) has never been used to analyse the aroma composition of this peculiar spice. As it requires no sample preparation and can be carried out directly on the crushed fruit of the plant, we reasoned that this technique would give us the actual profile of the volatile components of this highly prized pepper. Driven by the interest to highlight the main differences between and the common *P. nigrum*, we carried out a comparative analysis of the HS-GC-MS profile of the two species, identifying the compounds using the NIST library. As expected, the most abundant compounds were monoterpenes, 98.15% in *P. borbonense*, similar to 97.67% in *P. nigrum*. The most abundant compound was δ -3-carene, which reached 40.91% in *P. borbonense*, whereas 50.23% in *P. nigrum*. The other major peaks were assigned to D-limonene, 1.90% higher in *P. nigrum*, and α -pinene, 2.48% higher in *P. borbonense*. Furthermore, the β -pinene content was 10.48% in *P. borbonense* and 12.56% in *P. nigrum*, while β -thujene reached 7.51% in *P. borbonense* and only 0.53% in *P. nigrum*. The total sesquiterpenoid composition was higher in *P. nigrum* (2.15% compared to 1.43% of *P. borbonense*), but interestingly *P. borbonense* showed a more variable composition, as in *P. nigrum* the 1.87% was identified as caryophyllene (0.44% in *P. borbonense*). Finally, phenylpropanoids were not detected in *P. nigrum*, whereas they were found in small amounts (0.32%) in *P. borbonense*. The major differences of the two profiles are reported in **Table 1.2**, while the detailed analysis is reported in Appendix E.

Table 1.2 The main differences detected between HS-GC-MS profiles of *P. borbonense* and *P. nigrum*

Compound type	compound name	Rt (min)	<i>P. borbonense</i> Area %	<i>P. nigrum</i> Area %
Major monoterpenes	α -pinene	12.47	12.21	9.73
	sabinene	13.68	1.33	0.20
	β -myrcene	13.92	1.66	2.74
	β -pinene	14.00	10.48	12.56
	δ -3-carene	14.80	40.91	50.23
	p-cymene	15.35	5.52	1.86
	D-limonene	15.43	15.91	17.81
	β -thujene	15.58	7.51	0.53
Total for monoterpenes			98.15	97.67
Major sesquiterpenes	caryophyllene	26.22	0.44	1.87
Total for sesquiterpenes			1.43	2.15
Total for phenylpropanoids			0.32	N.d.

1.7 Conclusions

Our series of analyses on Madagascan *P. borbonense* led to the first phytochemical characterisation of this plant, highlighting some of the main differences between this rare species and the common *P. nigrum*, such as the aroma composition and sesamin content. In particular, the LC-MS/MS-based Feature-Based Molecular Networking analysis led to the phytochemical fingerprint of the piperamides present in the extract, which guided the subsequent purification steps, while the HS-GC-MS analysis revealed a rich and diverse terpenic composition of *P. borbonense*, as well as the presence of phenylpropanoids absent in *P. nigrum*. Indeed, the mono- and sesquiterpenes were a relevant class of compounds in the extract, as also revealed by the MPLC purification, which led to the isolation of several terpenes previously unreported in the *Piper* genus. In addition, a rich variety of lignans was identified, including the very abundant sesamin, which was quantified. Finally, two previously undescribed metabolites were characterised in detail. In conclusion, the typical aroma of this spice can be related to the variety of monoterpenes and sesquiterpenes detected, while its peculiar taste can be attributed to the rich abundance of fatty acid amides bearing a Z-olefinic group and a surprisingly high content of sesamin. Thus, *P. borbonense* showed a peculiar profile that can justify its high price on the market when compared to *P. nigrum*, and it would be interesting to evaluate in more detail how the lignan content varies between the species and how this could affect the organoleptic properties of these spices. Given the impressive number of isolated metabolites belonging to different structural families, we are currently conducting

a detailed study of the data in the literature to evaluate the most interesting MetS-related targets to test our compound library.

1.8 Experimental section

1.8.1 Plant Material and Extraction

Piper borbonense fruits and its acetonic extract were provided from Professor Giovanni Appendino of the University of Piemonte Orientale, Novara, Italy.

1.8.2 LC-MS/MS Analysis and Molecular Networking

The acetonic extract of *P. borbonense* was subjected to a LC-MS/MS analysis on an LTQ Orbitrap XL FT-MS mass spectrometer (Thermo Fisher Scientific Spa) equipped with an Ultimate 3000 HPLC (Agilent Technology). Chromatographic separation was achieved on a Kinetex 2.6 μm polar C18 100 Å (100 $\times$ 3 mm) column (Phenomenex). The LC-MS/MS experiment was carried out using a combination of A (0.1% formic acid in H₂O, v/v) and B (MeCN) as mobile phase; the gradient elution was optimized as follows: 50% B for 3 min, 50% to 95% B over 20 min, held 10 min before returning to the initial conditions. The total run time, including the column wash and the equilibration, was 30 min, flow rate 0.5 mL/min, injection volume 5 μL . The MS and MSⁿ spectra, in positive mode, were recorded in data-dependent acquisition mode, inducing fragmentation of the five most intense peaks for each scan. Source conditions were as follows: spray voltage, 3.5 kV (positive mode); capillary voltage, 25 V; source temperature, 320 °C; normalized collision energy, 25. The acquisition range was *m/z* 150-1500. The LC-MSMS data were converted from .raw to .mzML using MSConvert part of the ProteoWizard 3.0 package before being processed on Mzmine 2.53 software using the following parameters: mass detection noise level 1000 for mass level 1; mass detection noise level was set as 50 for nose level 2; the ADAP chromatogram builder was used with a minimum of 5 consecutive scans with a minimum intensity of 1000 and a minimum absolute height of 1000, the *m/z* tolerance was set at 0.005 or 15 ppm; deconvolution was performed with the ADAP resolver with a signal to noise ratio of 30, a minimum feature height of 1000, a coefficient/area threshold set at 10, a peak duration range of 0.0 to 1.0 min and the RT wavelet range set between 0.01 and 0.10; the isotope peak grouper was used with the *m/z* tolerance set at 0.005 or 15 ppm, a RT tolerance of 0.1 min and the maximum number of charges set at 2. The molecular

network was generated on MetGem, setting m/z tolerance 0.04, minimum matched peaks 3, keeping peak outside ± 17 Th, keeping peaks above 1% and in top 6 in the 50 Th window. Then, the maximum neighbour number was set at 10, the minimal score value at 0.50, and max connected component size at 1000. The molecular networks were visualized using Cytoscape_v3.7.2 software.³

1.8.3 GC-MS Analysis and quantitative analysis

GC-MS analysis was performed on an Agilent 6850 Ser. II apparatus, fitted with a fused silica HP-5 capillary column (30 m x 0.25 mm), 0.33 mm film thickness, coupled to an Agilent Mass Selective Detector MSD 5973; ionization voltage 70 eV; electron multiplier energy 2000 V. Helium was used as carrier gas (1 mL/min). Gas chromatographic conditions were column temperature: 40°C, with 5 min initial hold, and then to 260°C at 2°C/min, 260°C (20 min). Interface temperature 295°C; mass range was m/z 29–600, ionization energy 70 eV, multiplier energy 2000 V, scan time 1 s. Peak identification was accomplished by comparison of their mass spectra with those stored on the GC-MS data bases (NIST 02 and Wiley 275) and reported in literature.

1.8.4 Isolation of pure compounds

The acetonic extract (4.7 g) was purified on MPLC-DAD on a Silica Cartridge 60Å 50µ - size 120 (180 g) – column volume (CV) 153 mL. The mobile phase was a mixture of A: 0.1% formic acid in H₂O, v/v, B: MeCN and C: methanol with the following gradient method: starting conditions: 80%A – 20%B for CV 0-2; 70%A – 30%B for CV 4-6; 60%A – 40%B for CV 8-10; 50%A – 50%B for CV 13-16; 40%A – 60%B for CV 18-20; 10%A – 90%B for CV 30-32; 5%A – 95%C for CV 33-35; 100%C for CV 37-42; flow rate was 40.0 mL/min. The UV detection wavelength was set at 270 nm. This separation afforded 26 fractions (A-Z). Fraction B (73.1 mg) was identified as 5-p-menthene-1,2-diol (**23**). Fraction D (18.7 mg) was purified on RI-HPLC using a Luna 5 µm C18 100 Å 250 × 4.6 mm, flow rate 1.0 mL/min, using the isocratic elution 50%A-50%C, affording compound **24** (0.8 mg). Fraction E (26.9 mg) was purified on RP-HPLC using a mixture of 55%A-45%C as mobile phase, on a Luna 5 µm C18 100 Å 250 × 4.6 mm, flow rate 1.0 mL/min, yielding to compound **30** (0.8 mg, Rt 12.4 min), **25** (0.6 mg, Rt 0.6 min), **26** (0.5 mg, Rt 33.4 min), the unprecedented **29** (1.4 mg, Rt 40.0 min) and **27** (1.0 mg, Rt 43.1 min). Fraction G (43.4 mg) was purified on a HPLC-UV, setting the UV detection wavelength at 270 nm, using a Luna 10 µm C18

100 Å 250 × 10 mm, flow rate 3.0 mL/min, to give **31** (1.5 mg, Rt 17.6 min), **32** (1.2 mg, Rt 19.2 min), **1** (2.3 mg, Rt 24.4 min) and **4** (10.3 mg, Rt 27.7 min). Fraction H (16.5 mg) was purified using an HPLC-UV, setting the UV detection wavelength at 230 nm, using Luna 5 µm C18 100 Å 250 × 4.6 mm, flowrate 1.0 mL/min, using a mixture of 50%A-50%C as mobile phase, affording **17** (0.4 mg, Rt 16.7 min), **33** (0.5 mg, Rt 16.7 min), **36** (0.3 mg, Rt 21.7 min), **2** (2.3 mg, Rt 23.4 min), **5** (0.6 mg, Rt 26.3 min) and **3** (1.0 mg, Rt 28.4 min). Piperine (**6**) was obtained from fraction I (495.6 mg) in pure form. Fraction J (29.4 mg) was purified on HPLC-UV, UV detection wavelength set at 270 nm, equipped with a Luna 5 µm C18 100 Å 250 × 4.6 mm, using a flowrate of 1.0 mL/min, using an isocratic elution of 35%A-65%C, yielding to **7** (0.5 mg, Rt 15.0 min), **35** (3.2 mg, Rt 17.1 min), **9** (0.8 mg, 21.1 min), and **8** (0.7 mg, 27.1 min). Sesamine was obtained in pure from fraction K (221.8 mg). Fraction L (92.7 mg) was purified on an MPLC using a Silica Cartridge 100Å 25µ - size 4 (5.4 g) – column volume (CV) 5 mL, setting UV wavelength at 270 nm and at 5 mL/min flowrate, using the following gradient method: 30%A-70%C for CV 0-5, 25%A-75%C for CV 15-25; 20%A-80%C for CV 35-45; 15%A-85%C for CV 55-60; 5%A-95%C for CV 65-70, affording compound **28** (14.0 mg, CV 9-11) and compound **18** (13.6 mg, CV 13-15). Fraction M (27.3 mg) was purified on an HPLC-UV, setting the UV wavelength at 270 nm, using a Luna 10 µm C18 100 Å 250 × 10 mm, flowrate 3.0 mL/min, using as mobile phase 25%A-75%C, obtaining compound **12** (8.8 mg, Rt 28.9 min). Fraction N (38.0 mg) was purified on HPLC-UV using a Luna 5 µm C18 100 Å 250 × 4.6 mm, flowrate 1.0 mL/min, UV wavelength set at 270nm, using 25%A-75%C, affording **19** (3.6 mg, 16.3 min), **10** (0.5 mg, Rt 20.3 min), **13** (1.9 mg, Rt 25.8 min). Fraction O (57.8 mg) was purified on a HPLC-UV, setting UV wavelength at 300nm, using a Luna 5 µm C18 100 Å 250 × 4.6 mm, flowrate 1.0 mL/min, with the isocratic method 20%A-80%C, obtaining **14** (1.8 mg, Rt 17.9 min) and the previously undescribed **37** (1.0 mg, Rt 18.8 min). Fraction P (80.6 mg) was purified on a MPLC using a Silica Cartridge 100Å 25µ - size 4 (5.4 g) – column volume (CV) 5 mL, setting UV wavelength at 270 nm, using the following gradient method: 50%A-50%C for CV 0-5, 20%A-80%C for CV 55-60; 5%A-95%C for CV 65-67, yielding to **38** (8.9 mg, CV 16-19) and **15** (6.4 mg, CV 20-23) at 5 mL/min flowrate. Fraction Q (29.7 mg) was purified on a HPLC-UV equipped with a Luna 10 µm C18 100 Å 250 × 10 mm, flowrate 3.0 mL/min, with the isocratic method 15%A-85%C, setting the UV wavelength at 270 nm, yielding to **11** (5.8 mg, 20.6 min) and **16** (10.5 mg, Rt 24.8 min). Fraction X (77.3 mg) was purified on MPLC-DAD using a Silica Cartridge 60Å 50µ - size 4 (4.0 g) – column volume (CV) 5 mL and the following a A= Hexane, and C= Ethyl Acetate a mobile phase with the following gradient method:

90%A-10%C for CV 0-5; 80%A-20%C for CV 10-15, 70%A-30%C for CV 20-25, setting UV wavelength at 220 and 270 nm and 10 mL/min as flowrate, yielding to compound **20** (10.5 mg, CV 3-4), **22** (3.8 mg, 5 CV). Fraction Y (98.9 mg) was purified on MPLC a Silica Cartridge 60Å 50µ - size 12 (12.0 g) – column volume (CV) 25 mL and using A= Hexane, and C= Ethyl Acetate as mobile phase with the following gradient method: 100%A for CV 0-5; 90%A-10%C for CV 10-15 and 80%A-20%C for CV 20-25, at 15 mL/min as flowrate, yielding to the unprecedented **21** (20.6 mg, CV 17-19).

1.8.5 Spectral data

Compound **29**: yellowish amorphous solid, $[\alpha]_D^{25} +99.3$ (c 0.12, MeOH); HRESIMS $[M+H]^+$ m/z found 249.1489 ($C_{15}H_{21}O_3$ requires 249.1485). ^1H NMR ($CDCl_3$, 600 MHz) 6.11 (1H, d, J 17.0 Hz, H-6), 5.89 (1H, d, J 17.0 Hz H-5), 5.11 (1H, brd, J 11.4 Hz, H-8), 4.94 (1H, brd, J 10.6 Hz, H-1), 2.90 (1H, brd, J 11.9 Hz, H-10a), 2.63 (1H, m, H-2a), 2.03 (2H, overlapped, H-2b, H-3a), 2.02 (1H, overlapped, H-10b), 1.88 (3H, s, H-13), 1.81 (3H, s, H-14), 1.77 (1H, overlapped, H-3b), 1.48 (3H, s, H-15); ^{13}C NMR ($CDCl_3$, 150 MHz) 174.9 (C-12), 159.2 (C-7), 152.9 (C-5), 139.6 (C-1), 127.1 (C-10), 121.6 (C-11), 116.3 (C-6), 77.8 (C-8), 75.5 (C-4), 49.3 (C-3), 46.1 (C-9), 27.5 (C-15), 23.9 (C-2), 16.9 (C-14), 8.8 (C-13).

Compound **37**: white amorphous solid, $[\alpha]_D^{25} +6.8$ (c 0.12, MeOH); ESIMS $[M-H]^-$ m/z found 297.1490 ($C_{19}H_{21}O_3$ requires 297.1485). ^1H NMR ($CDCl_3$, 600 MHz) 7.64 (1H, d, J 15.9 Hz, H-3), 7.43 (2H, d, J 8.6 Hz, H-5, H-9), 6.83 (2H, d, J 8.6 Hz, H-6, H-8), 6.30 (1H, d, J 15.9 Hz, H-2), 5.58 (1H, m, H-4'), 4.59 (1H, d, J 12.7, 1.3 Hz H-2'a), 4.55 (1H, dd, J 12.7, 1.3 Hz, H-2'b), 2.41 (1H, m, H-7'a), 2.33 (1H, m, H-5'a), 2.26 (1H, m, H-5'b), 2.17 (1H, dt, J 5.5, 1.2 Hz, H-8'), 2.11 (1H, m, H-6'), 1.30 (3H, s, H-10'), 1.21 (1H, overlapped, H-7'b), 0.85 (3H, s, H-11'); ^{13}C NMR ($CDCl_3$, 150 MHz) 167.3 (C-1), 157.4 (C-7), 144.3 (C-3), 143.1 (C-3'), 130.1 (C-5, C-9), 127.4 (C-4), 121.1 (C-4'), 115.9 (C-6, C-8), 115.8 (C-2), 66.9 (C-2'), 43.6 (C-8'), 40.6 (C-6'), 38.8 (C-7'), 31.2 (C-5'), 27.1 (C-9'), 26.0 (C-10'), 21.1 (C-11').

1.8.6 Computational Details

A preliminary conformational search for the **29** model was conducted using simulated annealing in the INSIGHT II package. The solution phase in MeOH was simulated by setting the corresponding dielectric constant. Conformational energy minimization was achieved through

steepest descent, followed by the quasi-Newton-Raphson method (VA09A). Restrained simulations were run for 500 ps using the CVFF force field, as implemented in Discover software (Accelrys, San Diego, CA). The simulation began at 1000 K, and the temperature was gradually decreased to 300 K. A final energy minimization step was performed to refine the obtained structures, using successive applications of steepest descent and quasi-Newton-Raphson (VA09A) algorithms. Both dynamic and mechanical calculations were conducted with flat-well distance restraints set to 1 (kcal/mol)/Å². A total of one hundred structures were generated. All optimizations were performed with the Gaussian 16W software package, using the DFT functional M06-2X and the basis set 6-311G(2d,2p). The rotatory strength values for electronic transitions from the ground state to singly excited states for the two conformers of E09-a were calculated using TDDFT at the M06-2X/6-311G(2d,2p) level with Gaussian 16W, including at least 30 excited states in each case and using the IEF-PCM model for MeOH. The resulting ECD spectra were then compared to the experimental spectra using SpecDis 1.71 software.

1.8.7 Head-space gas-chromatography

Head-space gas-chromatography analysis was carried out on a Perkin-Elmer Clarus 600 gas chromatograph equipped with a Perkin-Elmer Clarus 600 mass spectrometer and a Perkin-Elmer TurboMatrix 40 headspace sampler. 120.0 mg of the sample (*P. borbonense* fruits) were crushed and insert in a sealed vial, which was heated in the sampler at 80°C for 10 minutes before the injection. Helium was used as gas carrier and 1 µL was injected to the transfer line at 150°C and 30 psi. The VOCs were carried by He gas through the transfer line maintained at 180°C for analysis in the GCMS system. The gas chromatograph was equipped with a DB-5 column (60 m x 0.25 mm x 0.25 mm inner diameter). The oven temperature was increased from 50 to 240°C at 6°C/min, to 310°C at 10°C/min, holding for 10 minutes, resulting in 46 minutes of measurement time. The total column pressure was kept constant at 26 psi in splitless mode. The mass spectra were obtained operating in electronic ionization mode (EI) at 70 eV and the ion source temperature 200 °C, with a scan mass range of *m/z* 40-500 and sampling rate of 0.2 seconds.

References

- Bezerra, D.P., Pessoa, C., Moraes, M.O.D., Silveira, E.R., Lima, M.A.S., Martins Elmiro, F.J. and Costa-Lotufo, L.V., **2005**. Antiproliferative effects of two amides, piperine and piplartine, from Piper species. *Zeitschrift für Naturforschung C*, 60(7-8), pp.539-543.
- Cai, C.C., Yan, M.C., Xie, H. and Pan, S.L., **2011**. Simultaneous determination of ten active components in 12 Chinese Piper species by HPLC. *The American Journal of Chinese Medicine*, 39(05), pp.1043-1060.
- Chang, F.R., Hsieh, T.J., Huang, T.L., Chen, C.Y., Kuo, R.Y., Chang, Y.C., Chiu, H.F. and Wu, Y.C., **2002**. Cytotoxic constituents of the stem bark of neolitsea a cuminatissima. *Journal of Natural Products*, 65(3), pp.255-258.
- Chen, J.J., Huang, Y.C., Chen, Y.C., Huang, Y.T., Wang, S.W., Peng, C.Y., Teng, C.M. and Chen, I.S., **2002**. Cytotoxic amides from Piper sintonense. *Planta medica*, 68(11), pp.980-985.
- Cheng, M.J., Lo, W.L., Huang, J.C., Yeh, Y.T., Hong, Z.L., Lu, Y.C., Chang, M.S. and Chen, C.Y., **2010**. Isolation of a new monoterpenic ester from the leaves of Michelia compressa (Maxim.) Sargent var. formosana Kanehira (Magnoliaceae). *Natural Product Research*, 24(7), pp.682-686.
- Dawid, C., Henze, A., Frank, O., Glabasnia, A., Rupp, M., Büning, K., Orlikowski, D., Bader, M. and Hofmann, T., **2012**. Structural and sensory characterization of key pungent and tingling compounds from black pepper (Piper nigrum L.). *Journal of agricultural and food chemistry*, 60(11), pp.2884-2895.
- Fortin, H., Vigor, C., Lohézic-Le Dévéhat, F., Robin, V., Le Bossé, B., Boustie, J. and Amoros, M., **2002**. In vitro antiviral activity of thirty-six plants from La Réunion Island. *Fitoterapia*, 73(4), pp.346-350.
- Geddo, F., Scandiffio, R., Antoniotti, S., Cottone, E., Querio, G., Maffei, M.E., Bovolín, P. and Gallo, M.P., **2019**. PipeNig®-FL, a fluid extract of black pepper (Piper nigrum L.) with a high standardized content of Trans- β -Caryophyllene, reduces lipid accumulation in 3T3-L1 preadipocytes and improves glucose uptake in C2C12 myotubes. *Nutrients*, 11(11), p.2788.
- Graidist, P., Martla, M. and Sukpondma, Y., **2015**. Cytotoxic activity of Piper cubeba extract in breast cancer cell lines. *Nutrients*, 7(4), pp.2707-2718.

- Hattori, M., Hada, S., Kawata, Y., Tezuka, Y., Kikuchi, T. and Namba, T., **1987**. New 2, 5-bis-aryl-3, 4-dimethyltetrahydrofuran lignans from the aril of *Myristica fragrans*. *Chemical and pharmaceutical bulletin*, 35(8), pp.3315-3322.
- İşcan, G., Kırimer, N., Demirci, F., Demirci, B., Noma, Y. and Başer, K.H.C., **2012**. Biotransformation of (-)-(R)- α -phellandrene: antimicrobial activity of its major metabolite. *Chemistry & Biodiversity*, 9(8), pp.1525-1532.
- Kikuzaki, H., Kawabata, M., Ishida, E., Akazawa, Y., Takei, Y. and Nakatani, N., **1993**. LC-MS analysis and structural determination of new amides from Javanese long pepper (*Piper retrofractum*). *Bioscience, biotechnology, and biochemistry*, 57(8), pp.1329-1333.
- Kubo, M., Ishii, R., Ishino, Y., Harada, K., Matsui, N., Akagi, M., Kato, E., Hosoda, S. and Fukuyama, Y., **2013**. Evaluation of constituents of *Piper retrofractum* fruits on neurotrophic activity. *Journal of natural products*, 76(4), pp.769-773.
- Kumar, S.R., Mohd Ramli, E.S., Abdul Nasir, N.A., Mohd Ismail, N. and Mohd Fahami, N.A., **2021**. Methanolic extract of *Piper sarmentosum* attenuates obesity and hyperlipidemia in fructose-induced metabolic syndrome rats. *Molecules*, 26(13), p.3985.
- Lee, H.S., **2005**. Pesticidal constituents derived from Piperaceae fruits. *Applied Biological Chemistry*, 48(2), pp.65-74.
- Lee, S.W., Rho, M.C., Park, H.R., Choi, J.H., Kang, J.Y., Lee, J.W., Kim, K., Lee, H.S. and Kim, Y.K., **2006**. Inhibition of diacylglycerol acyltransferase by alkamides isolated from the fruits of *Piper longum* and *Piper nigrum*. *Journal of agricultural and food chemistry*, 54(26), pp.9759-9763.
- Ley, J.P., Hilmer, J.M., Weber, B., Krammer, G., Gatfield, I.L. and Bertram, H.J., **2004**. Stereoselective enzymatic synthesis of cis-pellitorine, a taste active alkamide naturally occurring in Tarragon. *European journal of organic chemistry*, 2004(24), pp.5135-5140.
- Li, Y., Zhang, Y., Huang, Z., Cao, X. and Gao, K., **2004**. Stereoselective synthesis of naturally occurring unsaturated amide alkaloids by a modified Ramberg-Bäcklund reaction. *Canadian journal of chemistry*, 82 (5), pp.622-630.
- Li, N., Wu, J.L., Hasegawa, T., Sakai, J.I., Bai, L.M., Wang, L.Y., Kakuta, S., Furuya, Y., Ogura, H., Kataoka, T. and Tomida, A., **2007**. Bioactive lignans from *Peperomia duclouxii*. *Journal of natural products*, 70(4), pp.544-548.
- Lister, I.N.E., Ginting, C.N., Girsang, E., Nataya, E.D., Azizah, A.M. and Widowati, W., **2020**. Hepatoprotective properties of red betel (*Piper crocatum* Ruiz and Pav) leaves extract

- towards H₂O₂-induced HepG2 cells via anti-inflammatory, antinecrotic, antioxidant potency. *Saudi Pharmaceutical Journal*, 28(10), pp.1182-1189.
- Lieder, B., Zaunschirm, M., Holik, A.K., Ley, J.P., Hans, J., Krammer, G.E. and Somoza, V., **2017**. The alkamide trans-pellitorine targets PPAR γ via TRPV1 and TRPA1 to reduce lipid accumulation in developing 3T3-L1 adipocytes. *Frontiers in pharmacology*, 8, p.316.
- Liu, X., Wu, Q.X. and Shi, Y.P., **2005**. Terpenoids from the flower of *Cacalia tangutica*. *Journal of the Chinese Chemical Society*, 52(2), pp.369-374.
- Lu, J.Y., Huang, K.C., Chang, L.C., Huang, Y.S., Chi, Y.C., Su, T.C., Chen, C.L. and Yang, W.S., **2008**. Adiponectin: a biomarker of obesity-induced insulin resistance in adipose tissue and beyond. *Journal of biomedical science*, 15, pp.565-576.
- Luca, S.V., Minceva, M., Gertsch, J. and Skalicka-Woźniak, K., **2021**. LC-HRMS/MS-based phytochemical profiling of Piper spices: Global association of piperamides with endocannabinoid system modulation. *Food research international*, 141, p.110123.
- Mgbeahuruike, E.E., Yrjönen, T., Vuorela, H. and Holm, Y., **2017**. Bioactive compounds from medicinal plants: Focus on Piper species. *South African Journal of Botany*, 112, pp.54-69.
- Mgbeahuruike, E.E., Fyhrquist, P., Vuorela, H., Julkunen-Tiitto, R. and Holm, Y., **2018**. Alkaloid-rich crude extracts, fractions and piperamide alkaloids of *Piper guineense* possess promising antibacterial effects. *Antibiotics*, 7(4), p.98.
- Morikawa, T., Matsuda, H., Yamaguchi, I., Pongpiriyadacha, Y. and Yoshikawa, M., **2004**. New amides and gastroprotective constituents from the fruit of *Piper chaba*. *Planta medica*, 70(02), pp.152-159.
- Mostashari, P. and Mousavi Khaneghah, A., **2024**. Sesame Seeds: A Nutrient-Rich Superfood. *Foods*, 13(8), p.1153.
- Nan-Jun, S., Ching-Jer, C. and Cassady, J.M., **1987**. A cytotoxic tetralone derivative from *Pararistolochia flos-avis*. *Phytochemistry*, 26(11), pp.3051-3053.
- Nothias, L.F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., Protsyuk, I., Ernst, M., Tsugawa, H., Fleischauer, M. and Aicheler, F., **2020**. Feature-based molecular networking in the GNPS analysis environment. *Nature methods*, 17(9), pp.905-908.
- Okumura, Y., Narukawa, M., Iwasaki, Y., Ishikawa, A., Matsuda, H., Yoshikawa, M. and Watanabe, T., **2010**. Activation of TRPV1 and TRPA1 by black pepper components. *Bioscience, biotechnology, and biochemistry*, 74(5), pp.1068-1072.

- Olivon, F., Elie, N., Grelier, G., Roussi, F., Litaudon, M. and Touboul, D., **2018**. MetGem software for the generation of molecular networks based on the t-SNE algorithm. *Analytical chemistry*, 90(23), pp.13900-13908.
- Palchetti, E., Biricolti, S., Gori, M., Nodari, G.R., Gandolfi, N. and Papini, A., **2018**. Two new Malagasy species of genus *Piper* L.(Piperaceae), *Piper malgassicum* and *Piper tsarasotrae*, and their phylogenetic position. *Turkish Journal of Botany*, 42(5), pp.610-622.
- Pandey, S. and Ovadge, P.U., University of Windsor, **2016**. *Long Pepper Extract an Effective Anticancer Treatment*. U.S. Patent Application 14/942,438.
- Ramesh, B., Sarma, V.U.M., Kumar, K., Babu, K.S. and Devi, P.S., **2015**. Simultaneous determination of six marker compounds in *Piper nigrum* L. and species comparison study using High-Performance Thin-Layer Chromatography–Mass Spectrometry. *Journal of Planar Chromatography*, 28(4), pp.280-286.
- Razafimandimby, H., Bénard, A.G., Andrianoelisoa, H., Tsy, J.L.P., Touati, G., Levesque, A., Weil, M., Randrianaivo, R., Ramamonjisoa, L., Queste, J. and Aubert, S., **2017**. Tsiperifery, the wild pepper from Madagascar, emerging on the international spice market whose exploitation is unchecked: Current knowledge and future prospects. *Fruits*, 72(6), pp.331-340.
- Saladino, R., Fiani, C., Crestini, C., Argyropoulos, D.S., Marini, S. and Coletta, M., **2007**. An efficient and stereoselective dearylation of asarinin and sesamin tetrahydrofurofuran lignans to acuminatolide by methyltrioxorhenium/H₂O₂ and UHP systems. *Journal of natural products*, 70(1), pp.39-42.
- Salehi, B., Zakaria, Z.A., Gyawali, R., Ibrahim, S.A., Rajkovic, J., Shinwari, Z.K., Khan, T., Sharifi-Rad, J., Ozleyen, A., Turkdonmez, E. and Valussi, M., **2019**. *Piper* species: A comprehensive review on their phytochemistry, biological activities and applications. *Molecules*, 24(7), p.1364.
- Schmid, R., Heuckeroth, S., Korf, A., Smirnov, A., Myers, O., Dyrland, T.S., Bushuiev, R., Murray, K.J., Hoffmann, N., Lu, M. and Sarvepalli, A., **2023**. Integrative analysis of multimodal mass spectrometry data in MZmine 3. *Nature biotechnology*, 41(4), pp.447-449.
- Seo, J., Lee, U., Seo, S., Wibowo, A.E., Pongtuluran, O.B., Lee, K., Han, S.B. and Cho, S., **2022**. Anti-inflammatory and antioxidant activities of methanol extract of *Piper betle*

- Linn.(Piper betle L.) leaves and stems by inhibiting NF- κ B/MAPK/Nrf2 signaling pathways in RAW 264.7 macrophages. *Biomedicine & Pharmacotherapy*, 155, p.113734.
- Soidrou, S.H., Farah, A., Satrani, B., Ghanmi, M., Jennan, S., Hassane, S.O.S., Lachkar, M., El Abed, S., Ibnsouda Koraichi, S. and Bousta, D., **2013**. Fungicidal activity of four essential oils from Piper capense, Piper borbonense and Vetiveria zizanoides growing in Comoros against fungi decay wood. *Journal of essential oil research*, 25(3), pp.216-223.
- Soidrou, S.H., Bousta, D., Lachkar, M., Hassane, S.O., Youbi-Hamsas, A.E., Mansouri, L.E., Benjilali, J., El-Hajaji, H. and Farah, A., **2014**. Immunomodulatory activity of phenolic fraction from Piper borbonense and Cassytha filiformis growing in Comoros Islands. In *Chemistry: The Key to our Sustainable Future* (pp. 105-112). Springer Netherlands.
- Tan, R.X., Tang, H.Q., Hu, J. and Shuai, B., **1998**. Lignans and sesquiterpene lactones from Artemisia sieversiana and Inula racemosa. *Phytochemistry*, 49(1), pp.157-161.
- Tang, G.H., Chen, D.M., Qiu, B.Y., Sheng, L., Wang, Y.H., Hu, G.W., Zhao, F.W., Ma, L.J., Wang, H., Huang, Q.Q. and Xu, J.J., **2011**. Cytotoxic amide alkaloids from Piper boehmeriaefolium. *Journal of Natural Products*, 74(1), pp.45-49.
- Ternes, W. and Krause, E.L., **2002**. Characterization and determination of piperine and piperine isomers in eggs. *Analytical and bioanalytical chemistry*, 374, pp.155-160.
- Tabuneng, W., Bando, H. and Amiya, T., **1983**. Studies on the constituents of the crude drug "Piperis Longi Fructus." On the alkaloids of fruits of Piper longum L. *Chemical and pharmaceutical bulletin*, 31(10), pp.3562-3565.
- Venkatasamy, R., Faas, L., Young, A.R., Raman, A. and Hider, R.C., **2004**. Effects of piperine analogues on stimulation of melanocyte proliferation and melanocyte differentiation. *Bioorganic & medicinal chemistry*, 12(8), pp.1905-1920.
- Wang, M., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y., Nguyen, D.D., Watrous, J., Kapon, C.A., Luzzatto-Knaan, T. and Porto, C., **2016**. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature biotechnology*, 34(8), pp.828-837.
- Wei, P., Zhao, F., Wang, Z., Wang, Q., Chai, X., Hou, G. and Meng, Q., **2022**. Sesame (Sesamum indicum L.): A comprehensive review of nutritional value, phytochemical composition, health benefits, development of food, and industrial applications. *Nutrients*, 14(19), p.4079.

- Weil, M., Sing, A.S.C., Méot, J.M., Boulanger, R. and Bohuon, P., **2017**. Impact of blanching, sweating and drying operations on pungency, aroma and color of Piper borbonense. *Food chemistry*, 219, pp.274-281.
- Weil, M., Remize, F., Durand, N., Alter, P., Hoarau, M. and Meile, J.C., **2020**. Effect of processing on microbial safety of wild pepper (Piper borbonense) from Reunion Island. *Food Control*, 111, p.107061.
- Weil, M., Boulanger, R., Morel, G., Servent, A., Sing, A.S.C. and Bohuon, P., **2020**. Quality, typicity and potential valorization of Piper borbonense, a poorly known wild pepper from Reunion Island. *Fruits*, 75(3), pp.95-103.
- Weil, M., PockTsy, J.M.L. and Razafimandimby, H., **2021**. Authenticating wild Piper species (peppers) originating from islands in the Indian Ocean on the basis of morphological, genetic and chemical characteristics. *Phytochemistry*, 190, p.112886.
- Wu, S., Sun, C., Pei, S., Lu, Y. and Pan, Y., **2004**. Preparative isolation and purification of amides from the fruits of Piper longum L. by upright counter-current chromatography and reversed-phase liquid chromatography. *Journal of Chromatography a*, 1040(2), pp.193-204.
- Yamaguchi, I., Matsuda, H., Zhang, H., Hamao, M., Yamashita, C., Kogami, Y., Kon'I, H., Murata, M., Nakamura, S. and Yoshikawa, M., **2014**. Adipogenic effects of piperlonguminine in 3T3-L1 cells and plasma concentrations of several amide constituents from Piper chaba extracts after treatment of mice. *Journal of natural medicines*, 68, pp.74-82.
- Zhang, H., Matsuda, H., Nakamura, S. and Yoshikawa, M., **2008**. Effects of amide constituents from pepper on adipogenesis in 3T3-L1 cells. *Bioorganic & medicinal chemistry letters*, 18(11), pp.3272-3277.
- Yin, W.T., Yang, C.J., Yang, H.J., Hu, B.B., Zhang, F., Liu, H.M. and Miao, H.M., **2024**. Sesame lignans modulate aroma formation in sesame oil through the Maillard reaction and lipid oxidation in model systems. *Food Chemistry*, p.140079.
- Zhang, C., Ji, J., Ji, M. and Fan, P., **2015**. Acetylcholinesterase inhibitors and compounds promoting SIRT1 expression from Curcuma xanthorrhiza. *Phytochemistry Letters*, 12, pp.215-219.

APPENDIX E

Here are reported the spectra of the previously undescribed compounds isolated from *Piper borbonense* (described in PART 4 – Chapter 1 of this manuscript).

APPENDIX D	281
Figure S1: ¹ H NMR spectrum of compound 29 (CDCl ₃ , 600MHz)	282
Figure S2: HSQC 2D NMR spectrum of compound 29 (CDCl ₃ , 600MHz)	282
Figure S3: HMBC 2D NMR spectrum of compound 29 (CDCl ₃ , 600MHz)	283
Figure S4: COSY 2D NMR spectrum of compound 29 (CDCl ₃ , 600MHz)	283
Figure S6: ¹ H NMR spectrum of compound 37 (CDCl ₃ , 600MHz)	284
Figure S7: HSQC 2D NMR spectrum of compound 37 (CDCl ₃ , 600MHz)	284
Figure S8: HMBC 2D NMR spectrum of compound 37 (CDCl ₃ , 600MHz)	285
Figure S9: COSY 2D NMR spectrum of compound 37 (CDCl ₃ , 600MHz)	285
Figure S10: NOESY 2D NMR spectrum of compound 37 (CDCl ₃ , 600MHz)	286
Table S1: HS-GCMS annotated compounds from <i>Piper</i> species.....	286

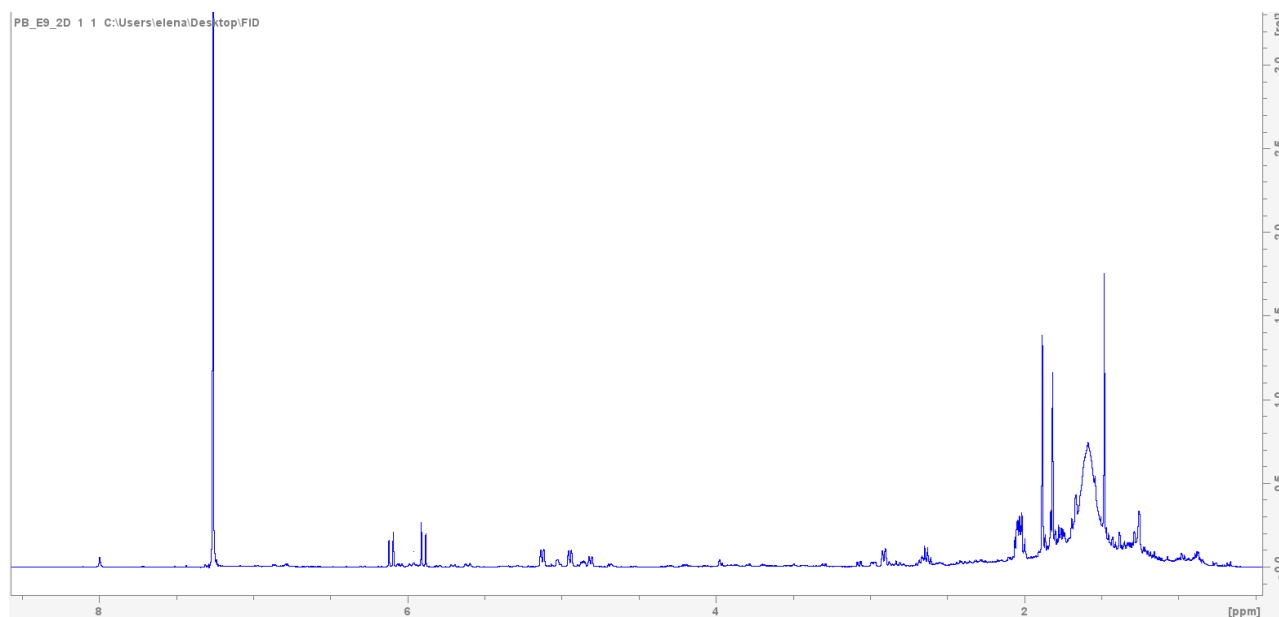
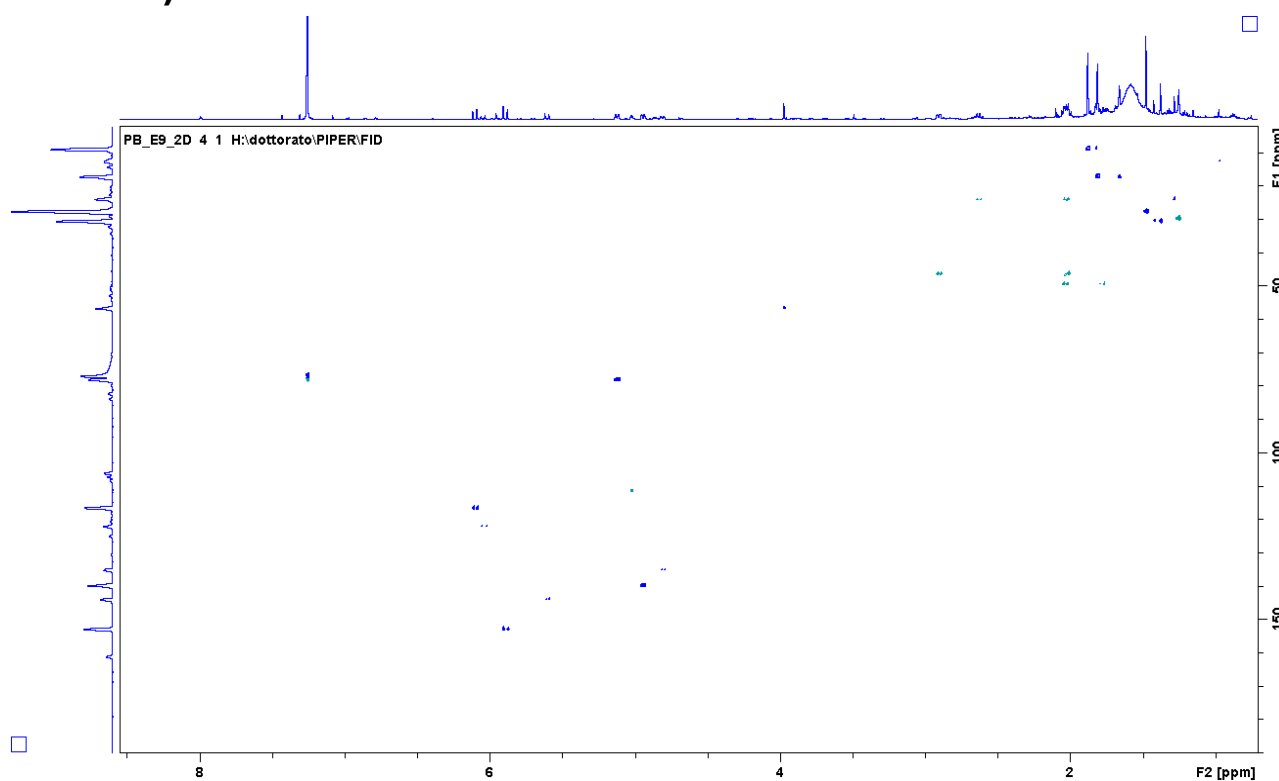
Figure S1: ^1H NMR spectrum of compound 29 (CDCl_3 , 600MHz)**Figure S2: HSQC 2D NMR spectrum of compound 29 (CDCl_3 , 600MHz)**

Figure S3: HMBC 2D NMR spectrum of compound 29 (CDCl₃, 600MHz)

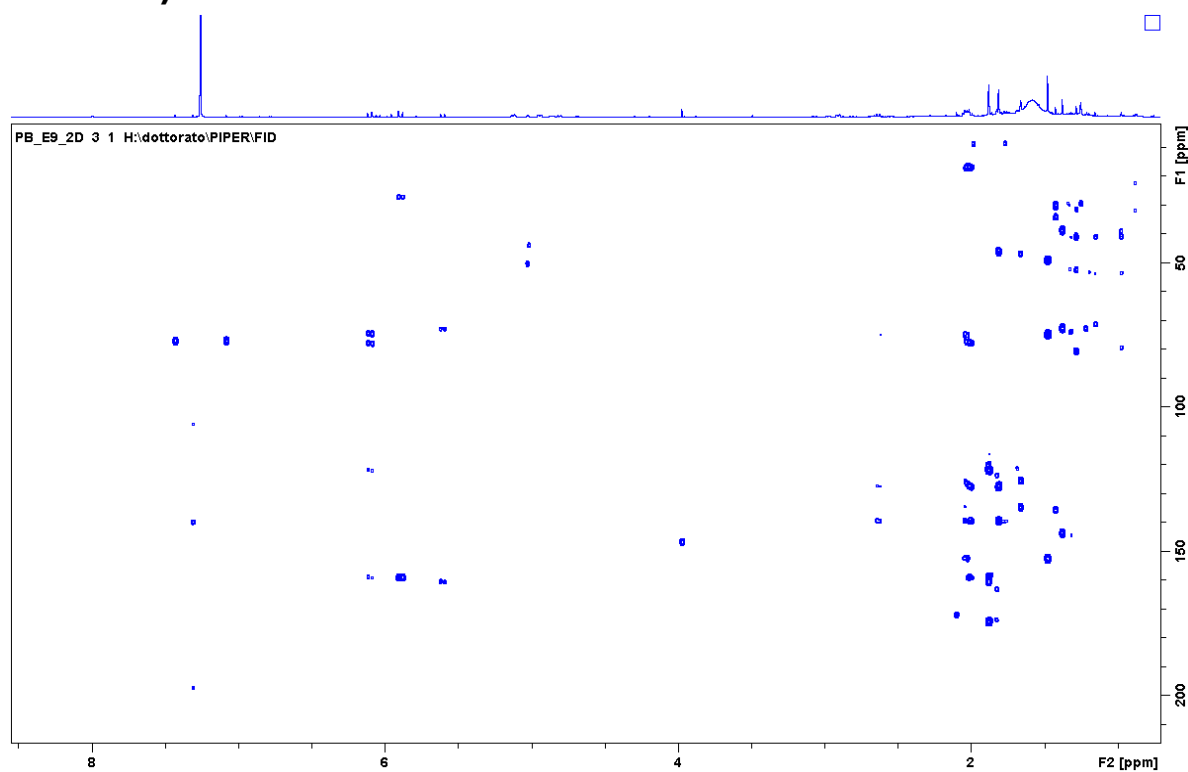


Figure S4: COSY 2D NMR spectrum of compound 29 (CDCl₃, 600MHz)

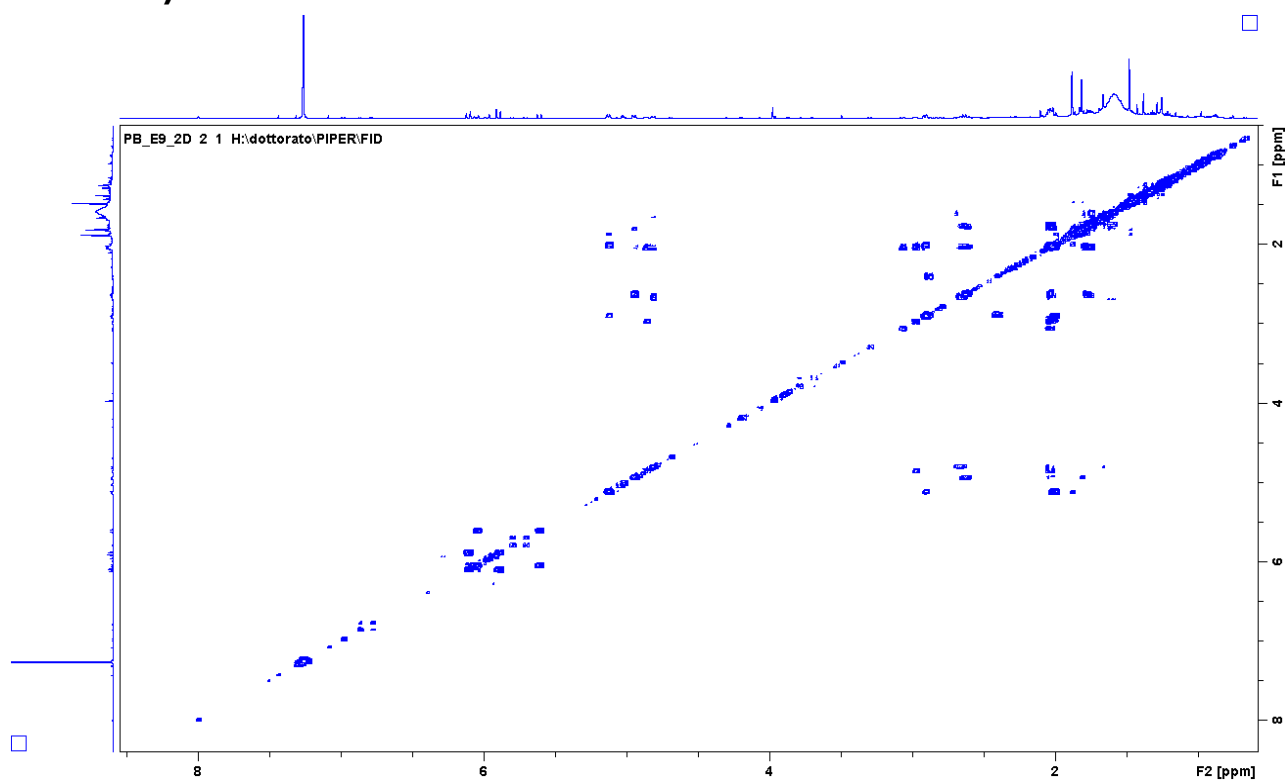


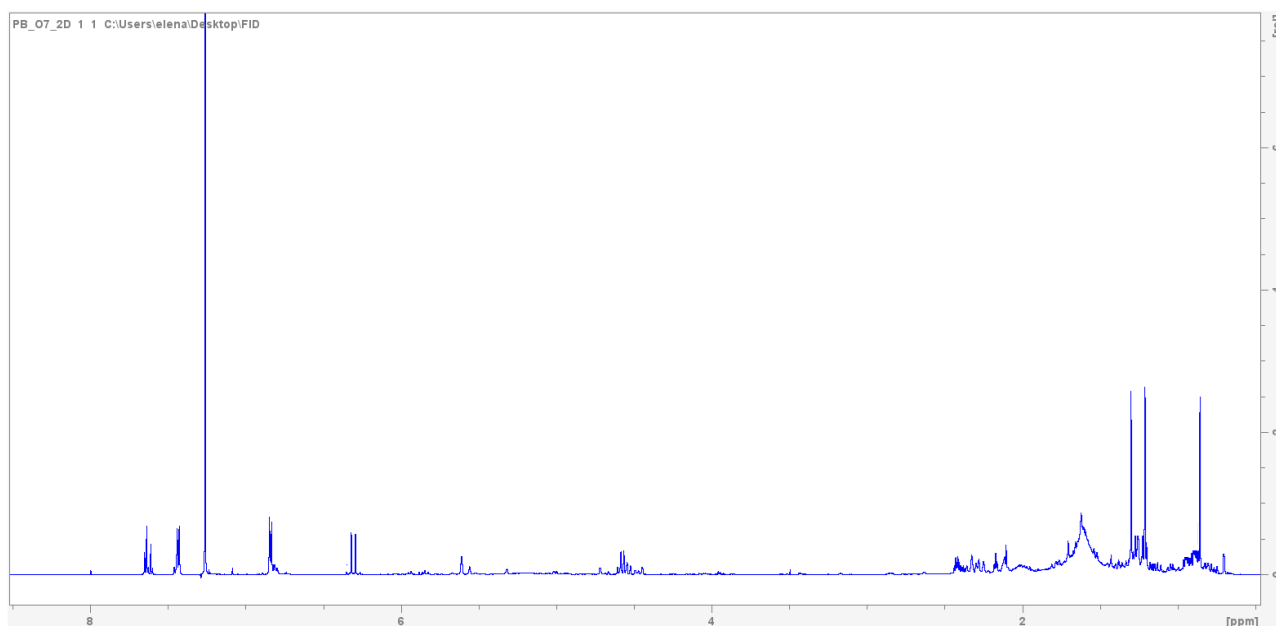
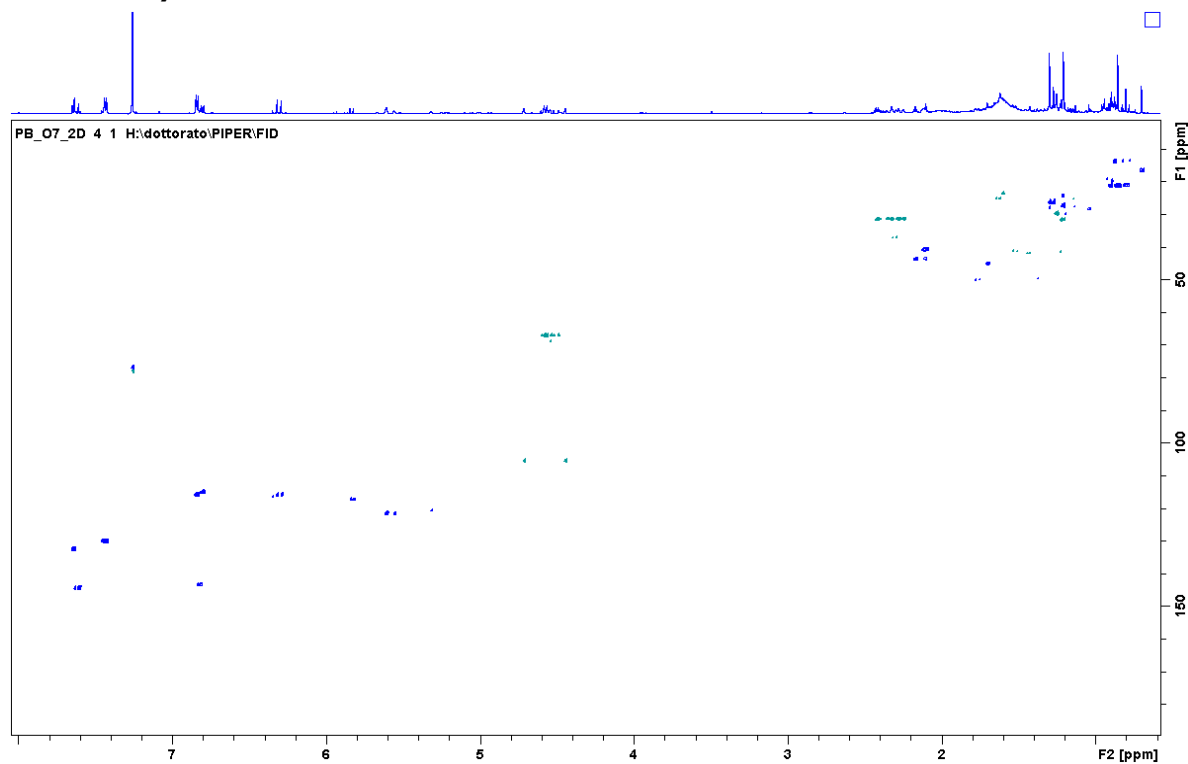
Figure S6: ^1H NMR spectrum of compound 37 (CDCl_3 , 600MHz)**Figure S7: HSQC 2D NMR spectrum of compound 37 (CDCl_3 , 600MHz)**

Figure S8: HMBC 2D NMR spectrum of compound 37 (CDCl₃, 600MHz)

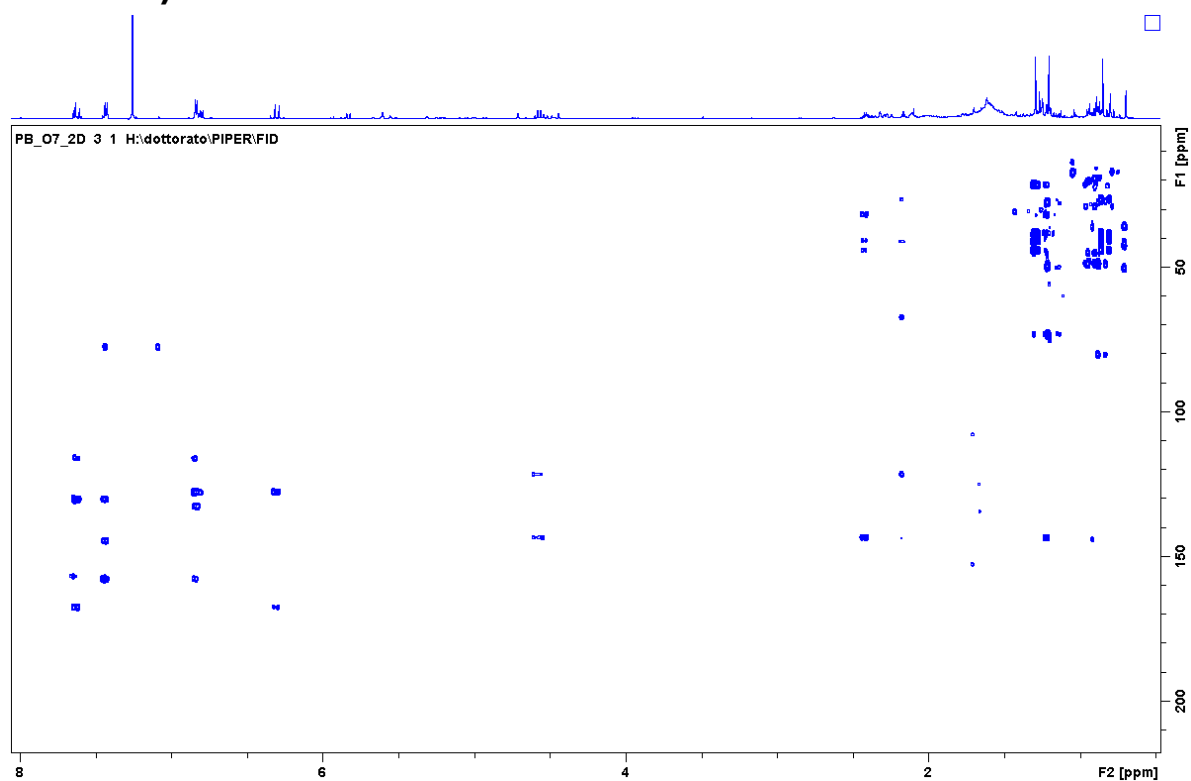


Figure S9: COSY 2D NMR spectrum of compound 37 (CDCl₃, 600MHz)

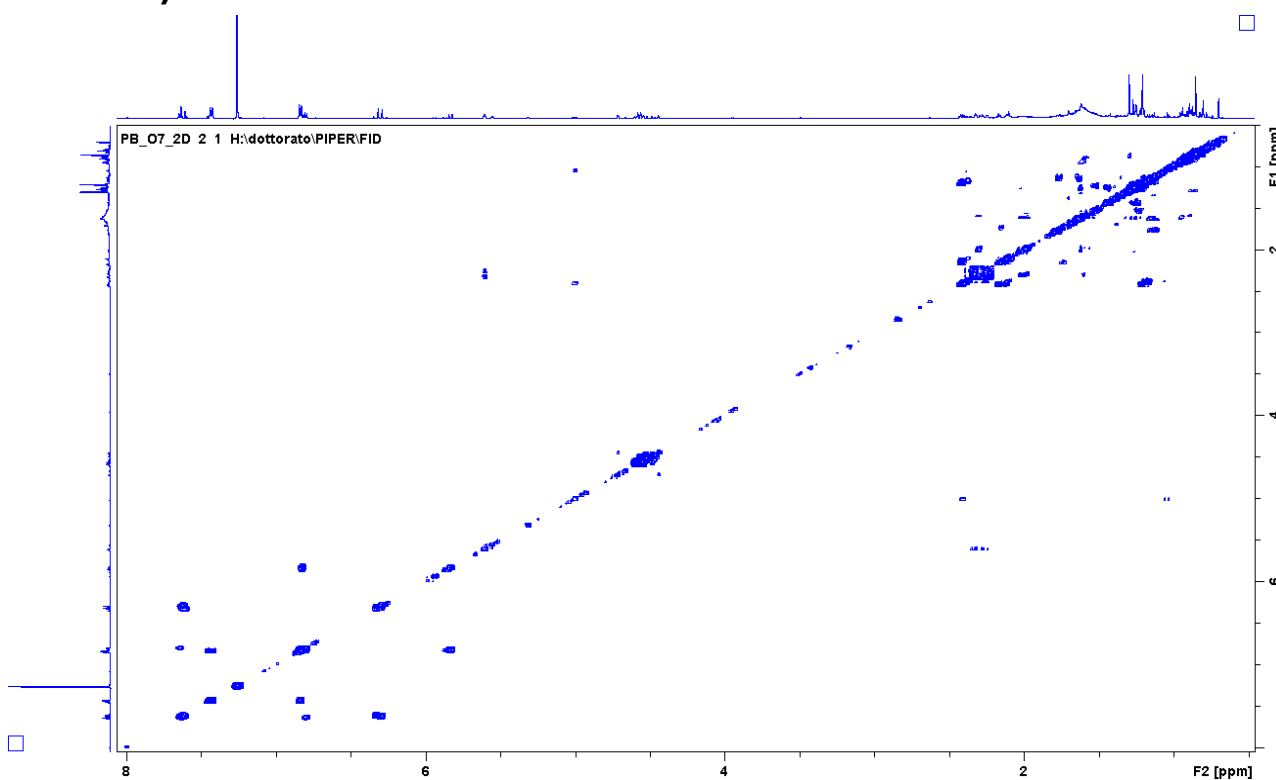


Figure S10: NOESY 2D NMR spectrum of compound 37 (CDCl₃, 600MHz)

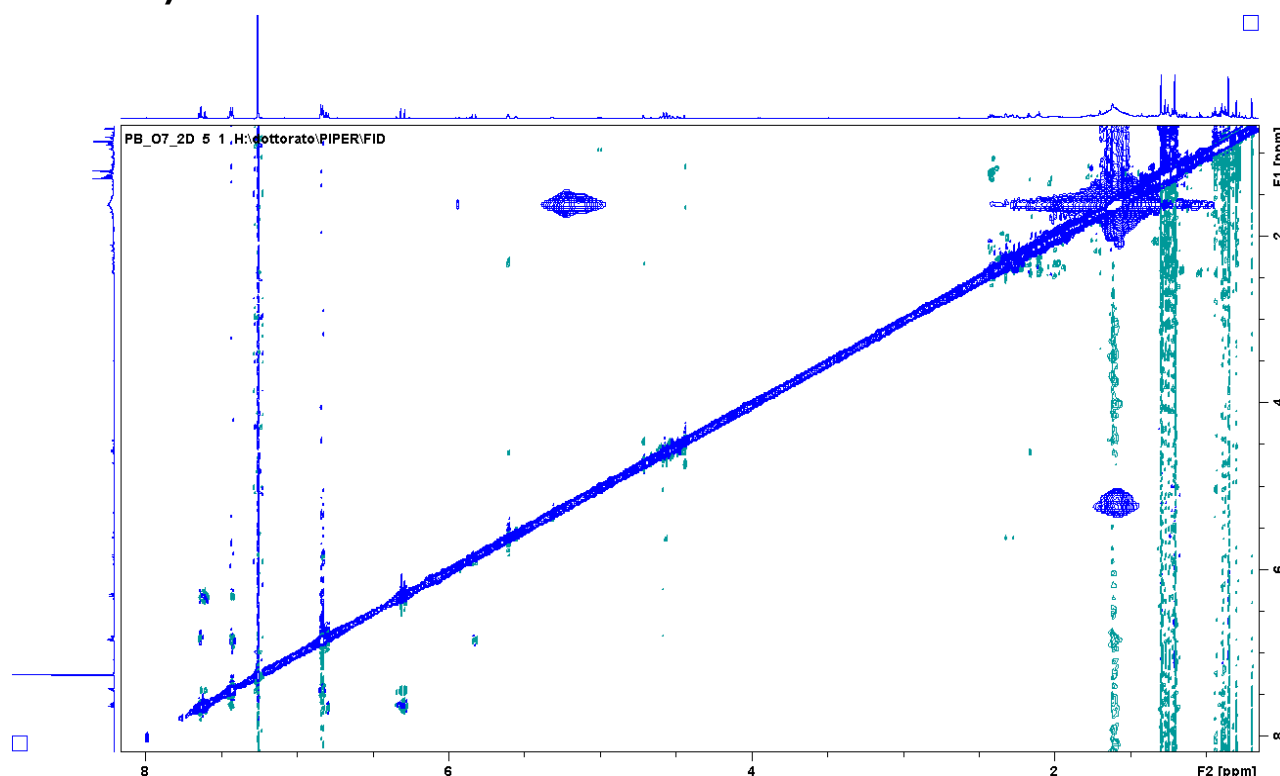


Table S1: HS-GCMS annotated compounds from *Piper* species.

Compound type	compound name	Rt (min)	<i>P. borbonense</i> Area %	<i>P. nigrum</i> Area %
monotrp	α-phellandrene	12.10	0.56	0.18
monotrp	α-pinene	12.47	12.21	9.73
monotrp	bicyclo[2.2.1]heptane, 7,7-dimethyl-2-methylene-	13.02	n.d.	0.02
monotrp	camphene	13.12	0.19	0.11
monotrp	sabinene	13.68	1.33	0.20
monotrp	1,3,5-cycloheptatriene, 3,7,7-trimethyl-	13.76	0.03	0.03
monotrp	β-myrcene	13.92	1.66	2.74
monotrp	β-pinene	14.00	10.48	12.56
monotrp	α-terpinolene	14.46	0.01	0.03
monotrp	δ-3-carene	14.80	40.91	50.23
monotrp	1,3-cyclohexadiene, 1-methyl-4-(1-methylethyl)-	15.02	0.07	0.09
monotrp	O-cymene	15.13	0.04	0.09
monotrp	cyclohexene, 1-methyl-5-(1-methylethenyl)-	15.24	0.01	0.08
monotrp	p-cymene	15.35	5.52	1.86
monotrp	D-limonene	15.43	15.91	17.81
monotrp	β-Thujene	15.58	7.51	0.53
monotrp	eucalyptol	15.66	0.59	0.11
monotrp	trans-thujene	16.04	0.01	0.06
monotrp	γ-terpinene	16.25	0.10	0.14
monotrp	trans-4-Thujanol	16.76	< 0.01	< 0.01

PART 4

monotrp	Isoterpinolene	16.91	0.05	0.32
monotrp	α -Terpinolene	17.08	0.64	0.70
monotrp	Linalool	17.32	0.09	0.05
monotrp	β,β -Dimethylstyrene	17.40	0.01	nd
monotrp	trans-sabinenehydrate	17.75	0.02	nd
monotrp	cosmene	18.04	nd	< 0.01
monotrp	(+)-4-carene	18.43	0.01	nd
monotrp	p-mentha-1,5,8-triene	18.76	< 0.01	< 0.01
monotrp	(+)-2-bornanone	19.48	0.01	nd
monotrp	p-mentha-1,5-dien-8-ol	19.60	0.01	nd
monotrp	trans-sabinenehydrate	20.10	0.01	nd
monotrp	benzenemethanol, $\alpha,\alpha,4$ -trimethyl-	20.29	0.01	nd
monotrp	α -terpinol	20.50	0.01	nd
monotrp	trans-sabinol	20.80	0.07	nd
monotrp	linalyl acetate	21.23	0.05	nd
monotrp	Isopiperitone	22.27	0.02	nd
Aromatic	3,5-dimethoxytoluene	22.37	0.01	nd
phenylprop	safrolene	23.22	0.15	nd
sesquitrp	δ -elemene	23.72	0.02	0.01
monotrp	camphenol	23.96	< 0.01	nd
sesquitrp	α -cubebene	24.05	0.02	< 0.01
sesquitrp	trans- α -Copaene	24.77	0.07	nd
sesquitrp	copaene	24.95	0.11	0.13
sesquitrp	elemene	25.15	0.07	0.03
phenylprop	methyleugenol	25.54	0.11	nd
sesquitrp	eremophylene	25.70	nd	0.01
sesquitrp	ylangene	26.08	0.01	nd
sesquitrp	caryophyllene	26.22	0.44	1.87
sesquitrp	α -guaiene	26.30	0.01	nd
sesquitrp	Copaen-4- α -ol	26.40	0.05	0.01
sesquitrp	muurola-3,5-diene	26.83	0.01	nd
sesquitrp	humulene	27.10	0.01	0.05
ditrp	geranyl- α -terpinene	27.21	0.01	nd
sesquitrp	γ -muurolene	27.33	0.02	nd
sesquitrp	δ -curcumene	27.50	0.02	nd
sesquitrp	germacrene D	27.63	0.13	nd
sesquitrp	α -guaiene	27.70	0.22	nd
sesquitrp	β -bisabolene	27.78	0.07	nd
sesquitrp	isodene	27.82	nd	< 0.01
sesquitrp	epi- α -selinene	27.92	0.03	0.03
sesquitrp	δ -guajene	27.99	0.03	nd
phenylprop	myristicine	28.08	0.01	nd
sesquitrp	7-epi- α -cadinene	28.25	0.03	0.01
sesquitrp	4-epi-cubebol	28.37	0.02	nd
sesquitrp	trans-calamenene	28.51	0.01	nd
phenylprop	trans-isomyristicin	28.67	0.02	nd
phenylprop	elemicin	28.78	0.03	nd

CHAPTER 2: Chemotaxonomy of *Annona* species and analysis of acetogenins

2.1 *Annona* genus

The family Annonaceae, one of the most varied and ancient families of angiosperms, contains 2,100 species in 123 genera and is widely distributed in Africa, Asia and Central America as shown in **Figure 2.1** (Leal et al., 2023). Ethnopharmacological studies have reported their use in folk medicine to treat various conditions, including malaria, gastrointestinal discomforts and parasitic infestations (Frausin et al., 2014), but they also have an important economic role as a source of edible fruits and oil (Biba et al., 2014). One of the most representative and studied genera of this family is the genus *Annona*, which contains approximately 120 species, almost native to tropical America (Quílez et al., 2018).

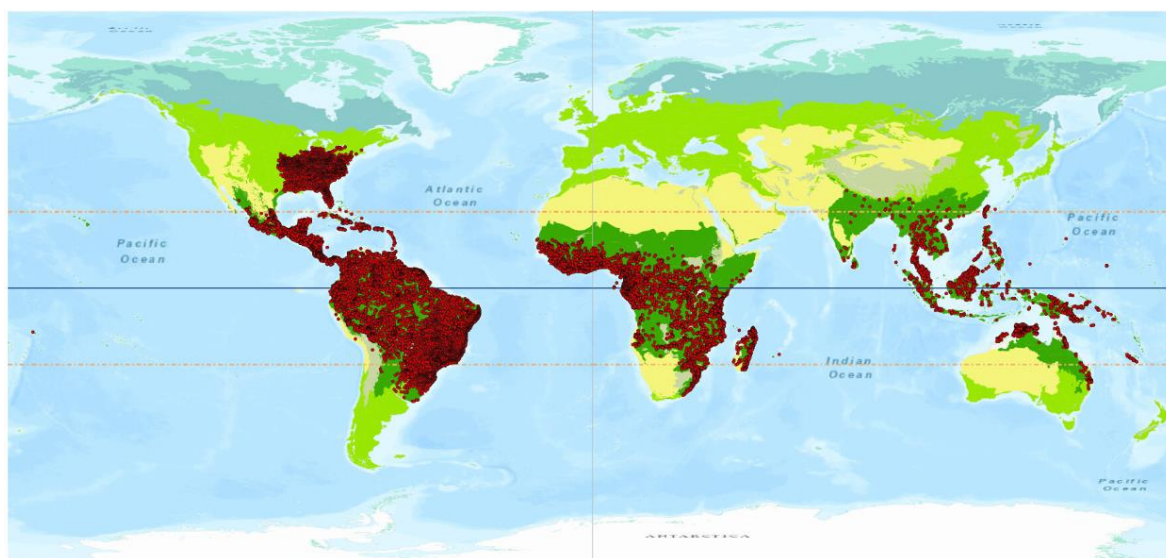


Figure 2.1 Distribution of Annonaceae plants worldwide (Erkens et al., 2023).

Annona sp. plants are generally small trees or shrubs with evergreen, alternate leaves. The flowers may be solitary or grouped in small clusters, while the fruits have a rough or scaly exterior and contain numerous hard, dark brown seeds embedded in a sweet and edible pulp (**Figure 2.2**).

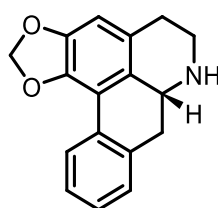


Figure 2.2 Fruits of *Annona* species. A) *A. muricata*; B) *A. squamosa*; C) *A. reticulata*; D) *A. glabra*.

The fruits of *Annona* sp. plants are widely consumed in the native areas and have high nutritional value, including *A. squamosa* L. (sugar apple), *A. muricata* L. (guanabana or soursop), *A. reticulata* L. (custard apple) and *A. glabra* L. (pond apple), which therefore have economic importance among the species (Al Kazman et al., 2022). In addition, different parts of *Annona* plants have been used in traditional medicine as infusions, decoctions or pastes for their medicinal properties. For example, the leaves of *A. muricata* have been used in a variety of remedies in different cultures: in Indonesia they were known for their analgesic properties, while in Burkina-Faso they were used to treat dysentery and fever; the bark and roots of this plant were also used to treat helminth parasitosis; and in Brazil, *A. dioica* was used for wound healing (Leite et al., 2020; Al Kazman et al., 2022). Recent studies have investigated the multiple pharmacological properties of these plants, including anticancer, antidiabetic, antibacterial and antifungal effects (Dev et al., 2021). Furthermore, extracts of *Annona* plants have been shown to have interesting effects on the nervous system, including antinociceptive, antiacetylcholinesterase, anticonsulant and antidepressant activities (Kusmardiyani et al., 2020). However, in Guadeloupe, Martinique, New Caledonia and French Guiana, an excessive consumption of *Annona* fruits has been linked to the pathological development of atypical Parkinson's disease (Lannuzel et al., 2007). The neurotoxicity of these fruits has been linked to the presence of a peculiar type of specialised metabolites found in Annonaceae plants, the acetogenins (das Chagas Lima et al., 2022). Thus, annonaceous acetogenins (AAGs) can be considered as natural neurotoxins present in these food plants. Other bioactive secondary metabolites found in *Annona* are flavonoids, megastigmanes, steroids and alkaloids.

2.1.1 The antidiabetic potential of *Annona* sp. plants

The antidiabetic properties of these plants have been evaluated through in vitro and in vivo experiments, demonstrating the potential of different parts of the plant extracts. For example, an aqueous extract of *A. muricata* leaves was tested in diabetic-induced mice, resulting in weight loss and a reduction in LDL cholesterol (Sasso et al., 2019). This extract has also been shown to reduce in vitro intestinal glucose absorption and improve glucose tolerance in vivo (Guevara-Vásquez et al., 2021). Furthermore, due to its hypoglycaemic and hypolipidaemic properties, the aqueous extract of *A. muricata* has been evaluated for use in combination with metformin, a commonly used drug for its anti-hyperglycaemic effects (Mohammed et al., 2021). The ethanolic leaf extract of *A. muricata* has also been investigated for its hypoglycaemic activity (Opara et al., 2021). In another study on the anti-obesity potential of this plant, the pharmacological effect was attributed to the presence of anonaine, an aporphine alkaloids similar to berberine (**Figure 2.3**), which could downregulate the fat mass and obesity associated gene (FTO) with concomitant upregulation of STAT-3 genes, thereby reducing food intake (Elekofehinti et al., 2020).



anonaine

Figure 2.3 Chemical structure of anonaine

A. squamosa and *A. reticulata* species have also been investigated for their anti-diabetic potential. The methanolic extract of *A. squamosa* leaves has shown in vitro antihypertensive and lipase inhibitory properties (Habib et al., 2022), and preliminary studies conducted on its phenolic-rich extracts have shown that they may be able to prevent the development of type 2 diabetes mellitus by improving insulin sensitivity and reducing body weight gain in high-fat diet-induced obesity (Alkhalidy et al., 2023). Moreover, *A. reticulata* leaf extract showed antihyperglycaemic activity (Pulivarthi et al., 2021). Therefore, given the potential use of these plants in the treatment of MetS, a detailed evaluation of the chemical and safety profile is needed.

2.1.2 Annonaceous acetogenins

Acetogenins are a class of natural products characterised by a polyketide backbone. Depending on their biomass origin, they may have a C-15 skeleton, as reported in some algae (Harizani et al., 2022; Kamanda et al., 2019; Philippus et al., 2018), or they may have a C-35/C-37 skeleton, as reported in plants of the Annonaceae family (Neske et al., 2020). Plant-derived acetogenins are characterised by a methyl-substituted α,β -unsaturated γ -lactone and a methyl terminus separated by a variously decorated long carbon chain. Indeed, the long aliphatic chain can bear unsaturation and hydroxylation, which can give rise to epoxides, tetrahydrofuranic (THF) and tetrahydropyranic moieties, differentiating these compounds into several chemical subtypes (Bermejo et al., 2005). To date, more than 500 acetogenins have been identified, highlighting an impressive chemical diversity between the different type of isomers (Jacobo-Herrera et al., 2017). Some of the common acetogenins structures are reported in **Figure 2.4**.

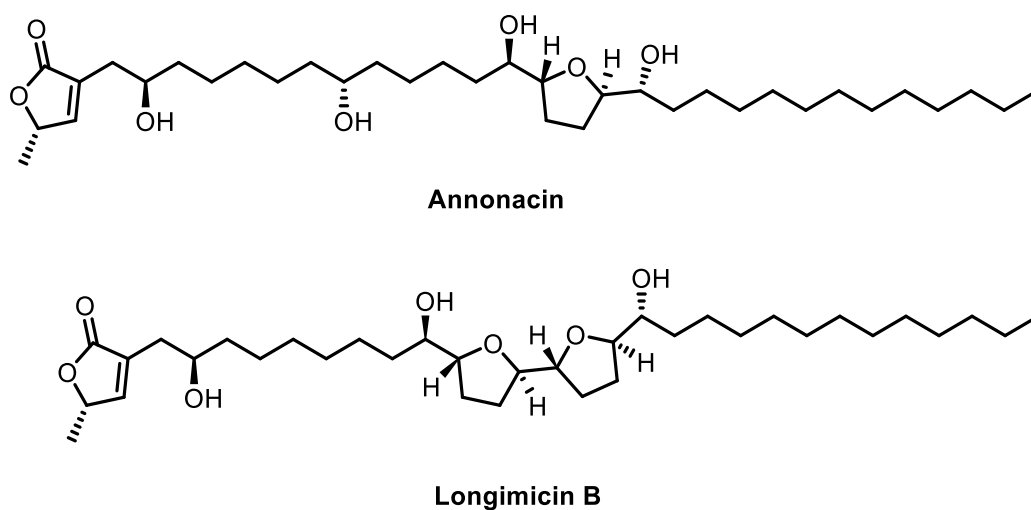


Figure 2.4 Structure of some common acetogenins.

The isolation of the compounds can be challenging due to the strong chemical similarities that occur along this family of analogs, and their structural characterisation by NMR can be tedious due to the presence of the long carbon chain. Therefore, the identification of AAGs in complex mixtures by LC-MS/MS can be of great use to improve their faster annotation. Nevertheless, the protonated or sodium-cationized species lead to only water and/or γ -lactone losses, resulting in poorly informative MS/MS spectra to get the correct metabolite annotated (**Figure 2.5**).

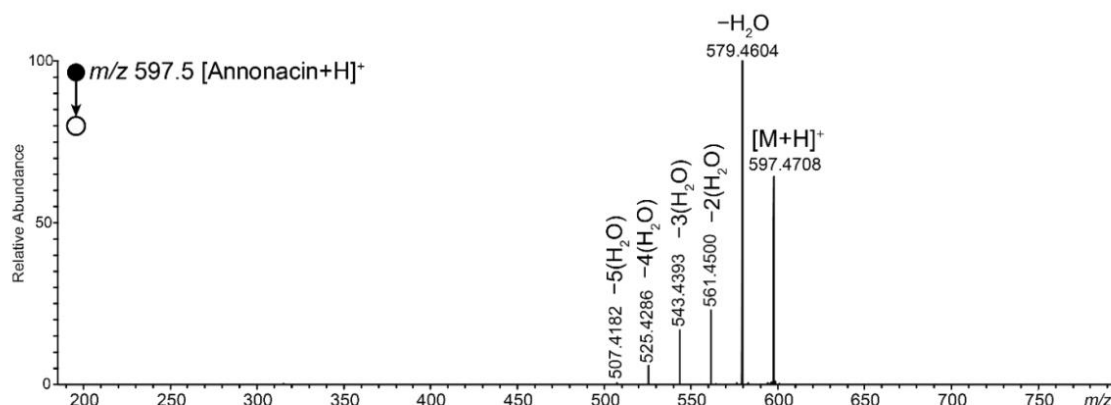


Figure 2. 5 MS/MS spectrum of [Annonacin+H]⁺ at 20 eV from ESI-LTQ-Orbitrap instrument (Poyer et al., 2022)

Recently, lithium (Le Ven et al., 2012) and copper (Poyer et al., 2022) induced fragmentation enriched the MS/MS spectra, providing more information on the structural features of the analysed samples (the MSMS spectra of annonacin with copper and lithium are reported in the Appendix F). In particular, copper post-column infusion resulted in highly informative MS/MS spectra compared to those obtained with lithium, as diagnostic C-C bond cleavages were observed, indicating the positions of THF rings on the acetogenin backbone. Considering the potential of copper-induced fragmentation of AAGs, involving a radical dissociation mechanism, we decided to carry out europium- and cerium-induced fragmentation, as these two lanthanides present several accessible redox states and exhibit high oxygen affinity. Post-column infusions with different salts in order to evaluate their influence on the fragmentation process, which, to our knowledge, has never been described in the literature.

2.2 Chemotaxonomy of French Guyana *Annona* species

Annonaceae is a member of the order Magnoliales, which comprises six families. Often referred to as living fossils, Annonaceae is one of the largest families in the order and has long been partially confused with Eupomatiaceae, which is also in the Magnoliales (Attiq et al., 2017). According to phylogenetic classifications, the Annonaceae family is divided into different subfamilies and tribes, and genus *Annona* belongs to the tribe Annoneae of the subfamily Annonoideae (Mezenes et al., 2021). As discussed above, genus *Annona* is one of the

largest genera in this family and its taxonomy has also undergone changes over the years. For example, some species were separated from the genus *Annona* and grouped in the new genus *Rollinia*, but then returned to the original classification. In fact, *Rollinia* species became redundant with those of *Annona*, and some are now considered synonymous (Leal et al., 2023). Conventional morphological classification based exclusively on botanical features or genetic data can sometimes lead to misinformation, so these studies can be supported by chemotaxonomy. This is a branch of taxonomy that integrates chemical data with traditional methods, using the chemical composition of organisms to classify them. Thus, chemotaxonomy is a powerful strategy to differentiate species that may appear morphologically similar by exploiting different metabolite profiles (Singh, 2016). In the literature, a few chemotaxonomic classifications have been performed on the family Annonaceae, considering diterpenes as molecular descriptors (Scotti et al., 2012; Mezenes et al., 2021). Meanwhile, thanks to the collaboration with the 'Herbier IRD de Guyane', we focused on the *Annona* plants collected in French Guïana to carry out a preliminary chemotaxonomic analysis of the species. Some of the botanical names of the specimens recorded in the herbarium have changed over the years according to different botanists. In addition, in some cases only the leaves of the plant were available, leading to uncertain classification. Thus, it would be interesting to develop and optimise a chemotaxonomic classification procedure that can validate and ensure the taxonomic classification of the specimens reported to the herbarium (Herbier IRD de Cayenne : <https://herbier-guyane.ird.fr/>). Therefore, small pieces (few mm²) of leaves were taken from the herbarium, extracted in methanol and subjected to LC-MSMS analyses. I initially analysed eight different species, including *Annona ambotay*, *A. cuspidata*, *A. foetida*, *A. hypoglauca*, *A. muricata*, *A. previostae*, *A. sericea* and *A. squamosa*, considering at least three samples per species. The LC-MS data were then processed on MZmine 3.6.0 and uploaded to MetaboAnalyst 6.0 for statistical analysis. **Figure 2.6** shows the best result obtained from hierarchical clustering analysis (panel A) and PLSDA (panel B). As can be seen, the classification worked well for almost all samples. In particular, as also shown by the PLSDA analysis, AP2L (*A. previostae*), previously classified as *A. paludosa* in the herbarium, showed a strong similarity to *A. hypoglauca*. It would therefore be interesting to reconsider its classification in the light of the new data.

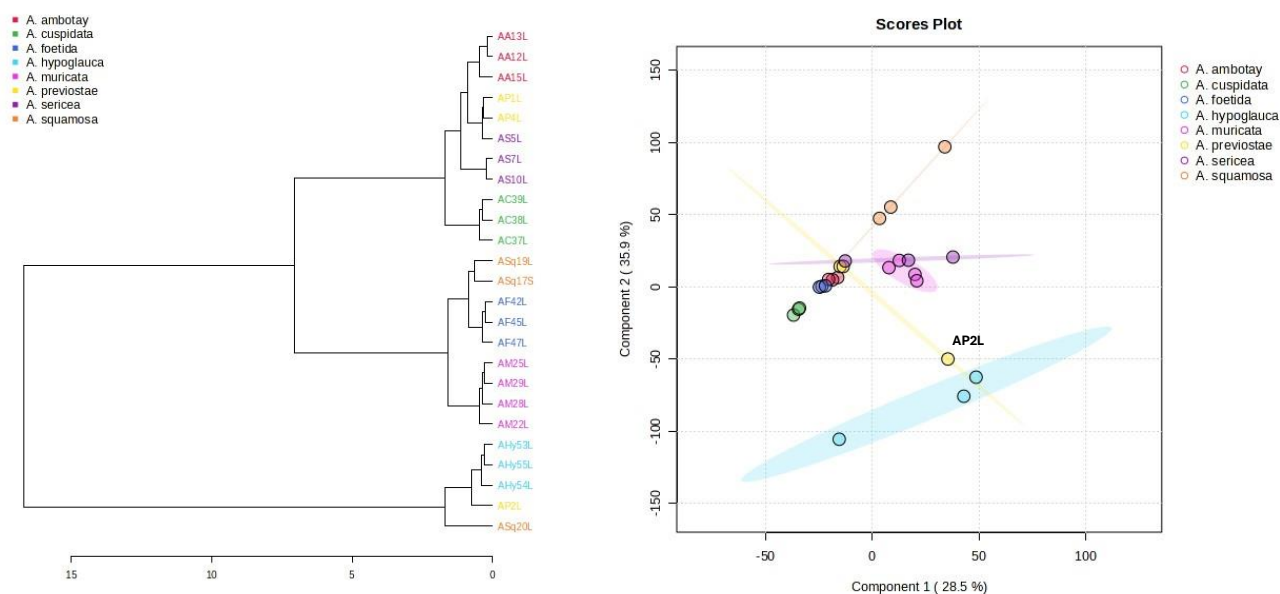


Figure 2. 6 Left) LCMS-based dendrogram obtained from *Annona* species. Right) PLSDA analysis.

This is a preliminary result that we would like to improve. In fact, we would like to include other *Annona* species collected in the herbarium by sampling other species, to extend the chemotaxonomy procedure to other Annonaceae, such as *Guatteria* and *Xylopia* genus, and to identify unknown species.

2.3 Cerium and europium induced fragmentation of acetogenins

2.3.1 Fragmentation pattern analysis of annonacin with different salts

To evaluate the effect of europium and cerium on the fragmentation of AAGs, we first optimised the analytical conditions on the timsToF instrument to obtain the same lithium- and copper-induced MS/MS spectra achieved previously (Le Ven et al., 2012; Poyer et al., 2020) using an annonacin standard available in the laboratory. Annonacin is one of the most common AAGs, mostly present in *Annona muricata*, and it is characterized by a C-35 scaffold bearing four hydroxylic groups and one THF ring. After obtaining the same results, as reported in the Appendix F, five different salts were tested, starting with the same experimental conditions, which were later optimised according to the results obtained. In particular, we tried three different europium salts, including europium (III) chloride, europium (III) nitrate and europium

(III) acetate, along with two cerium salts including cerium chloride and ammonium cerium nitrate (CAN). For each tested salt, we were able to detect annonacin cationized adducts that are summarized in **Table 2.1** and their isotope pattern is reported in Appendix F. In addition, we carefully evaluated the protonated/cationized ratio to evaluate which salt was more suitable.

Table 2.1 Annonacin adducts with the different tested salts.

Salt	Adduct	Chemical formula	Detected m/z	Theoretical m/z	Δ ppm
CuSO ₄	[M-H+Cu ^{II}] ⁺	C ₃₅ H ₆₃ O ₇ Cu ⁺	658.3858	658.3864	-0.9
Eu(CH ₃ CO ₂) ₃	[M+EuCO ₂] ⁺	C ₃₅ H ₆₄ O ₇ EuCO ₂ ⁺	791.3740	791.3743	-0.4
EuCl ₃	[M+EuCl ₂] ⁺	C ₃₅ H ₆₄ O ₇ EuCl ₂ ⁺	817.3208	817.3222	-1.7
Eu(NO ₃) ₃	[M+Eu(NO ₃) ₂] ⁺	C ₃₅ H ₆₄ O ₇ Eu(NO ₃) ₂ ⁺	871.3612	871.3601	1.3
(NH ₄) ₂ [Ce(NO ₃) ₆]	[M+Ce(NO ₃) ₂] ⁺	C ₃₅ H ₆₄ O ₇ Ce(NO ₃) ₂ ⁺	860.3450	860.3457	-0.8
CeCl ₃	[M+CeCl ₂] ⁺	C ₃₅ H ₆₄ O ₇ CeCl ₂ ⁺	806.3081	806.3078	0.4

Europium acetate gave a cationised adduct at m/z 791.3740, suggesting C₃₅H₆₄O₇EuCO₂⁺ as molecular formula, whereas europium chloride and europium nitrate showed adducts at m/z 817.3208 and m/z 871.3612, suggesting C₃₅H₆₄O₇EuCl₂⁺ and C₃₅H₆₄O₇Eu(NO₃)₂⁺ as molecular formula, respectively. Each of the europium adduct displayed a complex isotopic pattern where the M+2 ion peak was the most intense of the isotopic cluster. This is due to the abundance of europium isotopes, where ¹⁵³Eu has 52.2% of natural abundance compared to 47.8% of ¹⁵¹Eu. Moreover, among europium salts, europium acetate displayed a high sodium/cationised ratio, as the MS spectrum showed a high intense sodium adduct at m/z 619.4522 and a low intense cationised adduct. On the other hand, cerium salts exhibit a simpler isotopic pattern, given the 88.4% natural abundance of ¹⁴⁰Ce. In particular, ammonium cerium nitrate exhibited a cationised adduct at m/z 860.3450, whereas cerium chloride gave m/z 806.308, thus implying C₃₅H₆₄O₇Ce(NO₃)₂⁺ and C₃₅H₆₄O₇CeCl₂⁺ as molecular formula, respectively. It should be noted that, to achieve the MS/MS spectra of annonacin with europium salts it was necessary to select a preferred mass list to induce fragmentation on the correct parent ion. Among the europium and cerium series, we did not observe a significant difference in the fragmentation pattern between the salts, as shown in Appendix F. Thus, as examples, here are reported the data of CAN-induced fragmentation of annonacin in **Figure 2.7**, panel A and **Table 2.2**, while the ones obtained with europium nitrate are reported in **Figure 2.7**, panel B and **Table 2.3**. The detail analysis of the MS/MS spectra showed that the detected fragments were diagnostic to rebuild the structure of annonacin.

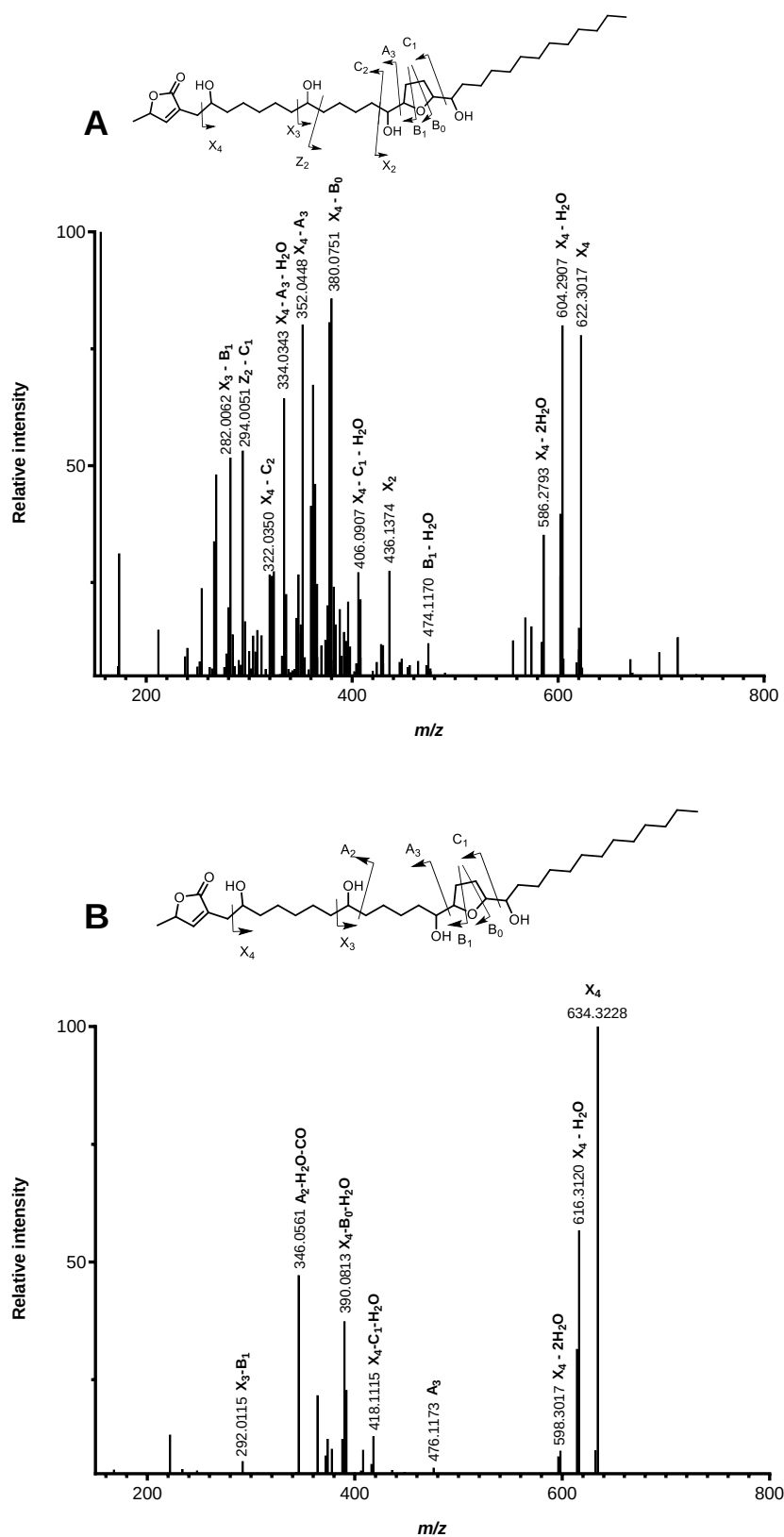


Figure 2. 7 Panel A) MS/MS (80eV) spectrum of annonacin with CAN post-column infusion; panel B) MS/MS (70eV) spectrum of annonacin with $\text{Eu}(\text{NO}_3)_3$ post-column infusion. The nomenclature of the fragments follows the rules established by Laprevoté et al., 1994.

Table 2.2 Annonacin standard fragments with CAN post-column infusion

Fragment ion (<i>m/z</i>)	Type of fragment	Chemical formula	Theoretical <i>m/z</i>	Δ ppm
622.3017	X ₄	C ₂₉ H ₅₄ O ₅ Ce ⁺	622.3020	-0.5
604.2907	X ₄ -H ₂ O	C ₂₉ H ₅₂ O ₄ Ce ⁺	604.2914	-1.2
586.2793	X ₄ -2H ₂ O	C ₂₉ H ₅₀ O ₃ Ce ⁺	586.2801	-1.4
474.1170	B ₁ -H ₂ O	C ₂₀ H ₃₀ O ₄ Ce ⁺	474.1193	-4.8
436.1374	X ₂	C ₁₈ H ₃₂ O ₃ Ce ⁺	436.1400	-6.0
406.0907	X ₄ -C ₁ -H ₂ O	C ₁₆ H ₂₆ O ₃ Ce ⁺	406.0931	-5.9
380.0751	X ₄ -B ₀ -H ₂ O	C ₁₄ H ₂₄ O ₃ Ce ⁺	380.0774	-6.0
352.0448	X ₄ -A ₃	C ₁₂ H ₂₀ O ₃ Ce ⁺	352.0461	-3.7
334.0343	X ₄ -A ₃ -H ₂ O	C ₁₂ H ₁₈ O ₂ Ce ⁺	334.0356	-3.9
322.0350	X ₄ -C ₂	C ₁₁ H ₁₈ O ₂ Ce ⁺	322.0356	-1.9
294.0051	Z ₂ -C ₁	C ₉ H ₁₄ O ₂ Ce ⁺	294.0043	2.7
282.0062	X ₃ -B ₁	C ₈ H ₁₄ O ₂ Ce ⁺	282.0043	6.7

Table 2.3 Annonacin standard with Eu(NO₃)₃ post-column infusion

Fragment ion (<i>m/z</i>)	Type of fragment	Chemical formula	Theoretical <i>m/z</i>	Δ ppm
634.3228	X ₄	C ₂₉ H ₅₅ O ₅ Eu ⁺	634.3242	-2.2
616.3120	X ₄ -H ₂ O	C ₂₉ H ₅₃ O ₄ Eu ⁺	616.3137	-2.7
598.3017	X ₄ -2H ₂ O	C ₂₉ H ₅₁ O ₃ Eu ⁺	598.3031	-2.4
476.1173	A ₃	C ₁₈ H ₂₉ O ₅ Eu ⁺	476.1208	-7.3
418.1115	X ₄ -C ₁ -H ₂ O	C ₁₆ H ₂₈ O ₃ Eu ⁺	418.1153	-9.1
390.0813	X ₄ -B ₀ -H ₂ O	C ₁₄ H ₂₃ O ₃ Eu ⁺	390.0854	-10.5
346.0561	A ₂ -H ₂ O-CO	C ₁₂ H ₁₉ O ₂ Eu ⁺	346.0578	-4.1
292.0115	X ₃ -B ₁	C ₈ H ₁₃ O ₂ Eu ⁺	292.0109	2.0

Annonacin adducts underwent to the neutral loss of two units of HNO₃, and the subsequent fragments allowed complete reconstruction of the structure. In fact, in CAN post-column infusion, the B₁-H₂O fragment at *m/z* 474.1170 indicated the position of the THF ring, while the X-type fragments at *m/z* 622.3017 (X₄) and at *m/z* 436.1374 (X₂) indicated the position of the hydroxylic groups at positions 4 and 15, respectively. Furthermore, the combination of several internal fragments detected, such as Z₂-C₁ at *m/z* 294.0051 and X₄-A₃ at *m/z* 352.0448, indicated the positions of the remaining functionalisation. The europium-induced fragmentation also led to a confident annotation of annonacin. Indeed, the X₄ fragment at *m/z* 634.3228 and A₃ at *m/z* 476.1173 indicated the position of the hydroxyl groups at positions 4 and 15. The internal fragments X₃-B₁ and X₄-C₁-H₂O indicated the position of the THF ring and the hydroxyl group at positions 10 and 20.

2.3.2 Analyses of complex mixtures

In order to test the developed method on complex mixtures, we perform the analyses on the methanolic extracts obtained from the seeds and leaves of *A. muricata*. The Base Peak Chromatograms (BPC) of the extracts is shown in **Figure 2.8**. As can be seen from the BPC comparisons, the extracts differ mainly in the first part, where the alkaloids elute (Rt 0-10 min), showing a selective abundance in the leaves, whereas the seeds are richer in acetogenins (20-35 min). We did not observe any cationic adduct between alkaloids and europium or cerium salts, which may be due to the affinity of these ions for oxygenated compounds (Comuzzi et al., 2002), resulting in selectivity for acetogenins. Thus, we focused on the seed extract.

Due to the isotopic pattern profile of europium, the analyses of the mixtures are more complicated as fragmentation is induced on the highest peak of the cluster. Thus, by using europium salts, this would lead in a tricky annotation procedure unless a preferred mass list is previously set, which is laborious and inconvenient for untargeted analyses. On the contrary, CAN infusion posed no such problem and can be used by simply setting an appropriate m/z range for the parent ions. The same analysis was performed using copper sulfate to compare the annotation obtained on the detected AAGs. The compounds annotated until now in the seed extract are reported in **Table 2.4**.

PART 4

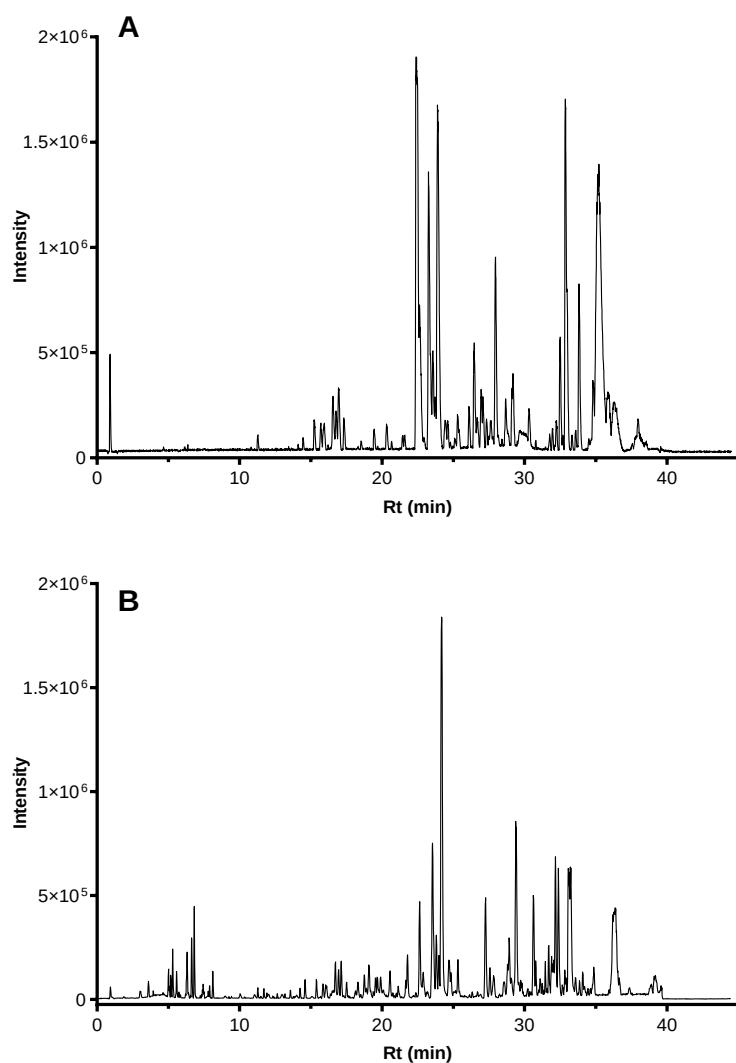
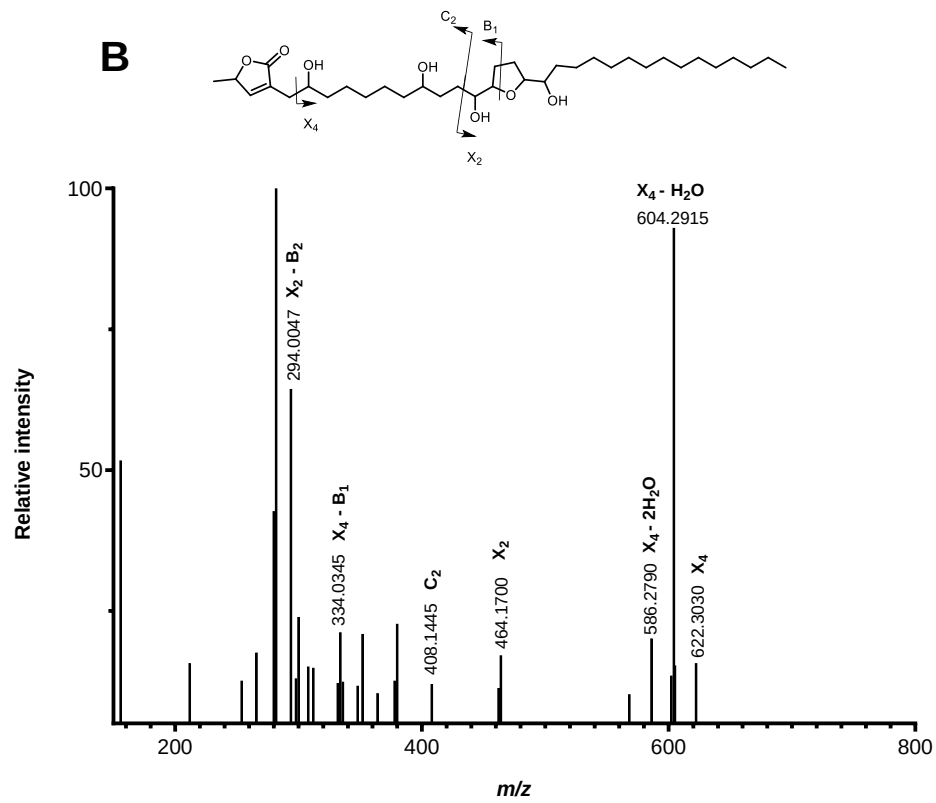
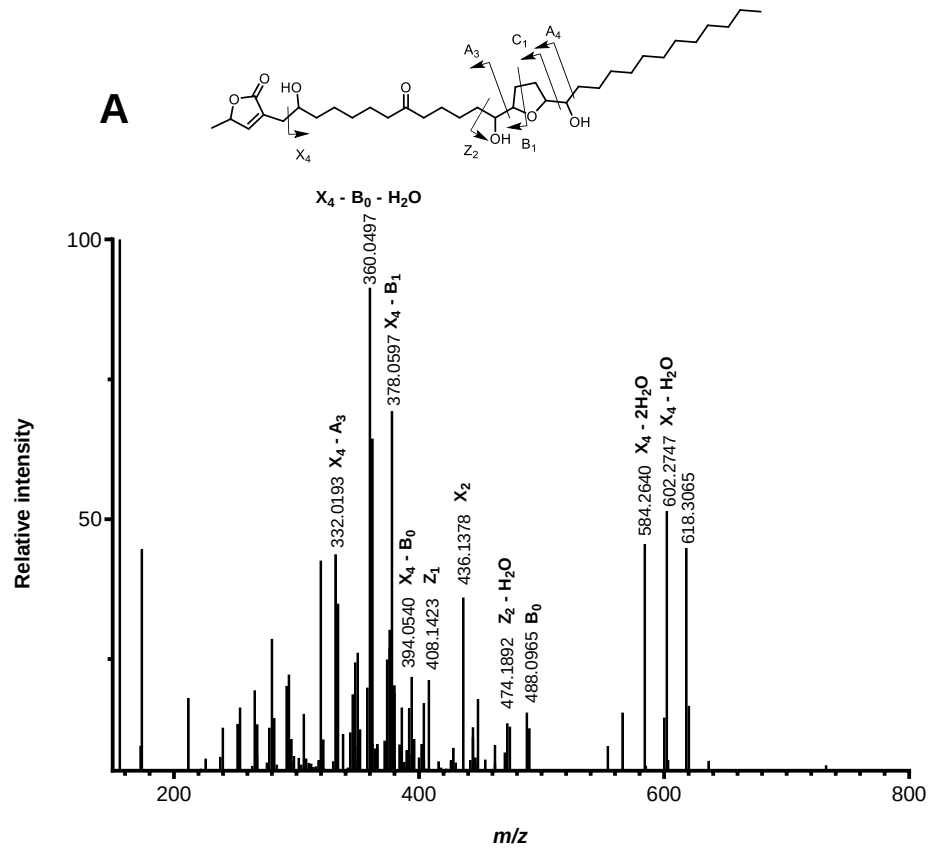


Figure 2.8 BPC of *A. muricata* seeds (panel A) and leaves (panel B) extracts

Table 2.4 Compounds annotated with CAN post-column infusion until now.

Rt (min)	[M+Ce(NO ₃) ₂] ⁺ m/z	Intensity	Chemical formula	Predicted m/z	Δppm	Compound name
22.4	860.3473	2.8 x 10 ⁵	C ₃₅ H ₆₄ O ₇ Ce(NO ₃) ₂	860.3457	1.8	Annonacin
23.3	858.3318	2.8 x 10 ⁵	C ₃₅ H ₆₂ O ₇ Ce(NO ₃) ₂	858.3301	2.0	Annonacinone
23.8	860.3469	1.8 x 10 ⁴	C ₃₅ H ₆₄ O ₇ Ce(NO ₃) ₂	860.3457	1.4	Goniothalamycin
24.6	888.3781	1.5 x 10 ⁴	C ₃₇ H ₆₈ O ₇ Ce(NO ₃) ₂	888.3770	1.2	Annomutacin (or isomers)
28.0	844.3523	1.1 x 10 ⁵	C ₃₅ H ₆₄ O ₆ Ce(NO ₃) ₂	844.3508	1.7	Murisolin

PART 4



PART 4

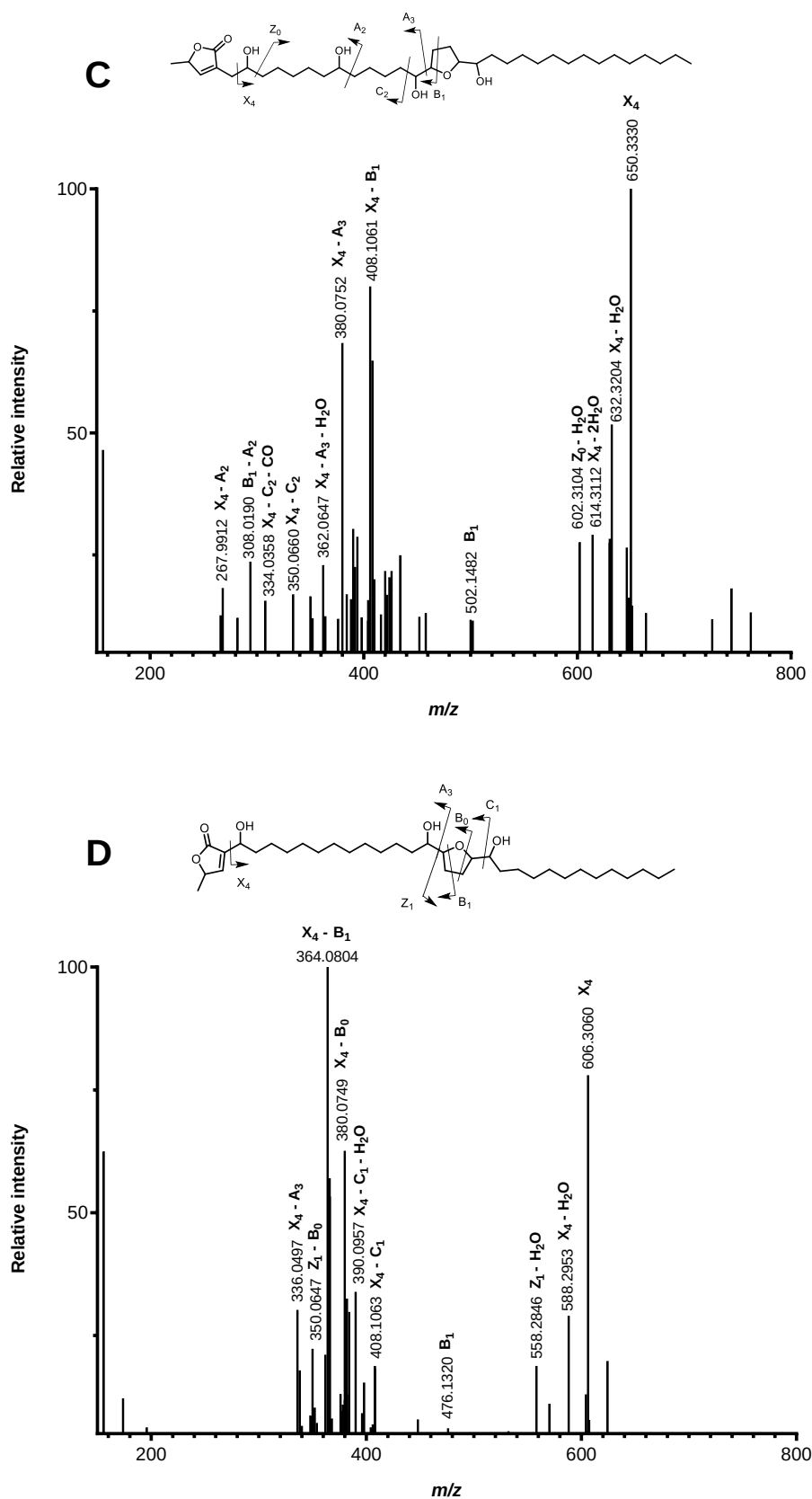


Figure 2.9 MS/MS spectra (80eV) of the compounds annotated with CAN post-column infusion until now. Panel A) Annonacinone; panel B) Goniotalamicin; panel C) Annonutacin; panel D) Murisolin

Interestingly, for annonacinone, another major acetogenin in *A. muricata*, cerium-induced fragmentation gave an advantage over that obtained with copper. Taking into account that annonacinone is the structural isomer of longimicin B, copper sulphate infusion did not yield fragments that could distinguish between the two isomers. On the contrary, cerium-induced fragmentation detected Z-type fragments, in particular Z_2-H_2O at m/z 474.1892, which excluded longimicin B. The annonacinone fragments detected with CAN and copper sulphate are given in Appendix F, together with the MS/MS spectrum of annonacinone obtained from the copper column infusion, while that obtained with CAN is given in **Figure 2.9**, panel A.

In conclusion, among the salts tested, CAN appears to be the most suitable for annotating acetogenins in complex mixtures and we look forward to continuing the annotation and hypothesising a mechanism of fragmentation.

2.3.3 Ion Mobility analyses of acetogenins.

Identification of acetogenins in complex mixtures is often complicated by co-elution of different isomers. The application of Ion Mobility Spectrometry (IMS) can give an advantage in separating them and identifying co-eluting metabolites. In fact, it can provide interesting information about the conformation of the ions in the gas phase, which depends on their stereochemical features. Thus, different stereoisomers will lead to a different conformation and a different mobility value. I first optimised the mobility parameters using an annonacin standard, whose mobility and CCS values are reported in **Table 2.5**. From the Extracted Ion Mobiliograms (EIM), annonacin shows only one conformation in the gas phase with each salt. Considering the atomic radius of the salts and the recorded mobility values, it could be assumed that annonacin adopts a cage-like conformation, which is not surprising given the ability of the considered cations to coordinate oxygen groups. However, the adducts with europium and cerium salts appear to be quite complex due to the presence of nitrate groups and specific calculations would be required to validate this hypothesis.

Table 2.5 Annonacin mobility and CCS value with different salts

Adduct	Molecular formula	Mobility value (1/ k_0)	CCS value ($\text{\AA}^2$)
$[M+H]^+$	$C_{35}H_{65}O_7^+$	1.246	255.3
$[M+Na]^+$	$C_{35}H_{64}O_7Na^+$	1.290	264.1
$[M+H+Cu]^+$	$C_{35}H_{63}O_7Cu^+$	1.270	263.7
$[M+Eu(NO_3)_2]^+$	$C_{35}H_{65}O_7Eu(NO_3)_2^+$	1.340	272.6
$[M+Ce(NO_3)_2]^+$	$C_{35}H_{65}O_7Ce(NO_3)_2^+$	1.340	272.7

In order to evaluate the use of ion mobility on complex mixtures of acetogenins, we tested the *A. muricata* seed extract with copper sulphate post-column infusion. At Rt 22.8 min, two different mobility values were detected for annonacin cationised adduct with copper. The MS/MS spectra were registered for each of the detected parent ion, showing one identical to annonacin for mobility and fragments, while the other showed a different mobility but almost identical MS/MS spectra to annonacin. As the annonacin standard gave only one conformation in the gas phase, we can assume that this could be a stereoisomer of annonacin, but further experiments would be needed to confirm this hypothesis. Furthermore, at Rt 23.9 min, two different mobility values were detected for the copper adduct of goniiothalamycin, and in this case the two gave different MS/MS spectra. As shown in **Figure 2.10**, as expected, one corresponded to goniiothalamycin (panel B), while the annotation for the other is under evaluation.

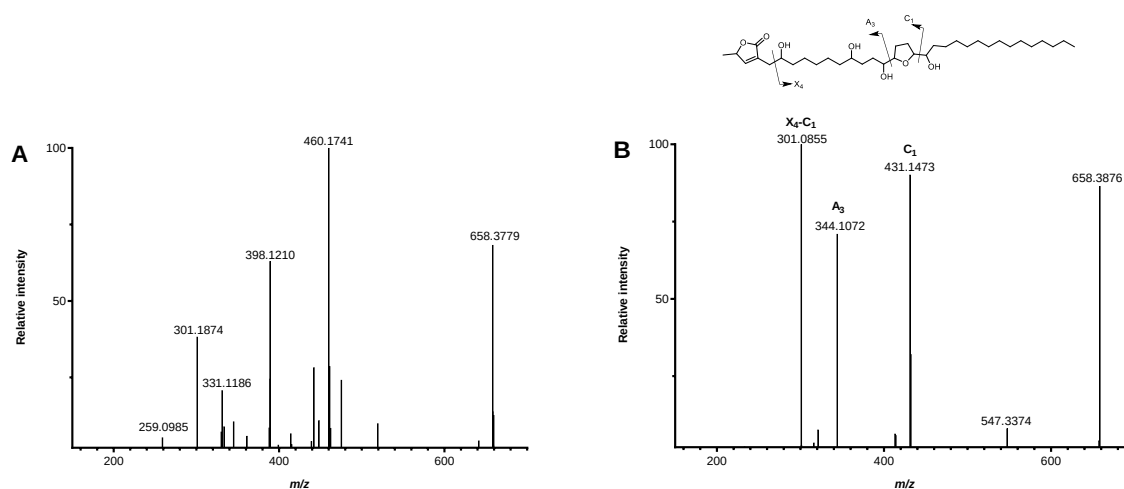


Figure 2.10 MS/MS spectra of compounds at Rt 22.9 min

The use of ion mobility could be a key tool to discriminate between acetogenin isomers in complex mixtures, and the different orientation of the hydroxylic groups could give rise to different conformations and thus different mobility values, providing insight into the stereochemical features of these compounds. However, further experiments would be required to prove its use.

2.4 Experimental Section

2.4.1 Chemicals

Cerium(III) chloride heptahydrate was purchased from Lancaster synthesis. Ammonium cerium(IV) nitrate was purchased from Acros. Europium(III) acetate hydrate and europium(III) nitrate pentahydrate were purchased from Sigma-Aldrich. Europium chloride, europium nitrate, copper sulfate, lithium chloride, methanol (MeOH LC-MS grade) and acetonitrile (MeCN LC-MS grade) were purchased from VWR (Rosny-sous-bois, France). Deionized water (18.2 MΩ) was obtained from PureLab Quest (Veolia water technologies, Wissous, France). Annonacin (C₃₅H₆₄O₇), was obtained as previously described (Laboureur et al., 2017). Stock solutions were prepared in methanol at a concentration of 100 μM.

2.4.2 Plant material and sample preparation.

The samples of *Annona* species used for the classification analysis were taken in the French Guiana Herbarium (Herbier IRD de Cayenne : <https://herbier-guyane.ird.fr/>)

The samples of *A. muricata* used for the MS/MS analysis were collected in French Guiana in October 2023 by Jean-Louis Smock. All the samples (around 15.0 mg of leaves and 60.0 mg of seeds) were sonicated for 1 hour with 1.5 mL of MeOH. The supernatant was filtrated with 0.45 μm filters and dried at rotavapor. The dried extracts were solubilized in MeOH, from which the sample for LC-MS/MS was prepared (1mg/mL).

2.4.3 LC-MSMS and ion mobility analyses

Liquid chromatography-tandem mass spectrometry with high resolution (LC-HR-MS/MS) experiments were performed on an 1260, Infinity II LC system (Agilent Technologies, France) coupled to a timsTOF mass spectrometer (Bruker, France) equipped with an electrospray ionization (ESI) source operated in the positive mode. The chromatographic separation was carried out on a Poroshell 120SB C18 (2.1x150 mm 2.7 μm) column heated at 40 °C, the flow rate was set at 400 μL/min with water as solvent A and acetonitrile as solvent B with 0.1% of formic acid in A and B. 2 μL of annonacin standard (10 μg/mL in MeOH) were injected using the following gradient: 0 min (5%B), 7 min (100%B), held for 2 minutes, 9.5 min (5%B) and 3 minutes of equilibration time. *A. muricata* extracts were injected (volume injected 5 μL) using the following gradient: 0 min (5% B), 10 min (50% B), 30 min (100%B), held for 7 minutes, 37.5 min (5%B) and 7 minutes of equilibration time. To perform the post-column infusion, we used a split

connected to a Waters 2690 HPLC separation module, infusing a 100 μM solution of the salts at 50 $\mu\text{L}/\text{min}$ flow rate.

Ion source parameters were set as follows: capillary voltage 4.2 kV, nebulizer 3.0 bar, dry gas 10.0 L/min, desolvation temperature 350°C. Nitrogen was used as drying gas in the source, in the mobility cell and for collision experiments. The instrument was daily calibrated in the 50-1350 m/z range using ESI-L low Concentration Tuning Mix solution (Agilent Technology) and for the mobility experiments (ions at m/z 118.0863, m/z 322.0481 and m/z 622.0290). The mobility range was set using detect mode ($1/K_0$) range of 1.12 - 1.45 $\text{V}\cdot\text{s}/\text{cm}^2$. The ramp time was set to 100 ms. For Auto MS/MS experiments, parameters were set as follows: precursor ion range m/z 500-900, No. of precursor 5, spectral rate 10.0 Hz, fixed CID 70 eV or 80 eV.

Elemental compositions, ion mobility and collision cross-section (CCS) for all ions were obtained using the instrument software DataAnalysis.

2.4.4 PLSDA and Hierarchical Cluster Dendrogram generation.

The LC-MS/MS data were acquired using the same chromatographic conditions and MS described above, but just the LC-MS data were considered for these analyses. Data were converted from .d to .mzML using MSConvert part of the ProteoWizard 3.0 package before being processed on Mzmine 3.6.0 software using the following parameters: mass detection noise level 1000 for mass level 1; the ADAP chromatogram builder was used with a minimum of 5 consecutive scans with a minimum intensity of 1000 and a minimum absolute height of 1000, the m/z tolerance was set at 0.005 or 15 ppm; deconvolution was performed with the ADAP resolver with a signal to noise ratio of 30, a minimum feature height of 1000, a coefficient/area threshold set at 10, a peak duration range of 0.0 to 0.8 min and the RT wavelet range set between 0.02 and 0.09; the isotope peak grouper was used with the m/z tolerance set at 0.005 or 15 ppm, a RT tolerance of 0.1 min and the maximum number of charges set at 2; the join aligner was used to align peak lists, the m/z tolerance was set at 0.005 or 15 ppm, weight for m/z was set at 80, while RT was set at 20, and RT tolerance was set at 0.3 minute; the gap filling was obtained setting 30% as intensity tolerance, m/z tolerance set at 0.005 or 15 ppm, and RT tolerance was set at 0.3 minute. Finally, the aligned gap-filled feature list was exported as an .csv file that was uploaded to MetaboAnalyst 6.0 to generate the 2D scores plot of PLSDA and hierarchical clustering analysis (HCA). The following steps were performed: data were filtered setting the reliability filter greater than 25%; variance filter as interquartile range at 40%; the

data were normalized by sum. HCA dendrogram was achieved setting the distance measure as Spearman and cluster algorithm as Ward.

References

- Al Kazman, B.S., Harnett, J.E. and Hanrahan, J.R., **2022**. Traditional uses, phytochemistry and pharmacological activities of annonaceae. *Molecules*, 27(11), p.3462
- Alkhalidy, H., Al-Nabulsi, A., Mhawish, R. and Liu, D., **2023**. Low-dose of phenolic rich extract from *Annona squamosa* Linn leaves ameliorates insulin sensitivity and reduces body weight gain in HF diet-induced obesity. *Frontiers in Nutrition*, 10, p.1146021.
- Attiq, A., Jalil, J. and Husain, K., **2017**. Annonaceae: breaking the wall of inflammation. *Frontiers in pharmacology*, 8, p.752.
- Bermejo, A., Figadère, B., Zafra-Polo, M.C., Barrachina, I., Estornell, E. and Cortes, D., **2005**. Acetogenins from Annonaceae: recent progress in isolation, synthesis and mechanisms of action. *Natural product reports*, 22(2), pp.269-303.
- Biba, V.S., Amily, A., Sangeetha, S. and Remani, P., **2014**. Anticancer, antioxidant and antimicrobial activity of Annonaceae family. *World J. Pharm. Pharm. Sci*, 3(3), pp.1595-1604.
- Comuzzi, C., Di Bernardo, P., Portanova, R., Tolazzi, M. and Zanonato, P., **2002**. Affinity of lanthanide (III) ions for oxygen-and mixed oxygen–nitrogen-donor ligands in dimethylsulfoxide: a thermodynamic and spectroscopic investigation. *Polyhedron*, 21(14-15), pp.1385-1391.
- das Chagas Lima, N.N., Faustino, D.C., Allahdadi, K.J., de Aragão França, L.S. and Pinto, L.C., **2022**. Acetogenins from Annonaceae plants: Potent antitumor and neurotoxic compounds. *PharmaNutrition*, 20, p.100295.
- Dev, A.A. and Joseph, S.M., **2021**. Anticancer potential of *Annona* genus: A detailed review. *Journal of the Indian Chemical Society*, 98(12), p.100231.
- Elekofehinti, O.O., Lawal, A.O., Ejelonu, O.C., Molehin, O.R. and Famusiwa, C.D., **2020**. Involvement of fat mass and obesity gene (FTO) in the anti-obesity action of *Annona muricata* Annonaceae: in silico and in vivo studies. *Journal of Diabetes & Metabolic Disorders*, 19, pp.197-204.
- Erkens, R.H., Blanpain, L.M., Carrascosa Jara, I., Runge, K., Verspagen, N., Cosiaux, A. and Couvreur, T.L., **2023**. Spatial distribution of Annonaceae across biomes and anthromes: Knowledge gaps in spatial and ecological data. *Plants, People, Planet*, 5(4), pp.520-535.

- Frausin, G., Lima, R.B.S., Hidalgo, A.D.F., Maas, P. and Pohlit, A.M., **2014**. Plants of the Annonaceae traditionally used as antimalarials: a review. *Revista Brasileira de Fruticultura*, 36, pp.315-337.
- Guevara-Vásquez, A.M., Campos-Florián, J.V. and Dávila-Castillo, J.H., **2021**. *Annona muricata* L. extract decreases intestinal glucose absorption and improves glucose tolerance in normal and diabetic rats. *Journal of Herbmed Pharmacology*, 10(3), pp.359-366.
- Habib, M.R., Igarashi, Y. and Rabbi, M.A., **2022**. In vitro inhibitory activity of *Annona squamosa* leaves against enzymes associated with metabolic disorders. *Journal of Research in Pharmacy*, 26(5).
- Harizani, M., Diakaki, D.I., Perdikaris, S., Roussis, V. and Ioannou, E., **2022**. New C15 Acetogenins from Two Species of *Laurencia* from the Aegean Sea. *Molecules*, 27(6), p.1866.
- Jacobo-Herrera, N., Pérez-Plasencia, C., Castro-Torres, V.A., Martínez-Vázquez, M., González-Esquinca, A.R. and Zentella-Dehesa, A., **2019**. Selective acetogenins and their potential as anticancer agents. *Frontiers in Pharmacology*, 10, p.783.
- Kamada, T., Phan, C.S. and Vairappan, C.S., **2019**. Nangallenes A and B, halogenated nonterpenoid C15-acetogenins from the Bornean red alga *Laurencia nangii*. *Journal of Asian natural products research*, 21(3), pp.241-247.
- Kusmardiyani, S., Suharli, Y.A., Insanu, M. and Fidrianny, I., **2020**. Phytochemistry and pharmacological activities of *Annona* genus: a review. *Current Research on Biosciences and Biotechnology*, 2(1), pp.77-88.
- Laboureur, L., Bonneau, N., Champy, P., Brunelle, A. and Touboul, D., **2017**. Structural characterisation of acetogenins from *Annona muricata* by supercritical fluid chromatography coupled to high-resolution tandem mass spectrometry. *Phytochemical Analysis*, 28(6), pp.512-520.
- Lannuzel, A., Höglinger, G.U., Verhaeghe, S., Gire, L., Belson, S., Escobar-Khondiker, M., Poullain, P., Oertel, W.H., Hirsch, E.C., Dubois, B. and Ruberg, M., **2007**. Atypical parkinsonism in Guadeloupe: a common risk factor for two closely related phenotypes?. *Brain*, 130(3), pp.816-827.
- Laprévôte, O. and Das, B.C., **1994**. Structural elucidation of acetogenins from Annonaceae by fast atom bombardment mass spectrometry.

- Leal, F. and Paull, R.E., **2023**. The genus *Annona*: Botanical characteristics, horticultural requirements and uses. *Crop Science*, 63(3), pp.1030-1049.
- Le Ven, J., Schmitz-Afonso, I., Lewin, G., Lapr v te, O., Brunelle, A., Touboul, D. and Champy, P., **2012**. Comprehensive characterization of Annonaceous acetogenins within a complex extract by HPLC-ESI-LTQ-Orbitrap[ ] using post-column lithium infusion. *Journal of Mass Spectrometry*, 47(11), pp.1500-1509.
- Leite, D.O., de FA Nonato, C., Camilo, C.J., de Carvalho, N.K., da Nobrega, M.G., Pereira, R.C. and da Costa, J.G., **2020**. *Annona* genus: traditional uses, phytochemistry and biological activities. *Current Pharmaceutical Design*, 26(33), pp.4056-4091.
- Menezes, R.P., Sessions, Z., Muratov, E., Scotti, L. and Scotti, M.T., **2021**. Secondary metabolites extracted from Annonaceae and chemotaxonomy study of terpenoids. *Journal of the Brazilian Chemical Society*, 32, pp.2061-2070.
- Mohammed, R.N., Ramadhan, H.H. and Shari, F.H., **2021**. Hypoglycemic, Hypolipidemic, renal protective and antioxidant activity of *Annona muricata* in streptozotocin-induced diabetic rats. *Research Journal of Pharmacy and Technology*, 14(12), pp.6484-6490.
- Neske, A., Hidalgo, J.R., Cabedo, N. and Cortes, D., **2020**. Acetogenins from Annonaceae family. Their potential biological applications. *Phytochemistry*, 174, p.112332.
- Opara, P.O., Enemor, V.H.A., Eneh, F.U. and Emengaha, F.C., **2021**. Blood Glucose–Lowering Potentials of *Annona muricata* Leaf Extract in Alloxan–Induced Diabetic Rats. *European Journal of Biology and Biotechnology*, 2(2), pp.106-113.
- Philippus, A.C., Zatelli, G.A., Wanke, T., Barros, M.G.D.A., Kami, S.A., Lhullier, C., Armstrong, L., Sandjo, L.P. and Falkenberg, M., **2018**. Molecular networking prospection and characterization of terpenoids and C 15-acetogenins in Brazilian seaweed extracts. *RSC advances*, 8(52), pp.29654-29661.
- Poyer, S., Laboureur, L., Hebra, T., Elie, N., Van der Rest, G., Salpin, J.Y., Champy, P. and Touboul, D., **2022**. Dereplication of acetogenins from *annona muricata* by combining tandem mass spectrometry after lithium and copper postcolumn cationization and molecular networks. *Journal of the American Society for Mass Spectrometry*, 33(4), pp.627-634.
- Pulivarthi, V., Josthna, P. and Naidu, C.V., **2021**. Ameliorative effect of *Annona reticulata* L. leaf extract on antihyperglycemic activity and its hepato-renal protective potential in streptozotocin induced diabetic rats. *Journal of Ayurveda and Integrative Medicine*, 12(3), pp.415-426.

- Quílez, A.M., Fernández-Arche, M.A., García-Giménez, M.D. and De la Puerta, R., **2018**. Potential therapeutic applications of the genus *Annona*: Local and traditional uses and pharmacology. *Journal of ethnopharmacology*, 225, pp.244-270.
- Sasso, S., Sampaio e Souza, P.C., Santana, L.F., Cardoso, C.A.L., Alves, F.M., Portugal, L.C., de Faria, B.B., da Silva, A.F., Motta-Castro, A.R.C., Soares, L.S. and Bandeira, L.M., 2019. Use of an extract of *Annona muricata* Linn to prevent high-fat diet induced metabolic disorders in C57BL/6 mice. *Nutrients*, 11(7), p.1509.
- Scotti, L., Tavares, J.F., Silva, M.S.D., Falcão, E.V., Silva, L.D.M., Soares, G.C.D.S. and Scotti, M.T., **2012**. Chemotaxonomy of three genera of the annonaceae family using self-organizing maps and ¹³C NMR data of diterpenes. *Química Nova*, 35, pp.2146-2152.
- Singh, R., **2016**. Chemotaxonomy: a tool for plant classification. *J. Med. Plants Stud*, 4(2), pp.90-93.

APPENDIX F

The supplementary data referred to in PART 4 - Chapter 2 of this manuscript are given here.

APPENDIX	311
Figure S1: MS/MS spectrum of annonacin with copper sulphate (panel A) and lithium chloride (panel B) infusion.	312
Table S1: MS/MS fragments of annonacin with copper infusion (CID 40eV)	312
Table S2: MS/MS fragments of annonacin with lithium infusion (CID 80eV)	312
Figure S2: Isotopic pattern of annonacin cationized adducts.....	313
Figure S3: Protonated and cationized adducts of annonacin (<i>m/z</i> 550-900)	314
Figure S4: Comparison of MS/MS spectra of annonacin with europium salts (70eV)..	315
Figure S5: Comparison of MS/MS spectra of annonacin with cerium salts	316
Table S3: MS/MS fragments of goniiothalamycin with CAN infusion (CID 80eV)	316
Table S4: MS/MS fragments of annomutacin with CAN infusion (CID 80eV)	317
Table S5: MS/MS fragments of murisolin with CAN infusion (CID 80eV)	317
Table S6: Annonacinone detected fragments with CAN post-column infusion	318
Table S7: Annonacinone detected fragments with copper post-column infusion	318

Figure S1: MS/MS spectrum of annonacin with copper sulphate (panel A) and lithium chloride (panel B) infusion.

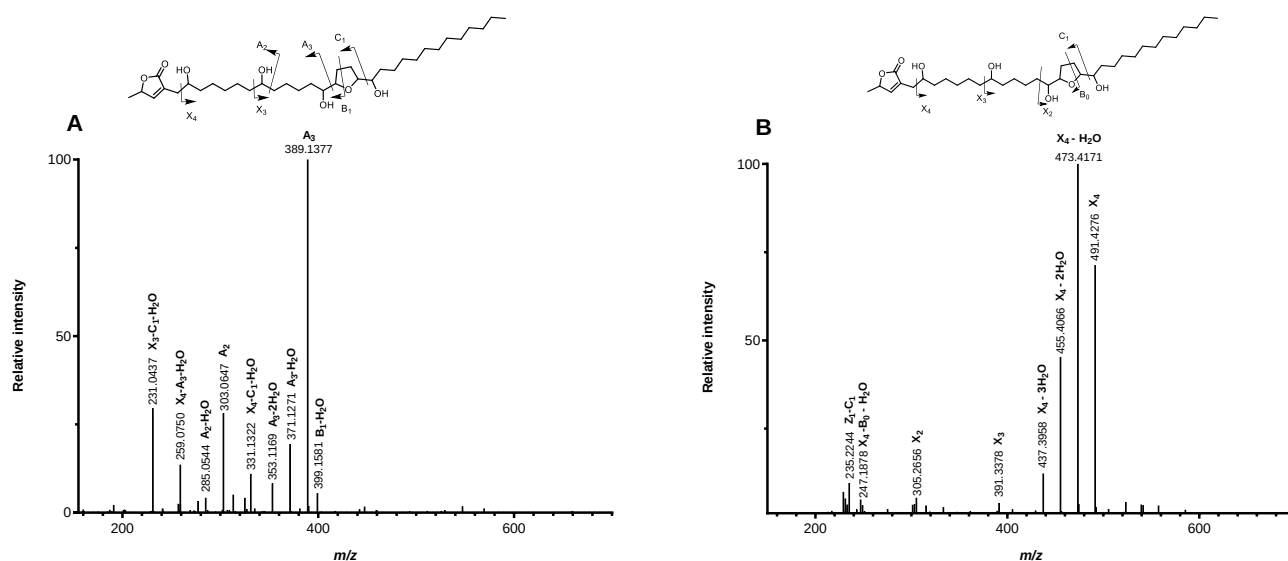
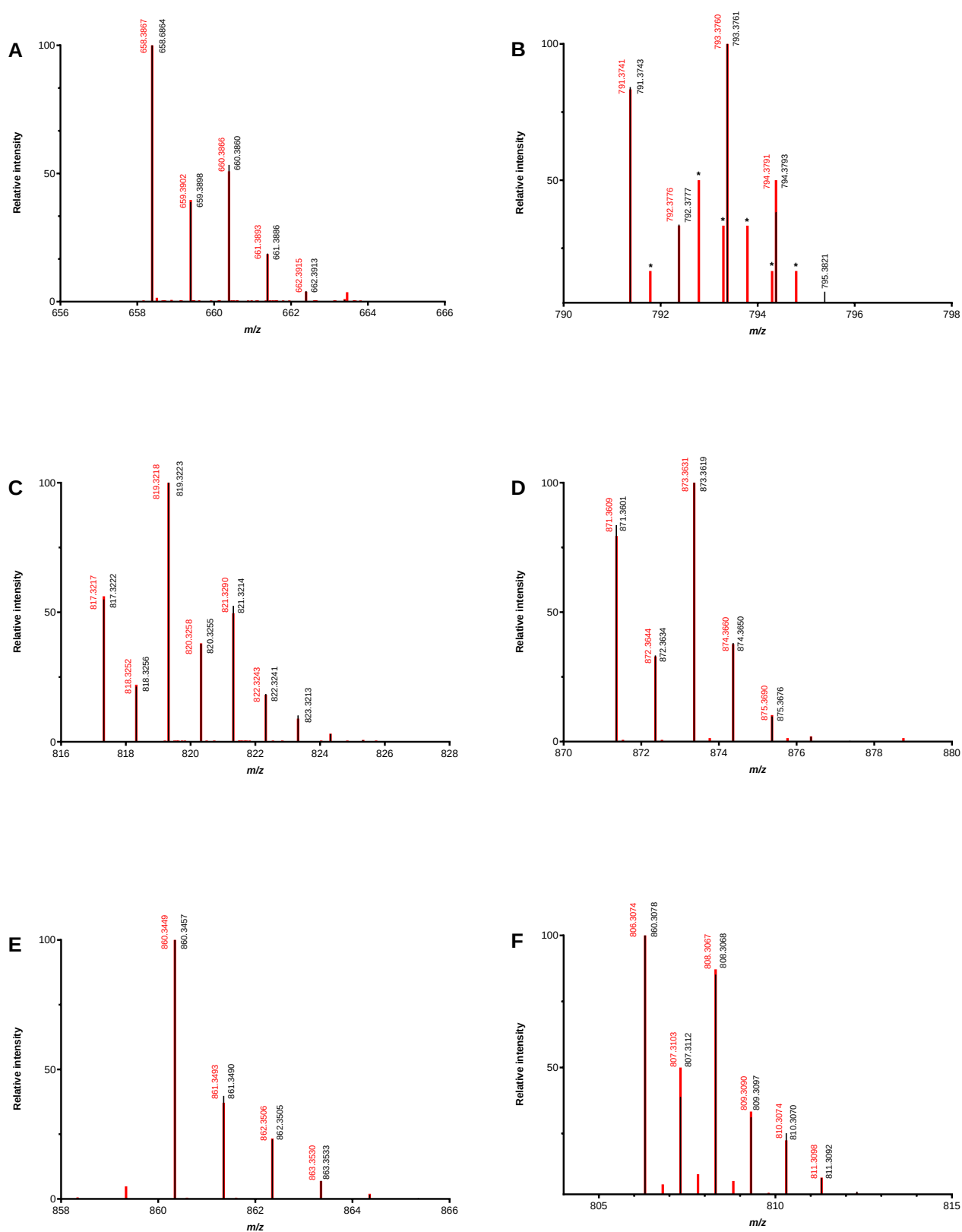


Table S1: MS/MS fragments of annonacin with copper infusion (CID 40eV)

Fragment ion (<i>m/z</i>)	Type of fragment	Chemical formula	Theoretical <i>m/z</i>	Δ ppm
399.1581	B ₁ - H ₂ O	C ₂₀ H ₃₂ O ₄ Cu ⁺	399.1591	-2.5
389.1377	A ₃	C ₁₈ H ₃₀ O ₅ Cu ⁺	389.1384	-1.8
371.1271	A ₃ - H ₂ O	C ₁₈ H ₂₈ O ₄ Cu ⁺	371.1278	-1.9
353.1165	A ₃ - 2H ₂ O	C ₁₈ H ₂₆ O ₃ Cu ⁺	353.1172	-2.0
331.1322	X ₄ -C ₁ -H ₂ O	C ₁₆ H ₂₈ O ₃ Cu ⁺	331.1329	-2.1
303.0647	A ₂	C ₁₃ H ₂₀ O ₄ Cu ⁺	303.0652	-1.6
285.0555	A ₂ -H ₂ O	C ₁₃ H ₁₈ O ₃ Cu ⁺	285.0546	3.1
259.0750	X ₄ - A ₃ - H ₂ O	C ₁₂ H ₂₀ O ₂ Cu ⁺	259.0754	-1.5
231.0437	X ₃ -C ₁ - H ₂ O	C ₁₀ H ₁₆ O ₂ Cu ⁺	231.0440	-1.3

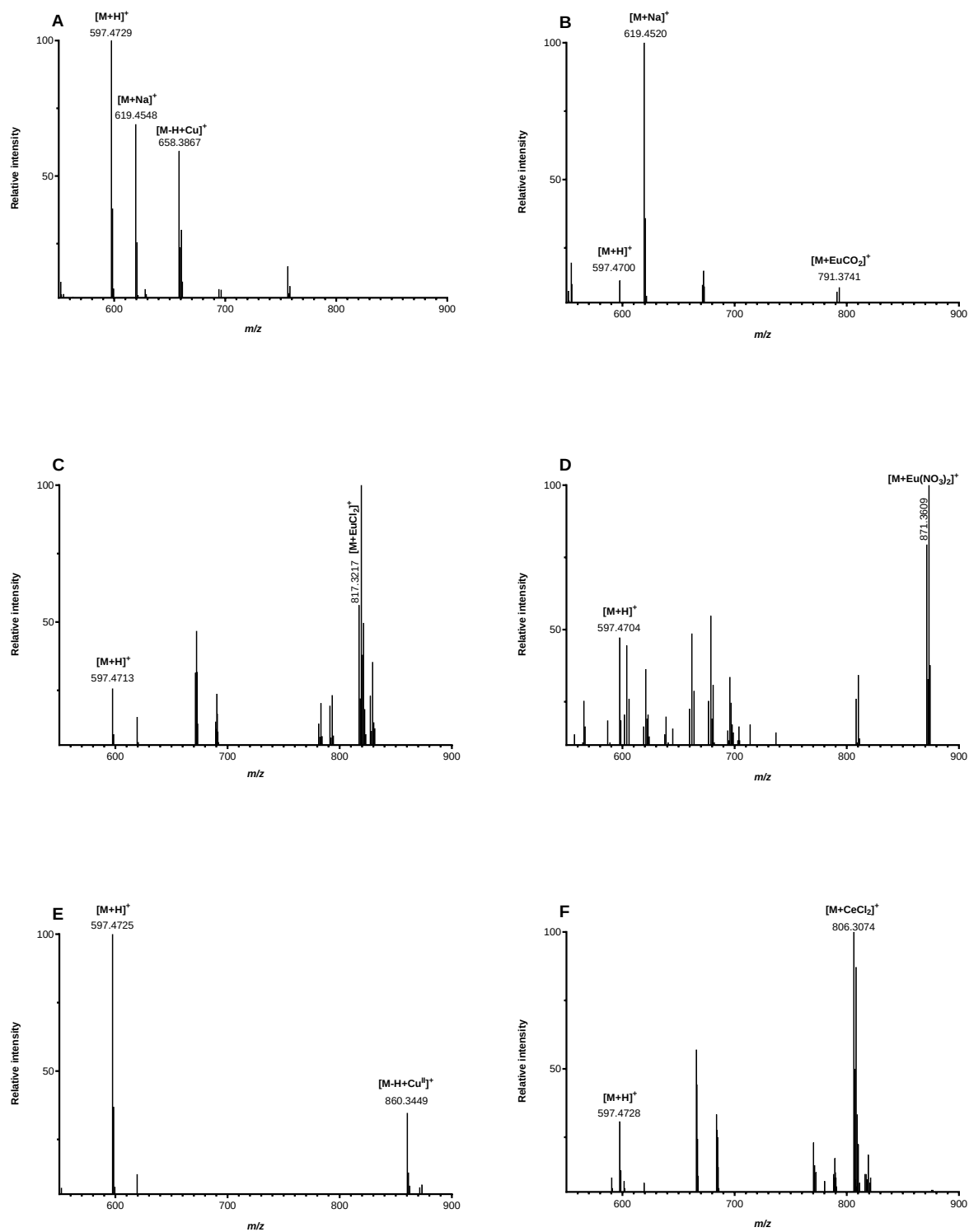
Table S2: MS/MS fragments of annonacin with lithium infusion (CID 80eV)

Fragment ion (<i>m/z</i>)	Type of fragment	Chemical formula	Theoretical <i>m/z</i>	Δ ppm
491.4276	X ₄	C ₂₉ H ₅₆ O ₅ Li ⁺	491.4283	-1.4
473.4171	X ₄ -H ₂ O	C ₂₉ H ₅₄ O ₄ Li ⁺	473.4177	-1.3
455.4066	X ₄ -2H ₂ O	C ₂₉ H ₅₂ O ₃ Li ⁺	455.4072	-1.3
437.3958	X ₄ -3H ₂ O	C ₂₉ H ₅₀ O ₂ Li ⁺	437.3967	-2.0
391.3378	X ₃	C ₂₃ H ₄₄ O ₄ Li ⁺	391.3395	-4.3
305.2656	X ₂	C ₁₈ H ₃₄ O ₃ Li ⁺	305.2663	-2.3
247.1878	X ₄ -B ₀ -H ₂ O	C ₁₄ H ₂₄ O ₃ Li ⁺	247.1880	-0.8
235.2244	Z ₁ -C ₁	C ₁₄ H ₂₈ O ₂ Li ⁺	235.2244	0.0

Figure S2: Isotopic pattern of annonacin cationized adducts

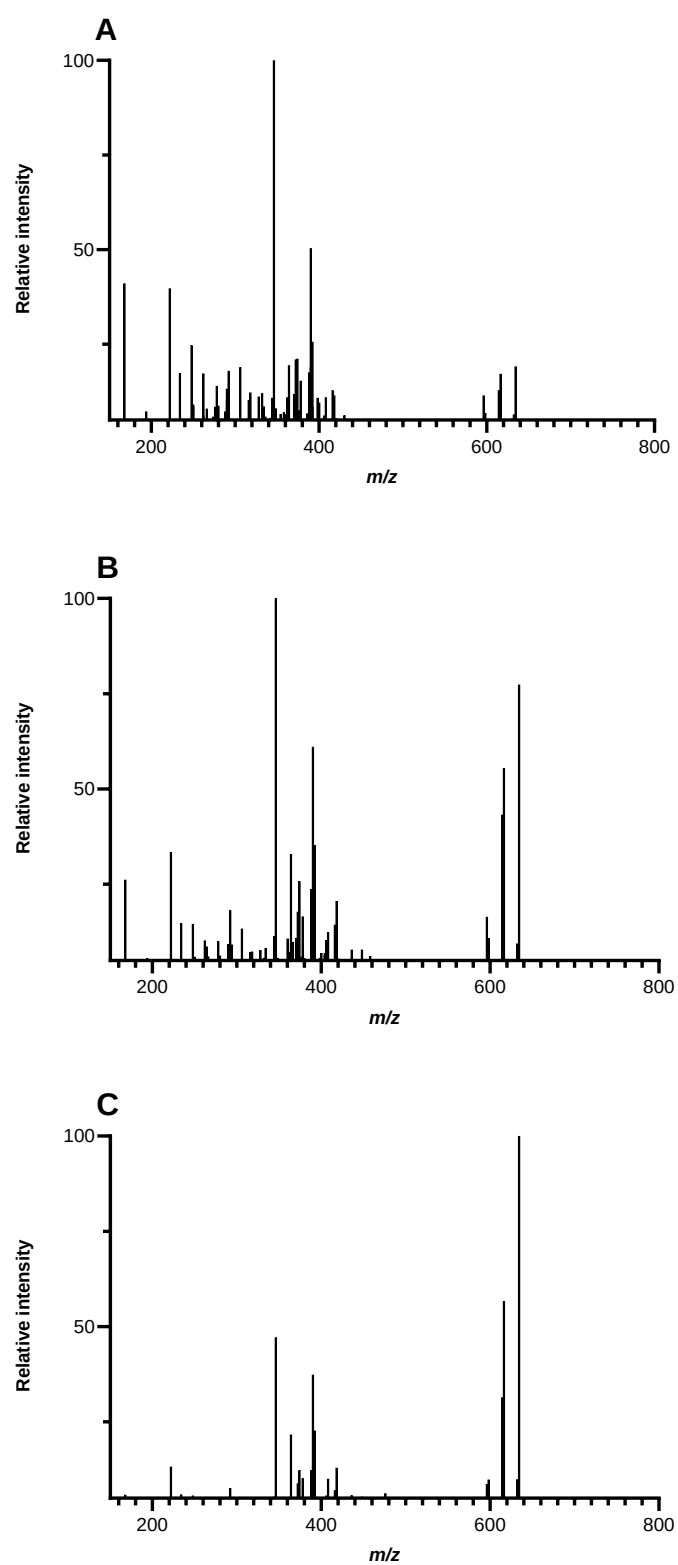
Panel A) [M-H+Cu^{II}]⁺ adduct; Panel B) [M+EuCO₂]⁺ adduct; Panel C) [M+EuCl₂]⁺ adduct; Panel D) [M+Eu(NO₃)₂]⁺ adduct; Panel E) [M+Ce(NO₃)₂]⁺ adduct; Panel F) [M+CeCl₂]⁺ adduct; in red: experimental data; in black: predicted data. (*) impurities.

Figure S3: Protonated and cationized adducts of annonacin (m/z 550-900)



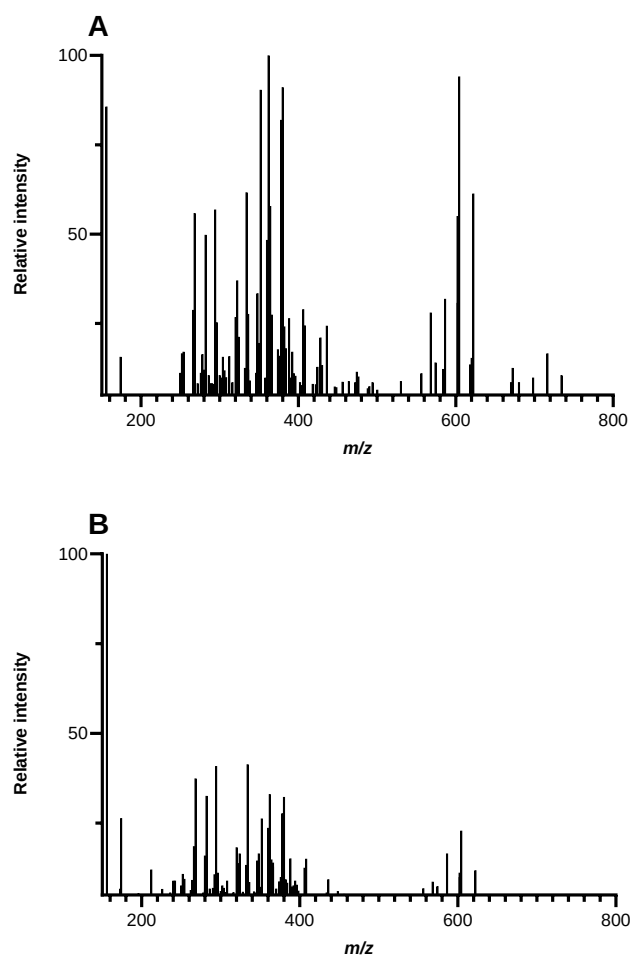
Panel A) $[M-H+Cu^{II}]^+$ adduct; Panel B) $[M+EuCO_2]^+$ adduct; Panel C) $[M+EuCl_2]^+$ adduct; Panel D) $[M+Eu(NO_3)_2]^+$ adduct; Panel E) $[M+Ce(NO_3)_2]^+$ adduct; Panel F) $[M+CeCl_2]^+$ adduct.

Figure S4: Comparison of MS/MS spectra of annonacin with europium salts (70eV)



Panel A) EuAc₃; panel B) EuCl₃; panel C) EuNO₃.

Figure S5: Comparison of MS/MS spectra of annonacin with cerium salts



Panel A) CAN; panel B) CeCl₃

Table S3: MS/MS fragments of goniotalamicin with CAN infusion (CID 80eV)

Fragment ion (<i>m/z</i>)	Type of fragment	Chemical formula	Theoretical <i>m/z</i>	Δ ppm
622.3030	X ₄	C ₂₉ H ₅₄ O ₅ Ce ⁺	622.3020	1.6
604.2915	X ₄ -H ₂ O	C ₂₉ H ₅₂ O ₄ Ce ⁺	604.2914	0.2
586.2790	X ₄ -2H ₂ O	C ₂₉ H ₅₀ O ₃ Ce ⁺	586.2809	-3.2
464.1700	X ₂	C ₂₀ H ₃₆ O ₃ Ce ⁺	464.1713	-2.8
408.1445	C ₂	C ₁₇ H ₃₂ O ₂ Ce ⁺	408.1451	-1.5
334.0345	X ₄ -B ₁	C ₁₂ H ₁₈ O ₂ Ce ⁺	334.0356	-3.3
294.0047	X ₂ -C ₂	C ₉ H ₁₄ O ₂ Ce ⁺	294.0043	1.4

Table S4: MS/MS fragments of annomutacin with CAN infusion (CID 80eV)

Fragment ion (<i>m/z</i>)	Type of fragment	Chemical formula	Theoretical <i>m/z</i>	Δ ppm
650.3330	X ₄	C ₃₁ H ₅₆ O ₅ Ce ⁺	650.3333	-0.5
632.3204	X ₄ -H ₂ O	C ₃₁ H ₅₆ O ₄ Ce ⁺	632.3228	-3.8
614.3112	X ₄ -2H ₂ O	C ₃₁ H ₅₄ O ₃ Ce ⁺	614.3122	-1.6
602.3104	Z ₀ -H ₂ O	C ₃₀ H ₅₄ O ₃ Ce ⁺	602.3122333	-3.0
502.1482	B ₁	C ₂₂ H ₃₄ O ₄ Ce ⁺	502.1506	-4.8
408.1061	X ₄ -B ₁	C ₁₆ H ₂₈ O ₃ Ce ⁺	408.1087	-6.4
380.0752	X ₄ -A ₃	C ₁₄ H ₂₄ O ₃ Ce ⁺	380.0774	-5.8
362.0647	X ₄ -A ₃ -H ₂ O	C ₁₄ H ₂₂ O ₂ Ce ⁺	362.0669	-6.1
350.0660	X ₄ -C ₂	C ₁₃ H ₂₂ O ₂ Ce ⁺	350.0669	-2.6
334.0358	X ₄ -C ₂ -CO	C ₁₂ H ₁₈ O ₂ Ce ⁺	334.0356	0.6
308.0190	B ₁ -A ₂	C ₁₀ H ₁₆ O ₂ Ce ⁺	308.0199	-2.9
267.9912	X ₄ -A ₂	C ₇ H ₁₂ O ₂ Ce ⁺	267.9886	9.7

Table S5: MS/MS fragments of murisolin with CAN infusion (CID 80eV)

Fragment ion (<i>m/z</i>)	Type of fragment	Chemical formula	Theoretical <i>m/z</i>	Δ ppm
606.3060	X ₄	C ₂₉ H ₅₄ O ₄ Ce ⁺	606.3071	-1.8
588.2953	X ₄ -H ₂ O	C ₂₉ H ₅₂ O ₃ Ce ⁺	588.2965	-2.0
558.2846	Z ₁ -H ₂ O	C ₂₈ H ₅₀ O ₂ Ce ⁺	558.2860	-2.5
476.1320	B ₁	C ₂₀ H ₃₂ O ₄ Ce ⁺	476.1350	-6.3
408.1063	X ₄ -C ₁	C ₁₆ H ₂₈ O ₃ Ce ⁺	408.1087	-5.9
390.0957	X ₄ -C ₁ -H ₂ O	C ₁₆ H ₂₆ O ₂ Ce ⁺	390.0981	-6.1
380.0749	X ₄ -B ₀	C ₁₄ H ₂₄ O ₃ Ce ⁺	380.0774	-6.6
364.0804	X ₄ -B ₁	C ₁₄ H ₂₄ O ₂ Ce ⁺	364.0825	-5.8
350.0647	Z ₁ -B ₀	C ₁₃ H ₂₂ O ₂ Ce ⁺	350.0669	-6.3
336.0497	X ₄ -A ₃	C ₁₂ H ₂₀ O ₂ Ce ⁺	336.0512	-4.5

Table S6: Annonacinone detected fragments with CAN post-column infusion

Fragment ion (<i>m/z</i>)	Type of fragment	Chemical formula	Theoretical <i>m/z</i>	Δ ppm
602.2747	X ₄ -H ₂ O	C ₂₉ H ₅₀ O ₄ Ce ⁺	602.2758	-1.8
584.2640	X ₄ -2H ₂ O	C ₂₉ H ₄₈ O ₃ Ce ⁺	584.2652	-2.0
488.0965	B ₀ -H ₂ O	C ₂₀ H ₂₈ O ₅ Ce ⁺	488.0986	-4.3
474.1892	Z ₂ -H ₂ O	C ₂₂ H ₃₈ O ₂ Ce ⁺	474.1921	-6.1
436.1371	X ₂	C ₁₈ H ₃₂ O ₃ Ce ⁺	436.1400	-6.9
408.1423	Z ₁	C ₁₇ H ₃₂ O ₂ Ce ⁺	408.1451	-6.9
394.0540	X ₄ -B ₀	C ₁₄ H ₂₂ O ₄ Ce ⁺	394.0573	-8.4
378.0597	X ₄ -B ₁	C ₁₄ H ₂₂ O ₃ Ce ⁺	378.0618	6.4
360.0497	X ₄ -B ₀ -2H ₂ O	C ₁₄ H ₂₀ O ₂ Ce ⁺	360.0512	-5.0
332.0193	X ₄ -A ₃ -H ₂ O	C ₁₂ H ₁₆ O ₂ Ce ⁺	332.0199	1.8

Table S7: Annonacinone detected fragments with copper post-column infusion

Fragment ion (<i>m/z</i>)	Type of fragment	Chemical formula	Theoretical <i>m/z</i>	Δ ppm
369.1099	A ₃ -H ₂ O	C ₁₈ H ₂₆ O ₄ Cu ⁺	369.1121	-6.0
351.0998	A ₃ -2H ₂ O	C ₁₈ H ₂₄ O ₃ Cu ⁺	351.1016	-5.1
341.1159	X ₄ -A ₄ -2H ₂ O	C ₁₇ H ₂₆ O ₃ Cu ⁺	341.1172	-3.8
329.1166	X ₄ -C ₁ -H ₂ O	C ₁₆ H ₂₆ O ₃ Cu ⁺	329.1172	-1.8
323.1061	X ₄ -A ₄ -3H ₂ O	C ₁₇ H ₂₄ O ₂ Cu ⁺	323.1067	-1.8
311.1070	X ₄ -C ₁ -2H ₂ O	C ₁₆ H ₂₄ O ₂ Cu ⁺	311.1068	0.6
285.0926	X ₄ -B ₁ -H ₂ O	C ₁₄ H ₂₂ O ₂ Cu ⁺	285.0910	5.6
275.0720	X ₄ -A ₃	C ₁₂ H ₂₀ O ₃ Cu ⁺	275.0708	4.3

General Conclusion

My PhD research focused on the phytochemical characterisation of less studied plants with anti-diabetic potential. Therefore, I worked on the extraction and isolation of secondary metabolites using chromatographic techniques (MPLC and HPLC) and applied spectroscopic and spectrometric methods (1D and 2D NMR, HR-MS) to characterise the compounds. The pure metabolites were then tested against targets involved in Metabolic Syndrome in order to identify the bioactive components.

A major part of my project involved the phytochemical characterisation of three different *Centaurea* species, including *C. kotschyi*, *C. sicula* and *C. patula*. This led to the isolation and structural characterisation of 58 pure secondary metabolites, including sesquiterpene lactones, lignans and polyphenols. In particular, these analyses led to the isolation of 14 unprecedented compounds (12 sesquiterpenes, one neolignane and one novel alkaloid). In addition, the antidiabetic potential of the isolated secondary metabolites was evaluated on GLUT transporters and on the α -glucosidase enzyme. The results of the studies on *C. kotschyi* and *C. sicula* were published in *Phytochemistry* in 2022 and in the *Journal of Natural Products* in 2024, while the results shown for *C. patula* are about to be submitted. Our analysis of the secondary metabolites of *Centaurea* species led to the identification of the bioactive components responsible for the hypoglycaemic effect of the extracts, revealing a promising activity of several compounds belonging to different structural families, some of which exploit a multi-target mechanism. Although further analysis is needed, it seems that these plants could be used as dietary supplements, taking advantage of the mixture of active ingredients that modulate blood glucose levels.

I then focused on the phytochemical composition of the aerial parts of *Glycyrrhiza foetida*, from which a total of 29 pure compounds were isolated and characterised, including 11 previously unreported amorfrutins and an unprecedented prenylated flavanone. All compounds were tested as GLUT promoters and mitochondrial stimulators, while the amorfrutins in particular were tested against PPARs targets. These results were published in the *Journal of Natural Products* in 2023 and in *Biomolecules* in 2024. *G. foetida* has emerged as a rich source of

bioactive compounds, thus not only can be considered as a potential functional food, but amorfrutins could be new leads for the development of new selective PPARs modulators. Therefore, amorfrutins are potential drug candidates for the treatment of diabetes and related diseases.

During the second year of my PhD, I have also contributed to the phytochemical analysis of Aglianico grape pomace, conducting NMR and untargeted LC-MS² experiments. This research was then focused on the anti-diabetic potential of oleanolic acid, one of the main components of the extract. This study was published in the *Journal of Functional Foods* in 2023.

During the last year of my PhD, I worked on the analytical analysis of the Malagasy *Piper borbonense*, a refined species of genus *Piper*, revealing the main differences in aroma composition between this wild species and the domesticated *P. nigrum*. Moreover, the untargeted LC-MS² analysis guided the subsequent purification steps, which led to the isolation of 38 compounds, including two unprecedented metabolites, and highlighted the surprising abundance of sesamin when compared to that present in *P. nigrum*. Given the potential of piperamides in metabolic syndrome, we look forward to testing our library of compounds against selected targets.

Finally, during my stay abroad, I worked on the dietary plants of the genus *Annona*, carrying out a chemotaxonomic analysis and optimising an LC-MS² method for the detection of acetogenins, neurotoxins present in the edible fruits.

Overall, plants remain a valuable source of bioactive compounds, many of which are still undiscovered, that can help manage global diseases. The use of dietary supplements and functional foods is now part of modern medicine, and their economic impact on the pharmaceutical industry is particularly relevant. Nutraceuticals are subject to less restrictive legislation than drugs, and while this makes them easy to introduce to the market, their indiscriminate use can often cause problems. It is therefore essential to carry out in-depth research into those plants that can actually have a beneficial effect on human health and then to use them in the most appropriate way, whether to support pharmacological therapies, as food supplements or as alternative medicine. This research can not only valorise the traditional use of the plants studied and have a positive economic impact in the regions where they are native, but can also lead to the discovery of new potential drugs, as in the case of amorfrutins.