Faculty of Pharmaceutical, Biomedical and Veterinary Sciences
Department of Pharmaceutical Sciences

Analytical, Antiviral and Chemical Investigations of Selected Classes of Alkaloids

Analytisch, Antiviraal en Chemisch Onderzoek op Geselecteerde Klassen van Alkaloïden

Proefschrift voorgelegd tot het behalen van de graad van Doctor in de Farmaceutische Wetenschappen aan de Universiteit Antwerpen te verdedigen door

Ngoc Thao Hien LE

Promotors

Prof. Dr. L. Pieters

Prof. Dr. E. Tuenter

Antwerp, 2023

“Learn the rules, break the rules, make up new rules, break the new rules.”

- Marvin Bell -

“Đôi khi cuộc sống thật đơn giản

Khi ta biết ta có một đam mê

Để ta có những đêm về thanh thản

Biết rằng mình đã sống trọn đam mê”

- Lê Ngọc Thảo Hiền –

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

Abbreviation	Explanation
CASE	Computer-assisted structure elucidation
COSY	Correlation spectroscopy
COVID-19	Coronavirus disease - 2019
DFT	Density functional theory
ECD	Electronic circular dichroism
GNPS	Global Natural Products Social Molecular Networking
HMBC	Heteronuclear multiple-bond correlation
HPLC	High-performance liquid chromatography
HSQC	Heteronuclear single-quantum coherence
MS	Mass spectrometry
NOESY	Nuclear Overhauser effect spectroscopy
NMR	Nuclear magnetic resonance
OR	Optical rotation
ORD	Optical rotation dispersion
RCF	Relative centrifugal force
rpm	Revolutions per minute
SARS-CoV-2	Severe acute respiratory syndrome coronavirus-2
SPE	Solid-phase extraction
TLC	Thin-layer chromatography
QToF	Quadrupole time-of-flight
TQD	Triple quadrupole
UPLC	Ultra-performance liquid chromatography
VCD	Vibrational circular dichroism

CHAPTER 1

General introduction

1.1. Alkaloids

'Alkaloid' is a term referring to low-molecular weight nitrogen-containing compounds, in which the nitrogen atom is usually derived from an amino acid. They are produced by microorganisms, plants and animals. In phytochemistry, alkaloids are the most diverse group of secondary metabolites with a wide variety of skeletons being reported in literature [1]. In general, alkaloids have basic properties owing to the presence of at least one nitrogen atom; however, certain alkaloids do not possess basic properties (e.g. colchicine, caffeine), or even have weak acidic properties, e.g. theophylline. Many plants with a high alkaloid content have been used in traditional folk medicine [2], including Chinese and Ayurvedic medicine and a large number of publications report the significant pharmacological activity of alkaloids. Therefore, research on alkaloids is still a compelling subject.

Amaryllidaceae alkaloids (AAs) are compounds belonging to the large group of isoquinoline alkaloids, derived from L-tyrosine and L-phenylalanine, and they are almost exclusively found in the Amaryllidaceae family [3]. These alkaloids show a wide variety of pharmacological activities such as anticancer, antiplasmodial, anti-inflammatory, etc [3]. In the pharmaceutical domain, galanthamine is being marketed worldwide for the treatment of cognitive decline in mild to moderate Alzheimer's disease and various other memory impairments [4]. With regard to their biosynthesis, AAs are considered to be formed biogenetically by intramolecular oxidative coupling of norbelladine [5]. To date, more than 600 compounds have been isolated and were described in the Amaryllidaceae family, and the number is still increasing [5]. **Figure 1** demonstrates the common skeleton-types of AAs [3-5]. Among these, the lycorine-type represents the most ubiquitous and chemically diverse scaffold.

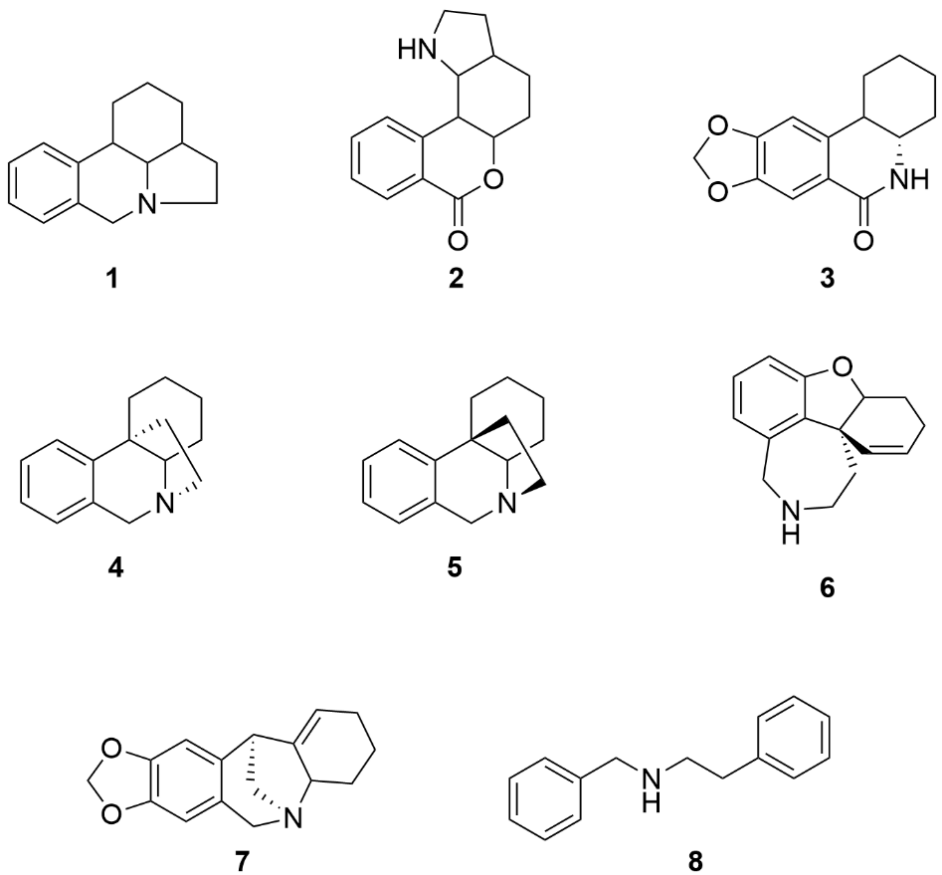


Figure 1. Common skeletal subclasses of Amaryllidaceae alkaloids: lycorine-type (1), lycorenine-type (2), narciclasine-type (3), haemanthamine-type (4), crinine-type (5), galanthamine-type (6), montanine-type (7) and norbelladine-type (8)

In this project, the phytochemical profiles of three selected Amaryllidaceae plant species were re-explored with the aim of building a library for testing antiviral activities since AAs are well-known as bioactive natural products [3]. The current investigation contributes to the fundamental chemical knowledge about this class of compounds in nature, which can be useful for future researchers in this field looking for new lead compounds or new therapeutic agents, and for elucidation of biosynthetic pathways aiming at unlocking and harnessing nature for human benefits.

The taxonomy of the plant species is summarized in **Figure 2** [6].

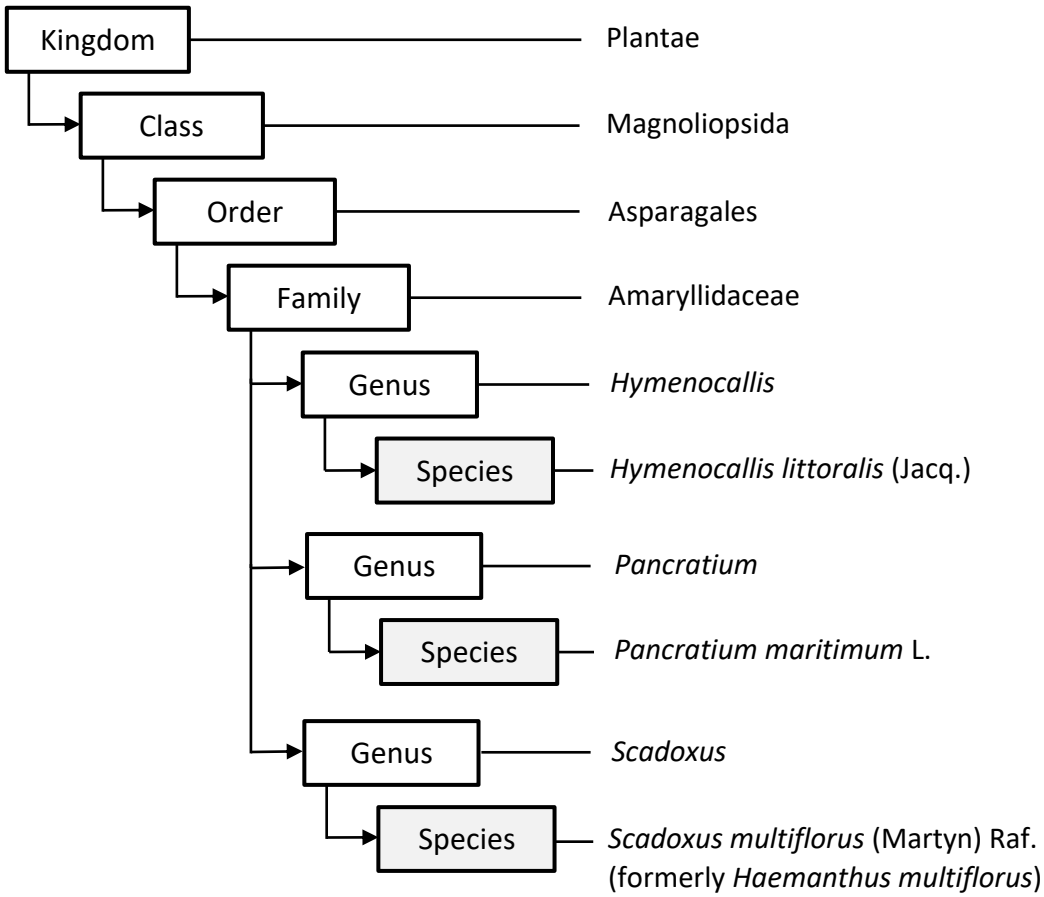


Figure 2. Taxonomic hierarchy of the selected Amaryllidaceae plant species

1.2. Antiviral natural products

Viral infections in humans caused millions of deaths around the globe and are accountable for many human diseases like HIV/AIDS, hepatitis, influenza, herpes, common cold, etc [7, 8]. Most recently, the COVID-19 pandemic has threatened mankind since 2019. Up to now, the pandemic has caused an immense global health crisis, with more than 6.5 million deaths according to the worldometers website [7]. SARS-CoV-2 (severe acute respiratory syndrome-related coronavirus-2), responsible for COVID-19 (coronavirus disease-2019), shares 79% sequence similarity with SARS-CoV, belonging to the genus Sarbecovirus (RNA viruses) [8]. SARS-CoV-2 is transmitted through respiratory droplets and aerosols. Although in

some cases the infection is asymptomatic, most patients suffer from mild-to-moderate respiratory disease, experiencing cough, fatigue, fever, headache and diarrhea, but also severe and life-threatening lung damage can develop [8]. Since the SARS-CoV-2 outbreak commenced in December 2019, medicinal plants as a complementary and/or alternative medicine to fight against SARS-CoV-2 have remained an intriguing research topic and a common practice at many places around the world, alongside the application of vaccines and synthetic drugs [9-11]. In addition, the pandemic has given a boost for testing plant extracts and libraries of natural compounds for potential activity against the virus.

Since ancient times, medicinal plants have been widely used for the treatment of many infectious and non-infectious diseases including viral infections [12, 13]. Viruses have been more resistant to treatment or prevention than any other form of life due to their constantly mutating ability. Currently, compared to antibacterial antibiotics, there are still relatively few drugs available on the market to treat viral diseases, like acyclovir, which is modelled based on a natural product [13].

Natural products as antiviral substances have been a subject of study for decades. Extensive reviews have demonstrated that several plant families could offer a rich reserve for drug discovery of antiviral natural products, such as Acanthaceae, Amaranthaceae, Amaryllidaceae, Apocynaceae, Asphodelaceae, Combretaceae, Lamiaceae, Fabaceae, Rutaceae, Piperaceae, etc. [14, 15]. Among these, weak-to-potent antiviral activities were reported for many Amaryllidaceae alkaloids against viruses in the Arenaviridae, Retroviridae, Piconaviridae, Togaviridae, Flaviviridae, Phenuiviridae, Paramyxoviridae, Coronaviridae, Rhabdoviridae, Filoviridae, Virgaviridae, Orthomyxoviridae, Herpesviridae and Poxviridae families [15, 16]. In the course of 2020-2022, lycorine, a main and widespread alkaloid in the Amaryllidaceae family, was reported as a highly potent anti-SARS-CoV-2 agent with an EC_{50} ranging from 0.18 - 0.87 µM [17, 18]. This has drawn scientific attention to Amaryllidaceae alkaloids as potential natural products for the treatment of SARS-CoV-2 infections. Thus, it was of our interest to investigate the anti-SARS-CoV-2 potency for a library of Amaryllidaceae alkaloids.

1.3. Quality control of plant-based food supplements

By definition, plant-containing commercial products can be classified into the following four categories according to European Union's regulations: (1) - Herbal substances or botanicals, (2) - Herbal or botanical preparations, (3) - Botanical or Plant Food Supplements and (4) - Herbal medicinal products [19]. Each classification requires different requirements for product marketing (safety, production, labelling and product composition). In this project, the focus is on Plant Food Supplements. According to European Food Safety Authority (EFSA), Food Supplements are legally defined as "*foodstuffs, the purpose of which is to supplement the normal diet and which are concentrated sources of nutrients or other substances with a nutritional or physiological effect, alone or in combination, marketed in dose form*" [20].

In Belgium, the two primary ways for introducing a medicinal plant to the market are: the registration process as a medicine (herbal medicinal product) or the notification process as a food supplement. Apart from that, plant preparations can also be marketed as medical device, cosmetic or homeopathic medicine. To register a new medicine, an application must be submitted to the Federal Agency for Medicines and Health Products, and for herbal medicines, it must be approved by the Committee on Herbal Medicinal Products, which comprises a panel of experts. When dealing with novel synthetic compounds, an application for marketing authorization necessitates a comprehensive report on efficacy, safety, pharmacokinetic properties, and other relevant factors. Conversely, for herbal preparations, a "simplified procedure" can be followed due to the exceedingly complex chemical composition of plant extracts and the limited knowledge of all active compounds present. In this application, safety and efficacy can be demonstrated through scientific literature, although detailed analytical data must be included to ensure the quality of the results [20, 21].

However, the "simplified procedure" above-mentioned is still demanding, and manufacturers often prefer to market herbal preparations as a food supplement rather than a herbal medicinal product. In Belgium, the legislation for food supplements containing plant extracts is described in the Belgian Royal Decree 31.08.2021 (published October 4th, 2021) which is an update to the Royal Decree

KB 29.08.1997. The Decree was published with three updated lists of plant species. Among these lists, list 3 contains an extensive series of plant species that require a notification before they can be marketed as a pre-dosed food supplement. For many of the plants in list 3, additional analytical documentation is required [21].

1.4. Aim of this work

The current project focuses on the phytochemical, biological and analytical investigation of selected classes of alkaloids. The thesis contains two distinct parts; the first one dealing with Amaryllidaceae alkaloids as potential antiviral agents, more in particular against SARS-CoV-2; and the second one dealing with quality control of alkaloid-containing food supplements based on Maca (*Lepidium meyenii*). Two main work packages were designed:

1. Phytochemical and antiviral investigations of Amaryllidaceae alkaloids. Firstly, a library of alkaloids will be composed from three selected plant species (*Hymenocallis littoralis*, *Pancratium maritimum*, *Scadoxus multiflorus*), because they are known to contain a wide variety of alkaloids belonging to this class. Secondly, *in vitro* screening of cytotoxicity and anti-SARS-CoV-2 activity will be performed for the library. Lastly, preliminary structure-activity relationships will be examined based on the obtained anti-SARS-CoV-2 results.
2. Phytochemical and analytical investigations of alkaloids from *Lepidium meyenii* Walp. (Maca). Firstly, the alkaloidal profile of Maca will be investigated. Isolated compounds will be used as standards for developing an analytical method for quality control of Maca-containing food supplements in the next step. The analytical work herein is part of the ANAHERBAFOOD (Analysis of Herbal Food Supplements) project, which aims to develop and validate analytical methods to detect or quantify a specific compound or class of compounds in plants mentioned in the Belgian Royal Decree. According to the list 3 of the Decree, Maca roots are allowed in food supplements, but analyses should confirm that there are no detectable quantities of potentially toxic alkaloids [21], [22].

Based on the obtained scientific results, the thesis hereby is outlined as follows:

1. Chapter 1 - General introduction to the PhD project. In this chapter, scientific reasoning behind the project and an overview of the studied subjects are provided. Detailed literature research for each work package can be found as Introduction section in chapters 3 – 8.
2. Chapter 2 - General materials and instrumentation. Information about solvents, reagents and instrumentation used for all experiments are listed in this chapter.
3. Chapter 3 - General workflow for structure elucidation. Since all investigations started with a phytochemical study, structure elucidation of isolated compounds is always the first task. This chapter describes a general workflow for structure elucidation, applied as example to the analysis of an epimeric mixture.
4. Chapter 4 - Amaryllidaceae alkaloids from *Hymenocallis littoralis*. In this chapter, the phytochemical profile and anti-SARS-CoV-2 activity of alkaloids isolated from *H. littoralis* was described.
5. Chapter 5 - Amaryllidaceae alkaloids from *Pancratium maritimum*. In this chapter, the phytochemical profile and anti-SARS-CoV-2 activity of alkaloids isolated from *P. maritimum* was reported.
6. Chapter 6 - Amaryllidaceae alkaloids from *Scadoxus multiflorus*. In this chapter, the phytochemical profile and anti-SARS-CoV-2 activity of alkaloids isolated from *S. multiflorus* was described.
7. Chapter 7 – Alkaloids from *Lepidium meyenii* (Maca). In this chapter, the phytochemical profile of alkaloids isolated from Maca was reported.
8. Chapter 8 – Quality control of food supplements containing Maca. This chapter is a continuation of chapter 7, describing a UPLC-TQD-MS/MS method for quality control of Maca-containing food supplements.
9. Chapter 9 – General discussion, conclusion and future perspectives

1.5. References

[1] Tadeusz Aniszewski, *Alkaloids: Chemistry, Biology, Ecology, and Applications*, Second. Helsinki: Elsevier, 2015.

[2] Biren N. Shah, *Textbook of Pharmacognosy and Phytochemistry*. Elsevier, 2010.

[3] Y. Ding, D. Qu, K. Zhang, X. Cang, Z. Kou, and W. Xiao, "Phytochemical and biological investigations of Amaryllidaceae alkaloids: a review," *J. Asian Nat. Prod. Res.*, vol. 6020, pp. 1–48, 2017.

[4] L. Marco and C. Carreiras, "Galanthamine , a Natural Product for the Treatment of Alzheimer's Disease," pp. 105–111, 2006.

[5] Berkov, S.; Osorio, E.; Viladomat, F.; Bastida, J. Chemodiversity, Chemotaxonomy and Chemoecology of Amaryllidaceae Alkaloids, 1st ed.; Elsevier Inc.: Amsterdam, The Netherlands, 2020; Volume 83.

[6] https://wfoplantlist.org/plant-list/ (checked 11/07/2023)

[7] https://www.worldometers.info/coronavirus/ (checked 11/07/2023)

[8] Lamers, M. M., & Haagmans, B. L. SARS-CoV-2 pathogenesis. *Nature reviews microbiology*, *20*(5), 270-284, 2022.

[9] Raman, K.; Rajagopal, K.; Islam, F.; Dhawan, M.; Mitra, S.; Aparna, B.; Varakumar, P.; Byran, G.; Choudhary, O.P.; Emran, T. Bin. Role of Natural Products towards the SARS-CoV-2: A Critical Review. *Ann. Med. Surg. 80*, 104062, 2022.

[10] Christy, M.P.; Uekusa, Y.; Gerwick, L.; Gerwick, W.H. Natural Products with Potential to Treat RNA Virus Pathogens Including SARS-CoV-2. *J. Nat. Prod. 84*, 161–182, 2021.

[11] Chakravarti, R.; Singh, R.; Ghosh, A.; Dey, D.; Sharma, P.; Velayutham, R.; Roy, S.; Ghosh, D. A Review on Potential of Natural Products in the Management of COVID-19. *RSC Adv. 11*, 16711–16735, 2021.

[12] El Sayed, Khalid A. "Natural products as antiviral agents." Studies in natural products chemistry 24: 473-572, 2000.

[13] Bachar, Sitesh C., et al. "A review of medicinal plants with antiviral activity available in Bangladesh and mechanistic insight into their bioactive metabolites on SARS-CoV-2, HIV and HBV." *Frontiers in Pharmacology* 12: 732891, 2021.

[14] Mukhtar, Muhammad, et al. "Antiviral potentials of medicinal plants." Virus research 131.2: 111-120, 2008.

[15] Nair, J.J.; van Staden, J. Antiviral Alkaloid Principles of the Plant Family Amaryllidaceae. *Phytomedicine*, *108*, 154480, 2022.

[16] Majnooni, M.B.; Fakhri, S.; Bahrami, G.; Naseri, M.; Farzaei, M.H.; Echeverría, J. Alkaloids as Potential Phytochemicals against SARS-CoV-2: Approaches to the Associated Pivotal Mechanisms. *Evid. Based Complement. Altern. Med.* 6632623, 2021.
[17] Zhang YN, Zhang QY, Li XD, Xiong J, Xiao SQ, Wang Z, Zhang ZR, Deng CL, Yang X Lou, Wei HP, Yuan ZM, Ye HQ, Zhang B. Gemcitabine, lycorine and oxysophoridine inhibit novel coronavirus (SARS-CoV-2) in cell culture. Emerg. Microbes. Infect. 9: 1170–1173, 2020.
[18] Jin, Y.H., Min, J.S., Jeon, S., Lee, J., Kim, S., Park, T., Park, D., Jang, M.S., Park, C.M., Song, J.H., Kim, H.R., Kwon, S.. Lycorine, a non-nucleoside RNA dependent RNA polymerase inhibitor, as potential treatment for emerging coronavirus infections. Phytomedicine 86, 153440, 2021.
[19] Silano, V., Coppens, P., Larrañaga-Guetaria, A., Minghetti, P., & Roth-Ehrang, R. Regulations applicable to plant food supplements and related products in the European Union. *Food & function*, *2*(12), 710-719, 2011.
[20] EFSA - Food Supplements. Available from: https://www.efsa.europa.eu/en/topics/topic/food-supplements.
[21] Koninklijk besluit tot wijziging van het koninklijk besluit van 29 augustus 1997 betreffende de fabricage van en de handel in voedingsmiddelen die uit planten of uit plantenbereidingen samengesteld zijn of deze bevatten, (31/08/2021).
[22] Pieters L, Hermans N. ANAHERBAFOOD: Development and validation of analytical methods for quality control of herbal food supplements, 2018.

CHAPTER 2

General Experimental Procedures

2.1. General Materials

All analytical-grade solvents, including ethanol (EtOH), dichloromethane (DCM), ethyl acetate (EtOAc), isopropanol (IPA), *n*-butanol (*n*-BuOH) were purchased from Acros Organics (Geel, Belgium) or from Fisher Scientific (Loughborough, UK). Solvents used for HPLC and UPLC, i.e. methanol (MeOH), acetonitrile (ACN) and isopropanol were HPLC and UPLC grade and were purchased from Fisher Scientific (Loughborough, UK) and Biosolve BV (Valkenswaard, the Netherlands), respectively. Water was dispensed and filtered by a Direct-Pure Up Ultrapure and RO water system (Rephile Bioscience, Belgium). For NMR experiments, methanol-d_4 (CD_3OD - 99.8% D), chloroform-*d* ($CDCl_3$ - 99.8% D) and dimethyl sulfoxide-d_6 (99.9% D) were purchased from Sigma-Aldrich (Merck, Germany).
Reagents, including TLC spraying reagents: ammonia 25% (NH_4OH 25%), formic acid (FA 98%), MS-grade formic acid (FA 99%), acetic acid (AA 98%), hydrochloric acid 25% (HCl 25%), potassium iodide (KI 99%), ammonium acetate (NH_4Ac 98%), diethylamine (DEA 99%) and bismuth (III) nitrate ($Bi(NO_3)_3$ 99.5%) were purchased from either Acros Organics (Geel, Belgium), Sigma-Aldrich (St. Louis, MO, USA) or Merck (Darmstadt, Germany).

2.2. Chromatographic and Spectroscopic Instrumentation

2.2.1. Thin Layer Chromatography (TLC)

TLC was performed on pre-coated silica gel F254 plates (Merck, Darmstadt, Germany), and the bands were observed under UV light (254 and 366 nm), as well as under visible light after spraying with the Dragendorff reagent. Dragendorff reagent was prepared by combining a mixture A and a solution B. Mixture A was prepared by suspending 0.85 g of bismuth subnitrate ($Bi_5O(OH)_9(NO_3)_4$) in 40 mL of water and 10 mL of glacial acetic acid. Solution B was prepared by dissolving 8 g of potassium iodide (KI) in 20 mL of water.

2.2.2. Flash Chromatography

Flash chromatography was performed on a Grace Reveleris X2 system (Columbia, MD, USA) using the Reveleris Navigator™ software. The system was equipped with a binary pump, a UV detector, an evaporative light scattering detector and a fraction collector. Isopropanol was supplied to the ELSD detector during operation, while DAD detection was set at 210 and 254 nm.

2.2.3. High-Performance Liquid Chromatography (HPLC)

HPLC analysis was performed on an Agilent 1200 series HPLC-DAD system (Agilent Technologies, Santa Clara, USA) with OpenLAB software version A.01.05. A Phenomenex Kinetex EVO C18 column (250 x 4.6 mm, 5 µm) was used for analytical purposes.

2.2.4. High-Performance Liquid Chromatography – Solid-phase extraction – Nuclear Magnetic Resonance Spectroscopy (HPLC-SPE-NMR)

For analytical and isolation purposes, an HPLC-PDA-SPE-NMR system was used, comprising an Agilent 1200 series HPLC system (Agilent Technologies, Santa Clara, CA, USA) with an in-line solvent degasser, quaternary pump, autosampler, column compartment, and a photodiode-array detector. The eluate was diluted with water (flow rate of 3 mL/min) by means of a K-120 HPLC pump (Knauer, Berlin, Germany). Samples were collected on 2 mm GP Resin cartridges using a Prospekt-2 SPE-unit (Spark Holland, Emmen, The Netherlands). Elution of samples from cartridges to NMR tubes was directed by a Gilson Liquid Handler 215 (Gilson, Middleton, WI, USA). The system was operated through Hystar version 3.2 software (Bruker, Daltonik) for HPLC analysis and trapping, and Prep Gilson ST version 1.2 software (Bruker, Biospin) for drying and elution. A Phenomenex Kinetex EVO C18 (250 x 4.6 mm, 5 µm) column was selected for all analyses (Phenomenex, Utrecht, the Netherlands).

2.2.5. Semi-preparative High-Performance Liquid Chromatography – Mass spectrometry (Semi-prep HPLC-MS)

For purification, a preparative HPLC-DAD-MS system was used. All its compartments were supplied by Waters (Milford, MA, USA), including MassLynx v.4.1 software. A Phenomenex Kinetex EVO C18 column (250 x 10.0 mm, 5 µm) was

selected for the semi-preparative separation (Phenomenex, Utrecht, the Netherlands).

2.2.6. Ultra-Performance Liquid Chromatography – Quadruple Time of Flight – Tandem Mass Spectrometry (UPLC-QToF-MS/MS)

Accurate mass-to-charge measurements were conducted on a UPLC–QTOF-MS/MS (Xevo G2-XS QToF mass spectrometer), comprising an Acquity UPLC (Waters Corporation, Milford, MA, USA) and using MassLynx software version 4.1. A BEH C18 column (100 x 2.1 mm, 1.7 µm, Waters)was used.

2.2.7. Ultra-Performance Liquid Chromatography – Triple Quadruple – Tandem Mass Spectrometry (UPLC-TQD-MS/MS)

An Acquity UPLC system, consisting of an autosampler and a binary pump (Waters, Milford, MA), was used. Compounds were separated on an Acquity BEH C18 column (2.1 mm × 100 mm, 1.7 mm; Waters, Milford, MA). The UPLC system was coupled to a triple quadrupole (TQD) mass spectrometer (Waters, Milford, MA) equipped with an electrospray ionization (ESI) source. All data were recorded and processed using the Quanlynx package in Masslynx software, version 4.1 (Waters, Milford, MA).

2.2.8. Nuclear magnetic resonance (NMR)

NMR spectra were recorded either on a DRX-400 or an Avance Nanobay III NMR instrument (Bruker BioSpin, Rheinstetten, Germany), both operating at 400 MHz for ^{1}H and 100 MHz for ^{13}C analysis. NMR data processing was performed using TopSpin software version 4.0.6 from Bruker.

2.2.9. Optical Rotation (OR)

Optical rotation was measured on a Jasco P-2000 spectropolarimeter (Easton, MD, USA), equipped with Spectra Manager™ software, at the sodium D-line wavelength (589.3nm).

2.2.10. Electronic Circular Dichroism (ECD)

ECD spectra were recorded using a Chirascan™ -Plus spectrophotometer (Applied Photophysics Ltd., Leatherhead, UK), equipped with a SUPRASIL® quartz cell (Hellma BeNeLux, Kruibeke, Belgium) with a path length of 2 mm. Measurement conditions and settings were as follows: continuous flushing of the instrument with

nitrogen gas (4 L.min^{-1} flow rate), temperature: 20 °C, bandwidth: 1 nm and acquisition time: 1 s.nm^{-1}.

2.2.11. Vibrational Circular Dichroism (VCD)

The experimental IR and VCD spectra were recorded on a Bruker Invenio FTIR spectrometer equipped with a PMA50 accessory (Bruker Optics, Ettlingen, Germany). All measurements were performed in $CDCl_3$ at a concentration of 0.12 M. A cell with 100 µm path length and BaF_2 windows was used. Both the sample and the solvent spectrum were recorded with a resolution of 4 cm^{-1}, totaling 30 000 scans each, with the PEMs optimized at 1400 cm^{-1}. The final baseline-corrected VCD spectrum was obtained through subtraction of the solvent spectrum.

2.3. Molecular networking

Molecular networks are visual displays showing the chemical patterns found in tandem mass spectrometry (MS/MS) experiments. This representation can detect groups of similar spectra (fragmentation) from both known and unknown compounds. In Global Natural Products Social Molecular Networking (GNPS), a web-based mass spectrometry ecosystem, each spectrum represents as a node and connects with related spectra by lines. Nodes can also have extra information, like matching known compounds or user-provided data about features like quantity, origin, or properties, which can change the size or color of the nodes. This map of related molecules is called a cluster. All clusters of a sample are considered a global molecular network [1].

2.4. DP4 and DP4+ probabilities

DP4 and DP4+ probabilities combine computational chemistry and mathematics to assist the assignment of relative configuration [2, 3]. By comparing mathematically the experimental chemical shifts with all possibilities computed by density functional theory (DFT) or other quantum chemical theories, relative configuration of a compound can be proposed [2, 3]. The computational resources used in this work were provided by the HPC core facility CalcUA of the University of Antwerp and VSC (Flemish Supercomputer Center).

2.5. Anti-SARS-CoV-2 assay

2.5.1. Cells and Virus

VeroE6-eGFP cells (provided by Dr. K. Andries, J&JPRD, Beerse, Belgium) were maintained in DMEM (Gibco cat no 41965-039) supplemented with heat-inactivated 10% FBS and 500 µg/mL Geneticin (Gibco cat no 10131-0275) and kept in a humidified 5% CO_2 incubator at 37 °C. For the production of virus stocks and during virus experiments, the cells were maintained in an assay medium. The VeroE6-eGFP assay medium consisted of DMEM supplemented with 2% FBS.
The SARS-CoV-2 isolate used in this study was the BetaCov/Belgium/GHB-03021/2020 (EPI ISL407976|2020-02-03), which was isolated from a Belgian patient returning from Wuhan in February 2020. The isolate was passaged seven times on Vero-E6 cells, which introduced two series of amino acid deletions in the spike protein [4]. The infectious content of the virus stock was determined by titration on Vero-E6 cells. The SARS-CoV-2-related work was conducted in the high-containment BSL3+ facilities of the KU Leuven Rega Institute (3CAPS) under licenses AMV 30112018 SBB 219 2018 0892 and AMV 23102017 SBB 219 2017 0589 according to institutional guidelines.

2.5.2. Antiviral Assay

The SARS-CoV-2 antiviral assay in Vero-E6 cells is derived from a previously established SARS-CoV-1 assay [5]. Stock solutions of the various compounds in DMSO (10 mM) were prepared. On day -1, the test compounds were serially diluted in assay medium and VeroE6-eGFP cells were plated corresponding to a final density of 25,000 cells per well in black 96-well plates (Greiner Bio-One, Vilvoorde, Belgium; Catalog 655090). The plates were incubated overnight (37 °C, 5% CO_2 and 95% relative humidity). On day 0, the cells with compound were infected with SARS-CoV-2 (at 20 $CCID_{50}$ per well). The plates were incubated in a humidified incubator at 37 °C and 5% CO_2. At 4 days, the wells were examined for eGFP expression using an argon laser-scanning microscope. The microscope settings were excitation at 488 nm and emission at 510 nm, and the fluorescence images of the wells were converted into signal values. The antiviral activity was expressed as EC_{50} defined as the concentration of compound achieving 50% inhibition of the virus-reduced eGFP signals as compared to the untreated virus-

infected control cells. The cytotoxicity of the compounds for VeroE6 cells in the absence of virus was evaluated in a standard MTS-assay as described previously [6].

2.6. References

1. Wang, Mingxun, et al. "Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking." *Nature biotechnology* 34.8 (2016): 828-837.
2. Smith, Steven G., and Jonathan M. Goodman. "Assigning stereochemistry to single diastereoisomers by GIAO NMR calculation: The DP4 probability." *Journal of the American Chemical Society* 132.37 (2010): 12946-12959.
3. Grimblat, Nicolás, María M. Zanardi, and Ariel M. Sarotti. "Beyond DP4: An improved probability for the stereochemical assignment of isomeric compounds using quantum chemical calculations of NMR shifts." *The Journal of organic chemistry* 80.24 (2015): 12526-12534.
4. Boudewijns, R.; Thibaut, H.J.; Kaptein SJ, F.; Li, R.; Vergote, V.; Seldeslachts, L.; Van Weyenbergh, J.; De Keyzer, C.; Bervoets, L.; Sharma, S.; et al. STAT2 Signaling Restricts Viral Dissemination but Drives Severe Pneumonia in SARS-CoV-2 Infected Hamsters. *Nat. Commun.* 2020, *11*, 1–10.
5. Ivens, T.; Van Den Eynde, C.; Van Acker, K.; Nijs, E.; Dams, G.; Bettens, E.; Ohagen, A.; Pauwels, R.; Hertogs, K. Development of a Homogeneous Screening Assay for Automated Detection of Antiviral Agents Active against Severe Acute Respiratory Syndrome-Associated Coronavirus. *J. Virol. Methods* 2005, *129*, 56–63.
6. Jochmans, D.; Leyssen, P.; Neyts, J. A Novel Method for High-Throughput Screening to Quantify Antiviral Activity against Viruses That Induce Limited CPE. *J. Virol. Methods* 2012, *183*, 176–179.

CHAPTER 3

Epimeric mixture analysis with DFT calculations

Publication:

Le Ngoc-Thao-Hien, Tom Vermeyen, Roy Aerts, Wouter A. Herrebout, Luc Pieters, and Emmy Tuenter. 2023a. "Epimeric Mixture Analysis and Absolute Configuration Determination Using an Integrated Spectroscopic and Computational Approach—A Case Study of Two Epimers of 6-Hydroxyhippeastidine". Molecules 28, no. 1: 214.

Abstract:

Structure elucidation has always been challenging, and misassignment remains a stringent issue in the field of natural products. The growing interest in discovering unknown, complex natural structures accompanies the increasing awareness concerning misassignments in the community. The combination of various spectroscopic methods with molecular modeling has gained popularity in recent years. In this work, we demonstrated, for the first time, its power to fully elucidate the connectivity and 3-dimensional structures of two epimers in an epimeric mixture of 6-hydroxyhippeastidine. DFT calculation of chemical shifts was first performed to assist the assignment of planar structures. Furthermore, relative and absolute configurations were established by three different ways of computer-assisted structure elucidation (CASE) coupled with ORD/ECD/VCD spectroscopies. In addition, the significant added value of OR/ORD computations to relative and absolute configuration determination was also revealed. Remarkably, the differentiation of two enantiomeric scaffolds (crinine and haemanthamine) was accomplished via OR/ORD calculations with the cross-validation by ECD and VCD spectroscopy.

Keywords:

6-hydroxyhippeastidine; epimeric mixture analysis; DFT calculation; OR/ORD; ECD; VCD; CASE

3.1. Introduction

Natural products chemistry is one of the oldest research areas. After more than a century of development, a wide variety of hit compounds are still being isolated from natural sources [1]. Thanks to the continuous advancement of chromatographic and spectroscopic techniques, scientists are able to explore minor and/or complex components. In recent years, the application of quantum chemical calculations has facilitated accurate, highspeed, theoretical methods to predict 2D and 3D molecular structures [2]. Reliable methods were established and optimized to compute NMR chemical shifts and coupling constants with minimal computational cost [2,3]. Regarding 3D structural identification, a breakthrough was first made by the Goodman group, introducing the CP3 parameters in 2009, followed by the DP4 probabilistic method in 2010 [4,5]. While the application of the former method is limited, the latter was successfully applied in many cases to confirm or correct the relative configuration of natural products of varying complexity [6–8]. More recently, the Sarotti group introduced two modified probabilistic methods based on the mathematical core of the DP4 probability: DP4+ and J-DP4 [9,10]. Many successful applications of the DP4+ probability method have already been reported in the field of natural products [7,11]. Most recently, the DP5 method was published by the Goodman group in 2022 [12]. Nonetheless, the computational effort associated with these methods can be significant, especially in the case of large and flexible molecules. In order to reduce this large workload, several (semi-)automatic procedures were developed, using freely accessible programming languages (Python, Bash and R), Excel (DP4+ and J-DP4 templates), machine learning and deep learning [9,13,14]. Software applications implementing Computer-Assisted 3D Structure Elucidation (CASE-3D) were also developed recently, such as the StereoFitter from Mestrelab MNova, CMC-se from Bruker, Structure Elucidator from ACD labs and Logic Structure Determination (LSD) from Nuzillard and Plainchont [15–18]. The considerable efforts made by theoretical chemists to introduce these methods to other disciplines, such as organic synthesis, biosynthesis and natural products chemistry, have resulted in a rapidly increasing number of citations within the last five years [2,11,19–20].

Over the last two decades, quantum chemical predictions of optical rotation (OR) or optical rotation dispersion (ORD) have become a fast and reasonably reliable tool to determine absolute configurations [21–23]. Several successful applications of this method were published recently, as such or in combination with other spectroscopic techniques such as electronic circular dichroism (ECD), vibrational circular dichroism (VCD) and/or Raman optical activity (ROA) [24–26]. As a result of these theoretical calculations, the information obtained from OR/ORD was extended and is no longer limited to just providing a single value and defining a compound as dextrorotatory or levorotatory. By comparing the computed and experimental OR/ORD values, possibly at multiple wavelengths, the sign of the rotation can be translated into an enantiomer described using R/S nomenclature.
In the structural elucidation of natural products field, misassignment remains a stringent and commonly unnoticed issue [27–32]. In the present study, a rational and general strategy using a combination of empirical and computational data was ap-plied, combining all the above-mentioned advancements (Figure 1). A major asset of this workflow lies in combining these various methods. Thus, each structural property of the molecule can be extracted from multiple sources, making the elucidation more robust and minimizing the probability of misassignment. For the first time, the effectiveness of the workflow was tested on an epimeric mixture of two 6-hydroxyhippeastidines purified from the plant species Hymenocallis littoralis (Amaryllidaceae). The crinine-type and haemanthamine-type scaffolds, widely distributed in the Amaryllidaceae family, are characterized by the arylhydroindole ring system and are composed of over 50 members in nature [33,34]. In this work, two epimers of 6-hydroxyhippeastidine, belonging to the crinine-type subgroup, were studied. Their 2D chemical structures were elucidated based on the basis of extensive 1D and 2D NMR spectroscopy and HRMS data. Relative and absolute configurations were determined by pooling the results obtained with optical rotation, nuclear Overhauser effect (NOE), DP4+, ECD and VCD analyses.

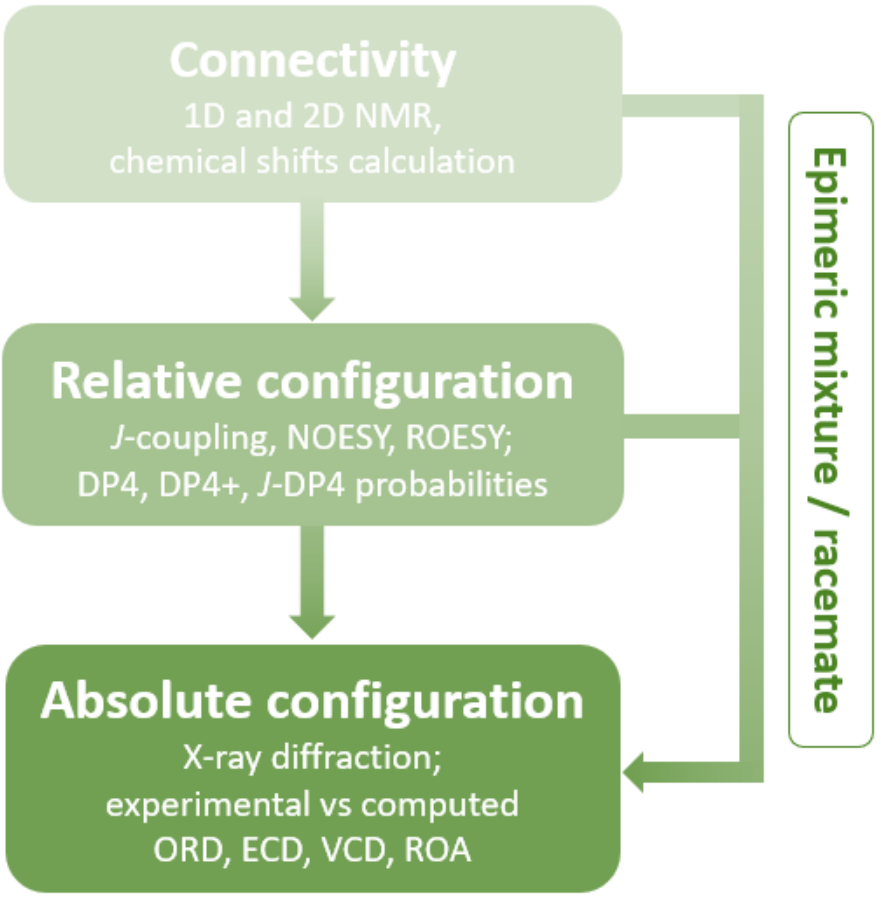


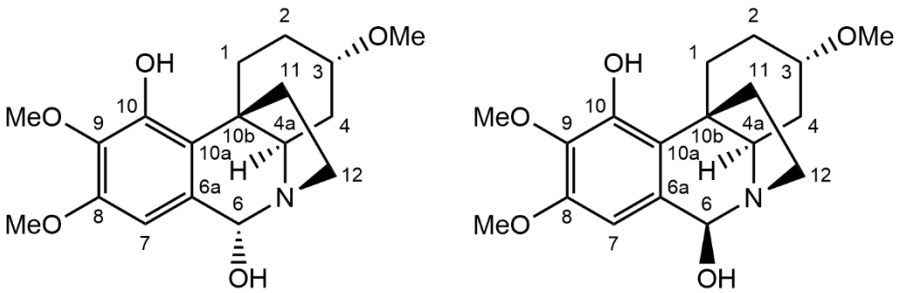
Figure 1. General Structural Elucidation Workflow.

3.2. Results and Discussion

3.2.1. Structure Elucidation of connectivity

Compounds 1 and 2 were isolated as an inseparable epimeric mixture (ratio ~ 3.15:1, averaged from the ratios of two corresponding pairs of the H-7 and H-6 ^{1}H-NMR signals of the two compounds, as shown in Figures 2 and 3). The crinine- or haemanthamine-type skeleton was first proposed for compounds **1** and **2** based on a comparison with the NMR data of haemanthidine in the literature [33]. Haemanthidine is a well-known mixture of two epimers and is widely distributed in the Amaryllidaceae family. In the case of haemanthidine, both 6-epimers were

present (adopting the same atom numbering system commonly used in the literature), as deduced from the difference in the chemical shift of H-6: the H-6 of the major epimer was found at 5.02 ppm, and the H-6 of the minor epimer was found at 5.65 ppm in $(CD_3)_2SO$ [34]. A large similarity was observed for the epimeric center of compounds **1** and **2**: the H-6 of the major epimer resonated at 4.72 ppm, while that of the minor epimer resonated at 5.43 ppm in $(CD_3)_2SO$. Continuing to use haemanthidine as the reference point, two noticeable differences were observed by comparing the ^1H NMR spectra: Firstly, a signal around 5.95 ppm with an integration of the two was present in the case of haemanthidine but was absent in the case of compounds **1** and **2**, indicating the absence of the methylenedioxy moiety. Secondly, in the downfield region, four signals corresponding to aromatic and olefinic hydrogens were observed in the ^1H NMR spectrum of haemanthidine, whilst only one aromatic hydrogen was found for compounds **1** and **2**. This infers that the double bond present in haemanthidine was absent in compounds **1** and **2** and that compounds **1** and **2** possess one more substituent on the benzene ring compared to haemanthidine. This assumption was supported by the appearance of two additional methoxy signals around 3.65–3.75 ppm and five additional aliphatic hydrogens in the upfield region from 1.00 to 2.50 ppm in the ^1H NMR spectrum of the mixture of epimers. Next, the 2D spectra were inspected. Five methylene and three methoxy groups were first determined in the ^1H NMR and HQSC spectra. Furthermore, two spin systems were observed in COSY: the first one for H-1 (2H), H-2 (2H), H-3 (1H) and H-4 (2H) and the second one for H-11 (2H) and H-12 (2H) (see Figure 4). A strong HMBC signal between the methoxy group (H at 3.24 ppm) and C-3 (77.6 ppm) inferred that they were connected. Notably, since the NMR signals of compound **2** were much less intense than those of compound **1** in the mixture, an additional long-range HMBC spectrum was recorded to observe the 4- and 5-bond HMBC correlations. This led to the observation of a peak between the H-6 and C-10b of compound **2** and to the identification of the signal corresponding to C-10b.



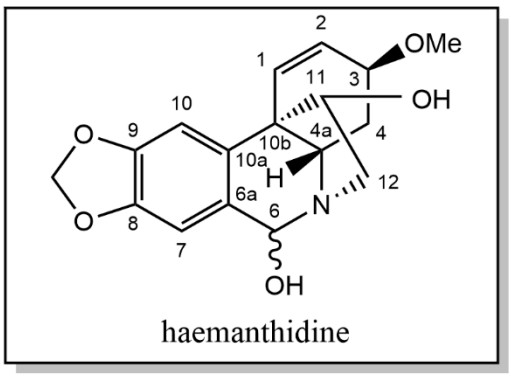


Figure 2. Structures of 6α- and 6β-hippeastidine (compounds **1** and **2**) and haemanthidine.

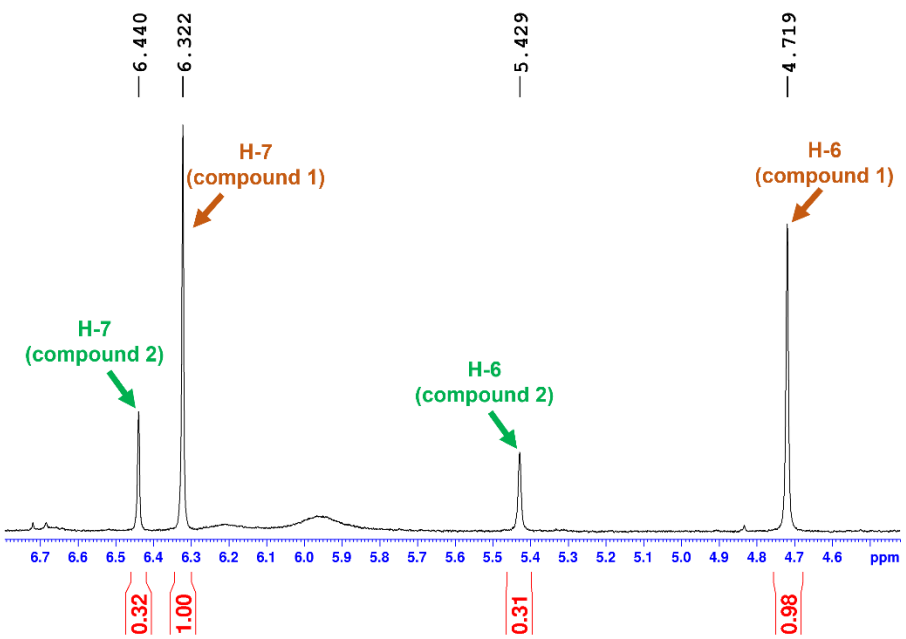


Figure 3. Epimeric ratio of the mixture of compounds **1** and **2**, deduced from the ^{1}H-NMR spectrum.

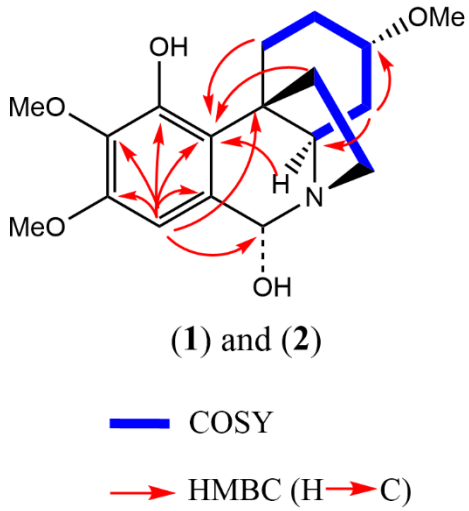
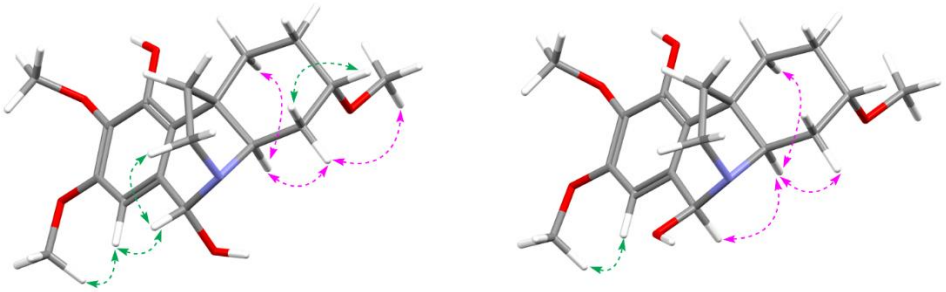


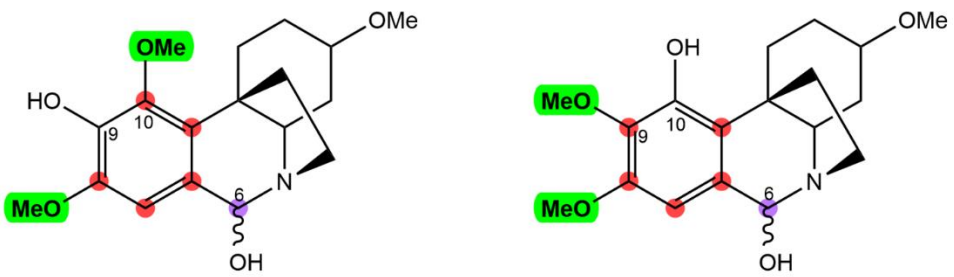
Figure 4. Key HMBC and COSY correlations for two 6-epimers of 6-hydroxyhippeastidine.

With regard to the aromatic moiety, the proton signal at 6.32 ppm only showed an HMBC correlation with C-6 and not with any other CH or CH_2 group of the aliphatic moiety, indicating that this proton signal was present at position 7. After assigning 3-OMe, two methoxy-groups remained, which were linked to the benzene ring. The position of the 8-OMe was confirmed by a NOESY cross-peak between H-7 and 8-OMe (Figure 5). Nonetheless, the HMBC spectrum confirmed that the last methoxy group was present at position 10, given the correlations of H-6 with C-6a, C-7, C-8, C10a and one additional signal of the methoxylated carbon, which most probably would be present in position 10. Hence, the hydroxy group would be present in the only remaining aromatic position, position 9, and the NMR signal of C-9 was found at 147.7 ppm (see Figure 6, structure a). Combining all this information and using HMBC and NOESY correlations to procure the assignments, compounds **1** and **2** were initially identified as the two 6-epimers of 6-hydroxyhippeastidine (see Figure 6, structure a).



(**1**) Major epimer (**2**) Minor epimer

Figure 5. Selected NOESY correlations for two 6-epimers of 6-hydroxyhippeastidine. Green arrows: NOESY correlations in front of the plane (β-orientation); pink arrows: NOESY correlations behind the plane (α-orientation).



(a) incorrect assignment (b) correct assignment

Figure 6. Observed HMBC correlations between H-6 (in purple) and carbons of the aromatic ring (in red) of compounds **1** and **2**: (a) incorrect assignment and (b) correct assignment.

However, surprisingly, the calculated chemical shifts of the proposed structures (Figure 6, structure a) were not in full agreement with the experimental data, in particular regarding the aromatic moiety. As listed in Tables 1 and 2, alarmingly high discrepancies (~10 ppm) were found for C-9 and C-10a, while their surrounding carbons showed 4–6 ppm errors (C-6a, C-8 and C-9). This raises a question about the assignment at the aromatic part of compounds **1** and **2** [11]. Based on the differences in shielding intensities of C-9 and C-10 (C-9 is more deshielded than C-10, according to the experimental data, while for the computations the opposite was observed), a new structural proposition (structure b) was considered, in which the C-9 and C-10 substituents were switched (Tables 1 and 2 and Figure 6). Finally, chemical shifts of structure b were computed following the approach reported by the authors [31]. Indeed, computed chemical shifts of

structure b showed an utter resemblance to the experimental values, with a CMAE of 1.04 ppm and a maximum outlier of 2.83 ppm for the carbons of compound **1** (Table 1) and a CMAE of 1.09 ppm and a maximum outlier of 2.94 ppm for the carbons of compound **2** (Table 2). Hence, structure b was confirmed to be correct, and compounds **1** and **2** were the two 6-epimers of structure b, which were reported in the plant species *Zephyranthes candida* by Shitara et al. (6α- and 6β-hydroxyhippeastidine) [35]. Remarkably, the HMBC correlation between H-6 and C-9 is a ^{5}J interaction, which is unexpected and rarely observed, while the ^{4}J interaction between H-6 and C-10 was absent in both the regular and long-range HMBC spectra (Figure 6, structure b).

Table 1. Experimental and Computed NMR Spectroscopic Data (400 MHz, $(CD_3)_2SO$) of Compound **2** (major epimer)

	Experimental		Calculated			
			Structure proposal 2a		Structure proposal 2b	
position	δ_C, type	δ_H (*J* in Hz)	δ_C	δ_H	δ_C	δ_H
1α	27.1, CH_2	1.55	27.7	1.60	27.9	1.63
1β	27.1, CH_2	3.12*	27.7	2.99	27.9	2.94
2α	27.9, CH_2	1.90 m	27.9	1.90	28.2	1.88
2β	27.9, CH_2	1.30 m	27.9	1.32	28.2	1.29
3	77.6, CH	3.10*	76.7	3.01	77.1	3.03
4α	33.7, CH_2	1.82 m	33.2	1.76	33.5	1.76
4β	33.7, CH_2	1.05 q (11.8)	33.2	1.10	33.5	1.08
4a	61.3, CH	3.07*	62.4	2.99	62.5	3.01
6	88.8, CH	4.72 s	88.9	4.68	89.0	4.75
6a	132.0, C	-	127.4	-	132.8	-
7	103.8, CH	6.32 s	106.2	6.45	103.5	6.36
8	150.9, C	-	146.3	-	151.4	-
9	135.8, C	-	144.6	-	134.8	-
10	147.7, C	-	140.9	-	147.9	-
10a	127.3, C	-	137.2	-	128.4	-
10b	43.3, C	-	46.2	-	46.1	-
11	33.3, CH_2	2.06 m	34.8	2.11	34.5	2.08
11	33.3, CH_2	1.51*	34.8	1.61	34.5	1.58
12	47.3, CH_2	3.12*	47.5	3.11	47.8	3.13
12	47.3, CH_2	2.52	47.5	2.55	47.8	2.56
3-OMe	55.3, CH_3	3.24 s	53.2	3.20	53.5	3.22
8-OMe	55.9, CH_3	3.72 s	52.9	3.68	53.2	3.65
9-OMe	60.6, CH_3	3.63 s	55.7	3.70	57.8	3.74
CMAE			3.0	0.06	1.04	0.06
Max. outlier			9.9	0.13	2.83	0.18

*overlapping signals

Table 2. Experimental and Computed NMR Spectroscopic Data (400 MHz, $(CD_3)_2SO$) of Compound **3** (minor epimer)

	Experimental		Calculated			
			Structure proposal 3a		Structure proposal 3b	
position	δ_C, type	δ_H (*J* in Hz)	δ_C	δ_H	δ_C	δ_H
1α	27.1	1.55	27.6	1.63	28.0	1.66
1β	27.1	3.12*	27.6	2.98	28.0	2.95
2α	27.9	1.90 m	27.8	1.89	28.0	1.88
2β	27.9	1.30 m	27.8	1.32	28.0	1.29
3	77.3	3.10*	76.6	3.01	76.8	3.03
4α	33.9	1.88 m	33.4	1.85	33.7	1.84
4β	33.9	1.13 q (11.8)	33.4	1.16	33.7	1.16
4a	66.3	2.95	67.0	2.85	67.1	2.88
6	86.8	5.43 s	87.8	5.37	87.9	5.42
6a	133.1	-	128.5	-	134.0	-
7	102.3	6.44 s	104.3	6.63	101.7	6.50
8	150.8	-	146.3	-	151.3	-
9	135.6	-	144.4	-	134.6	-
10	147.4	-	140.7	-	147.7	-
10a	126.5	-	136.0	-	127.0	-
10b	44.3	-	47.2	-	47.2	-
11	35.0	2.13 m	37.1	2.20	36.6	2.17
11	35.0	1.48*	37.1	1.54	36.6	1.52
12	41.9	3.17*	41.6	3.20	42.2	3.21
12	41.9	2.76	41.6	2.69	42.2	2.72
3-OMe	55.3	3.24 s	53.1	3.21	53.4	3.23
8-OMe	55.9	3.72 s	52.8	3.69	53.1	3.65
9-OMe	60.6	3.63 s	55.6	3.70	57.7	3.74
CMAE			3.1	0.06	1.09	0.06
Max. Outlier			9.5	0.16	2.94	0.17

*overlapping signals

3.2.2. Relative Configuration Determination

Further analysis of the NOE correlations (Figure 5) yielded the proposed relative configuration of the two compounds: (3***R***,4a***R***,6***R***)-6α-hydroxyhippeastidine or its enantiomer for compound **1** (major epimer) and (3***R***,4a***R***,6*S*)-6β-hydroxyhippeastidine or its enantiomer for compound **2** (minor epimer). Briefly, starting from the epimeric center at C-6, the H-6 (4.72 ppm) of compound 1 correlated with the H-12 of the ethanobridge, while the H-6 (5.43 ppm) of compound **2** did not, indicating that the H-6 of compound **2** is located on the same side of the ethanobridge (β-orientation). On the other hand, the H-6 (5.43 ppm) of compound 2 correlated with the H-4a (2.95 ppm), while the H-6 (4.72 ppm) of compound **1** did not, indicating that the H-4a of compounds **1** and **2** is on the opposite side of the ethanobridge and that the H-6 (5.43 ppm) of compound **2** and the H-4a were located on the same side (α-orientation). Finally, the H-4a correlated with the other hydrogens on the α-face (H-1α, H-2α, H-4α and 3-OMe), but not with the H-3, indicating the β-configuration of the H-3.

Before conducting DP4+ as the next step, it is noteworthy that the crinine type and haemanthamine type are two enantiomeric skeletons, defined by the orientation of the 5,10b-ethanobridge moiety: in the case of the haemanthamine skeleton, the ethanobridge is α-oriented, and in the case of crinine-type alkaloids, the ethanobridge is β-oriented [33,35]. However, computed chemical shifts of enantiomers will be identical, and, thus, only calculation of one set of enantiomers is required. The set of crinine-type diastereomers was selected to perform the calculations in this DP4+ probability assessment. Given the presence of the ethanobridge, which in the crinine-type skeleton is assigned the β-configuration, the number of stereogenic centers to consider is reduced from five (including the N at position 5) to three. Thus, eight diastereomers were computed for compounds **1** and **2**. According to Table 3, the DP4+ probabilities were in complete accordance with the preceding NOESY correlations. Indeed, the computed chemical shifts of the two proposed diastereomers showed the highest resemblance to the experimental chemical shifts. Thus, the relative configurations of compounds **1** and **2** were defined by two means: NOE correlations and DP4+ probability. The relative configuration assignment is clearly robust, leaving only the absolute configurations to be assigned.

Table 3. DP4+ probabilities of two 6-epimers of 6-hydroxyhippeastidine. Two bold letters R stand for the fixed configuration of the 5,10b-ethanobridge with respect to the crinine-type skeleton (3,4,5***R***,6,10b***R***).

Compound 1		**Compound 2**	
Diastereomer	Probability (%)	Diastereomer	Probability (%)
RR**R**R**R**	100	RR**R**R**R**	0
RR**R**S**R**	0	RR**R**S**R**	100
RS**R**R**R**	0	RS**R**R**R**	0
RS**R**S**R**	0	RS**R**S**R**	0
SR**R**R**R**	0	SR**R**R**R**	0
SR**R**S**R**	0	SR**R**S**R**	0
SS**R**R**R**	0	SS**R**R**R**	0
SS**R**S**R**	0	SS**R**S**R**	0

3.2.3. Absolute Configuration Determination

Using the relative configurations established in the previous section, the optical rotations were computed for both compounds. As indicated in Table 4, the optical rotation values were computed for four configurations, representing two pairs of enantiomers. More specifically, the diastereomers with the RRRRR and RRRSR configurations belong to the crinine-type skeleton, whilst the compounds with the SSSSS and SSSRS configurations belong to the haemanthamine-type skeleton. Since compounds **1** and **2** were obtained as a mixture with a ratio of 3.15:1.00, the calculated ORs had to be adjusted accordingly. In the case of the ORs computed by the Pople (6-311++G(3df,2dp)) basis set:

- the averaged OR of the crinine-type = $\frac{80.3*3.15+(-54.2)*1}{(3.15+1)}$ = 47.9;
- the averaged OR of the haemanthamine-type = $\frac{(-80.3)*3.15+54.1*1}{(3.15+1)}$ = −47.9.

In the case of the ORs computed by the Dunning (aug-cc-pVTZ) basis set:

- the averaged OR of the crinine-type = $\frac{80.1*3.15+(-53.7)*1}{(3.15+1)}$ = 47.8;
- the averaged OR of the haemanthamine-type = $\frac{(-80.1)*3.15+53.5*1}{(3.15+1)}$ = −47.9.

After comparison of the above computed ORs with the experimental value (43.7), it is apparent that compounds **1** and **2** are of the crinine type. As a result, the absolute configuration of compound **1** is (3***R***,4***R***,5***R***,6***R***,10b***R***) and that of compound **2** is (3***R***,4***R***,5***R***,6*S*,10b***R***). More interestingly, this result led to the conclusion that the OR/ORD calculation is capable of differentiating the crinine-type skeleton from the haemanthamine-type skeleton. As a technical remark, although the results obtained with the 6-311++G(3df,2dp) and aug-cc-pVTZ basis sets show a slight difference, the two basis sets both provided decent and conclusive results.

Table 4. Experimental and computed optical rotation values for two 6-epimers of 6-hydroxyhippeastidine (compounds **1** and **2**).

	Diastereomer	**OR**	**Theory Level**
Experimental		43.7	
Computed	RRRRR	80.3	B3LYP/6-31++G(d,p)// 6-311++G(3df,2dp)
	RRRSR	-54.2	
	SSSSS	-80.3	
	SSSRS	54.1	
	RRRRR	80.1	B3LYP/6-31++G(d,p)// aug-cc-pVTZ
	RRRSR	-53.7	
	SSSSS	-80.1	
	SSSRS	53.5	

The assignment of the absolute configuration of compounds **1** and **2** was confirmed by electronic circular dichroism (ECD) and vibrational circular dichroism (VCD) analyses. To this end, an ECD spectrum was recorded, following the conventional way to distinguish crinine-type and haemanthamine-type skeletons. In 1996, Wagner et al. reported a study on ECD for one crinine-type (crinidine) and two haemanthamine-type alkaloids (vittatine and 11-hydroxyvittatine), which later on has been used as an empirical rule for the determination of crinine- and haemanthamine-type skeletons [36]. Crinidine shows a maximum Cotton effect around 240 nm and a minimum Cotton effect around 290 nm. Vittatine and 11-hydroxyvittatine, on the other hand, display opposite Cotton effect patterns.

Figure 7 demonstrates the ECD spectrum recorded for the mixture of compounds **1** and **2**, which shows similar Cotton effects to crinidine. In comparison with other known crinine-type alkaloids, a larger similarity in Cotton patterns was observed for aulicine, an Amaryllidaceae alkaloid that is structurally close to compounds **1** and **2** [37]. Hence, it was confirmed that compounds **1** and **2** possess the crinine-type skeleton and that the results obtained from the OR/ORD calculation were the same as those from ECD.

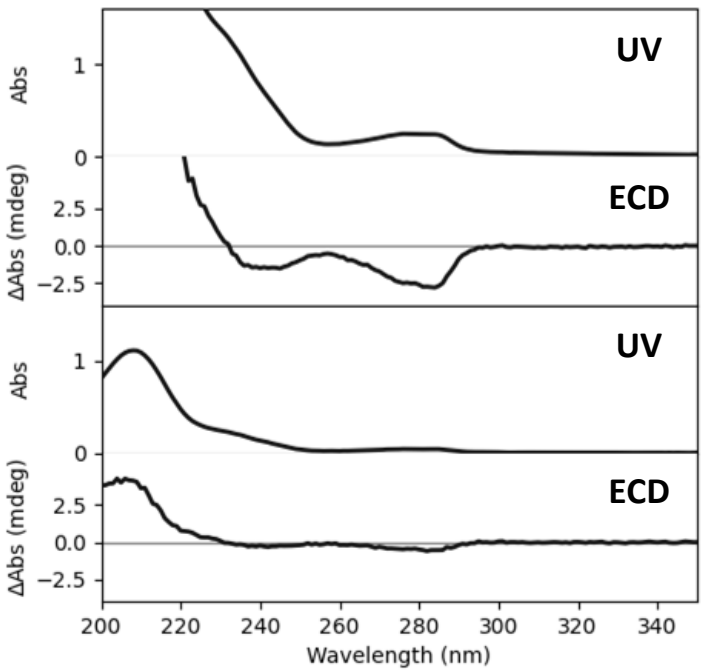


Figure 7. UV and ECD spectra of the mixture of compounds **1** and **2**. Top: spectra recorded for a high concentrated sample. Bottom: spectra recorded for a low concentrated sample.

The results of the VCD analyses are summarized in Figure 8. In the first step of the calculations, Boltzmann weighted IR and VCD calculations were performed at the B3LYP/6-31++G(d,p) level for both the major and minor epimers. This approach is similar to the one used recently to determine the absolute stereochemistry of the antibiotic polyketide tatiomicin [38]. Subsequently, the resulting spectra were combined, and a weighted IR and VCD spectrum was obtained using the 3.15:1.00 ratio, as determined from the NMR intensities. The experimental data obtained, the calculated IR and VCD spectra for the epimers involved and the resulting

spectra for the mixture are shown in Figure 8. A good agreement supporting the absolute configurations of compounds **1** and **2** proposed above is found between the experimental data in the upper panels and the final calculated data shown in the lower panels.

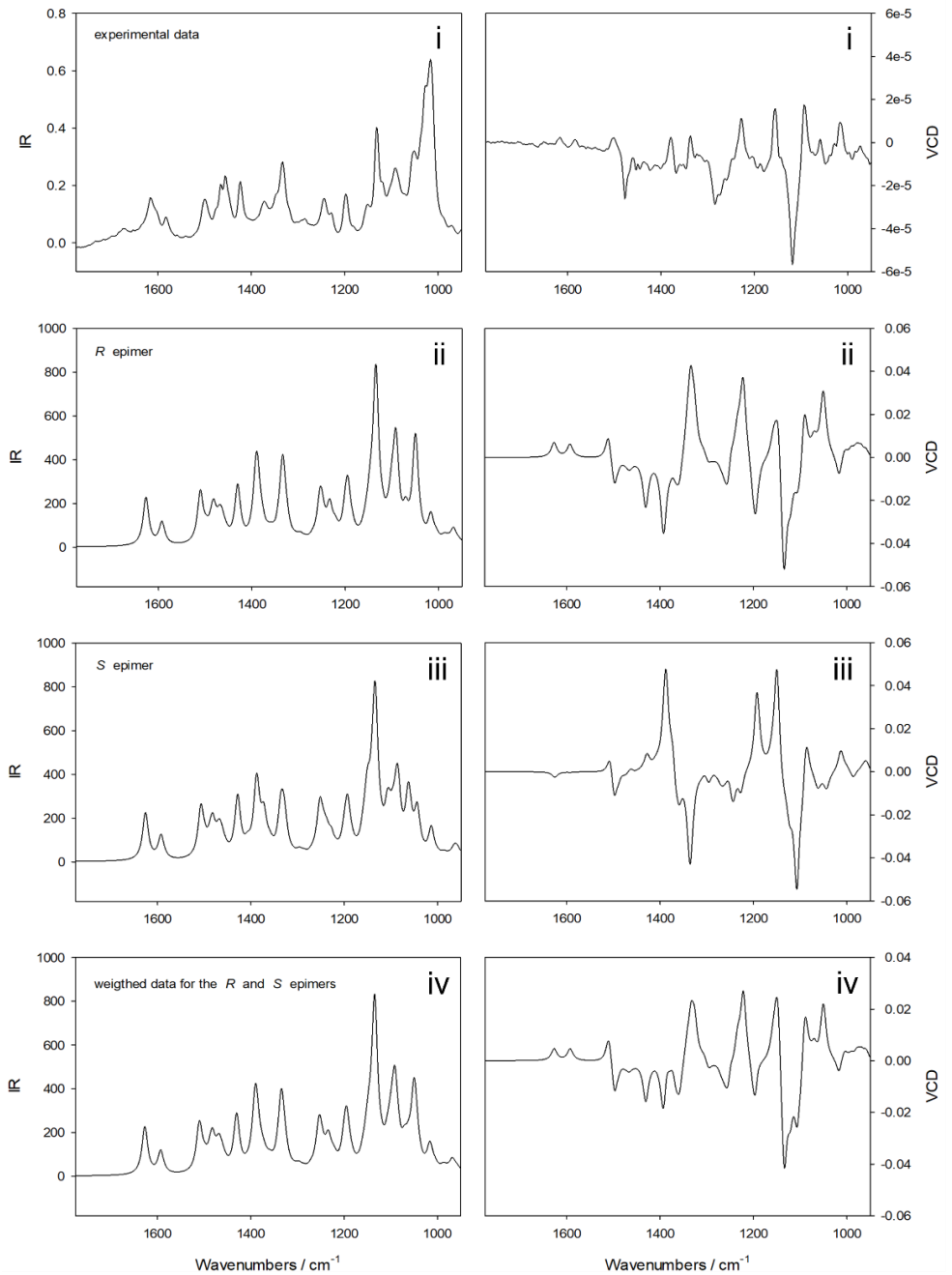


Figure 8. Experimental IR and VCD spectra confirming the absolute configurations of compounds **1** and **2**. From top to bottom, the data shown refer to (i) the experimental data obtained for the sample, (ii) the Boltzmann weighted IR and VCD spectra for the (3*R*,4*R*,5***R***,6*R*,10b***R***) epimer, (iii) the Boltzmann weighted IR and VCD spectra for the (3*R*,4*R*,5***R***,6*S*,10b***R***) epimer, and (iv) the weighted IR and VCD spectra obtained using the 3.15:1.00 ratio for major and minor epimers as derived from the NMR data.

The general workflow demonstrated in Figure 1 was successfully applied to the epimeric mixture in this study. Thus, this work showed the added value of quantum chemical calculations to predict chemical shifts in the structure elucidation of the 2D structure, in addition to the interpretation of the experimental 1D and 2D NMR data. When the discrepancies between the experimental and calculated data are >1 ppm for the proton data and >10 ppm for the carbon data, caution should be taken [2,11]. The experimental OR value of the epimeric mixture was successfully reproduced by quantum chemical calculations when taking into account the ratio of the two epimers in the mixture.

Whilst ECD still prevails and vibrational optical activity (VOA)-based methods (VCD and ROA) are gaining popularity, in our opinion OR/ORD is underrated and is favorable in terms of simplicity, accessibility and cost-effectiveness for the absolute configuration determination of relatively rigid, small chiral molecules in natural products research. OR/ORD can be applied to support ECD, VCD and ROA results, while in some cases OR/ORD data alone are even sufficient to draw conclusions regarding the absolute configuration of a certain compound [39–42]. Moreover, interpretation of OR/ORD data is straightforward, which makes this method suitable for application as a routine practice in natural products chemistry. Nevertheless, it is vital to emphasize that there is no one-size-fits-all technique, and the required type(s) of spectroscopic methods should be assessed on a case-by-case basis [41]. From our perspective, the following five criteria are essential for the proper application of OR/ORD for absolute configuration determination: (1)—experimental data are obtained for the purified compound (or pure mixture) in an appropriate amount depending on instrumental sensitivity and molecular properties; (2)—NOE correlations are used for relative configuration determination or as a filter to reduce the pool of computed diastereomers [43];

(3)—identical or strictly equivalent conformer pools of a pair of enantiomers are required; (4)—at least two different types of basis sets of a sufficiently large size should be used for computing OR/ORD; (5)—the need for other chiroptical/spectroscopic techniques should be considered when experimental OR/ORD values are close to zero, as is the case for racemic mixtures.

3.3. Conclusion

Epimeric mixture analysis has always been challenging in natural products research, since the identification and isolation of single compounds is highly demanding. This work served as a case study to explore the power of combining various spectroscopic methods and computational chemistry in the structural elucidation procedure, from planar structure to absolute configuration, of an epimeric mixture. In addition, the potential of OR/ORD simulation is highlighted for various applications.

Another interesting finding from our experiments is that crinine and haemanthamine skeletons (two enantiomeric scaffolds) can be differentiated by OR/ORD calculation, without the need for an ECD and/or VCD experiment. This will aid in simplifying the structural elucidation process of compounds possessing crinine or haemanthamine skeletons. Hence, this could be considered as a new benchmark for the distinguishment of the two skeleton types, since the publication of Wagner et al. in 1996 [36]. For the future, since this work only dealt with an example, it would be interesting to investigate a library of crinine-type and haemanthamine-type analogues to further confirm this finding.

3.4. Materials and Methods

3.4.1. Isolation and Purification

The mixture of two epimers was isolated from the bulbs of *Hymenocallis littoralis*. Briefly, dried plant material was macerated with 80% MeOH to obtain crude extract. The crude extract was acidified by HCl 5% to pH < 3 and, thereafter, partitioned with dichloromethane (x3). The crude dichloromethane extract was fractionated by flash chromatography to obtain 15 fractions. The two epimers were purified from fraction 10 by preparative HPLC using the following conditions:

column Kinetex C18 (250 × 4.6 mm, 5 µm), solvents (0.1% FA in H_2O and ACN), flow rate: 3 mL/min and fraction trigger: 336.3 *m/z*.

3.4.2. Computational Details

Conformational search was performed with PCMODEL (version 10.0) using the Monte Carlo algorithm and Merck Molecular Force Field (MMFF94). Two energy windows (8 and 7 kcal.mol^{-1}) were set for two consecutive conformational search cycles. Next, the Density Functional Theory (DFT) method, utilizing the exchange-correlation functional B3LYP (hybrid three-parameter Becke–Lee–Yang–Parr), was applied for the following calculations. All resulting conformers were subjected to geometry optimization using B3LYP/6–31G (d) level of theory in gas-phase, and isotropic shielding tensors were computed at the PCM//mPW1PW91/6-311+G(d,p) level of theory by the GIAO (gauge-independent atomic orbital) method. Boltzmann-averaged shielding tensors were used as input for carrying out the DP4+ probability. The DP4+ probability method using Bayes' theorem was applied in order to provide additional proof of relative configurations.

After dereplication, Boltzmann averaging was applied using the sum of electronic and thermal free energies at 298.15 K, and only conformers with energies within an energy window of 2.5 kcal.mol^{-1} from the global minimum were considered as contributing to the population. Geometry optimization, frequency, shielding tensor and specific rotation calculations were performed by Gaussian16. Avogadro (version 1.2.0) was used for the visualization of computed outputs. Detailed information about results of DP4+ analysis, Boltzmann distributions and 3D coordinates of contributing conformers can be found in Supporting Information.

Specific optical rotations were computed for the sodium D-line wavelength (589.3 nm). All conformers were optimized at the B3LYP/6-31++G(d,p) level of theory. The polarizable continuum model (PCM) was selected to improve the OR calculation accuracy (chloroform was used for the solvation model). The DFT method using the B3LYP functional combined with the two following basis sets was applied: a Pople 6-311++G (3df,2dp) and a Dunning (aug-cc-pVTZ), as recommended by Yu et al. (2012) and Stephens et al. (2001), respectively [44,45].

3.5. Supporting information – Chapter 3

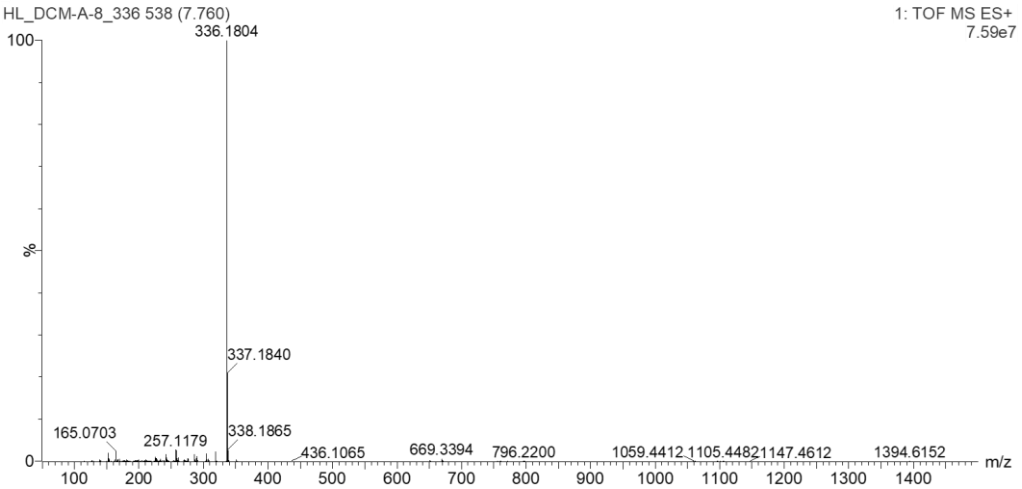


Figure S1. HRMS spectrum of compounds 1 and 2

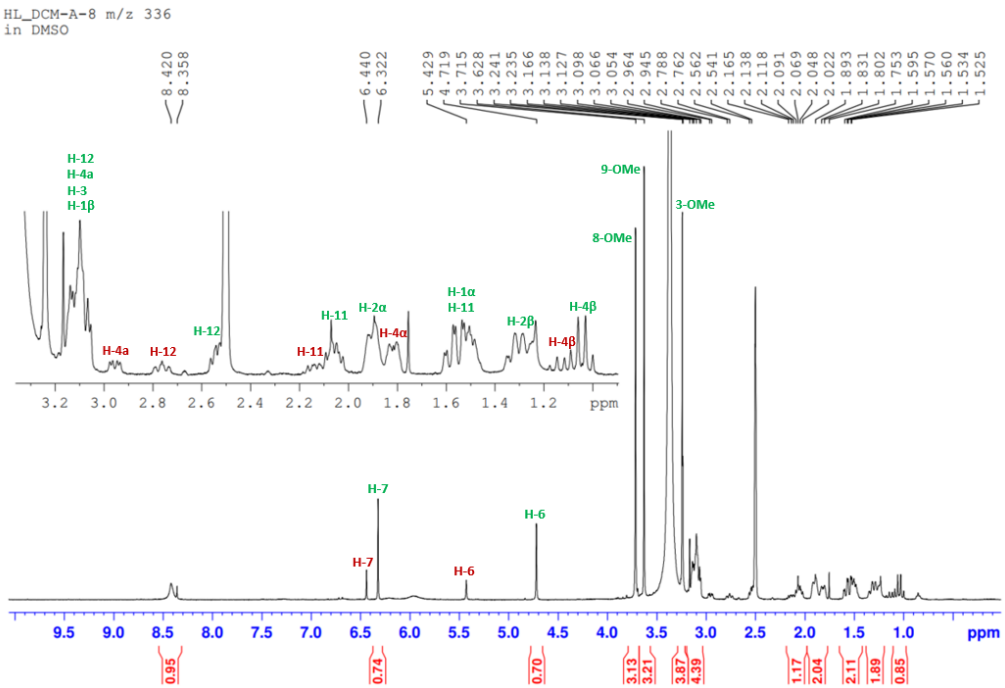


Figure S2. ^{1}H spectrum of compounds **1** and **2** in $(CD_3)_2SO$ (minor epimer, major epimer and overlapping signals)

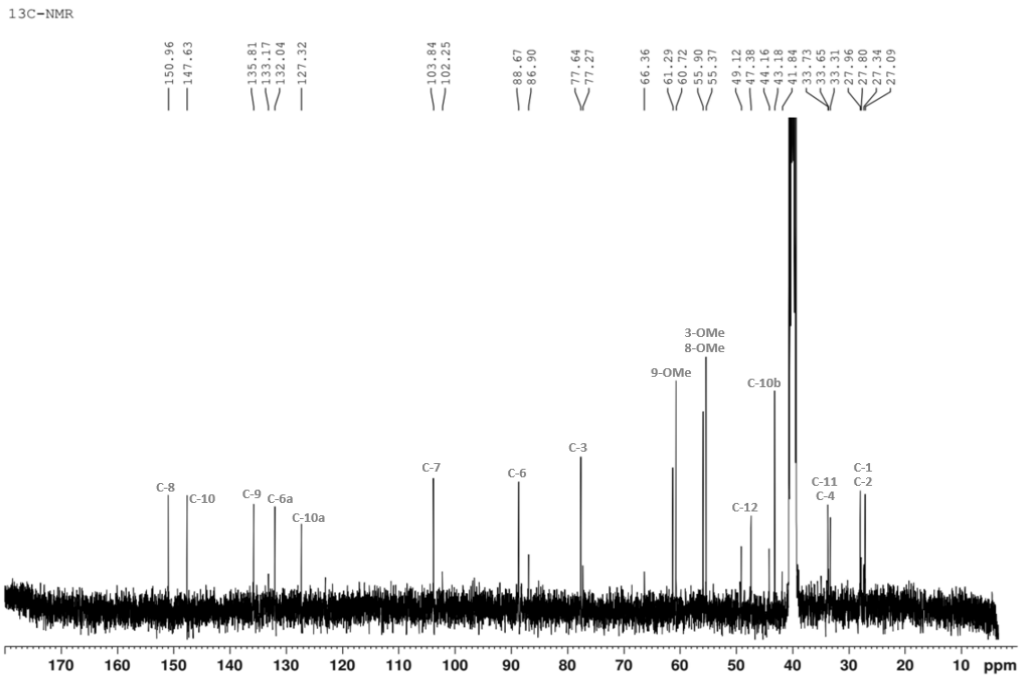


Figure S3. ^{13}C spectrum of compounds **1** and **2** in $(CD_3)_2SO$. Carbon signals of the major epimer was labelled.

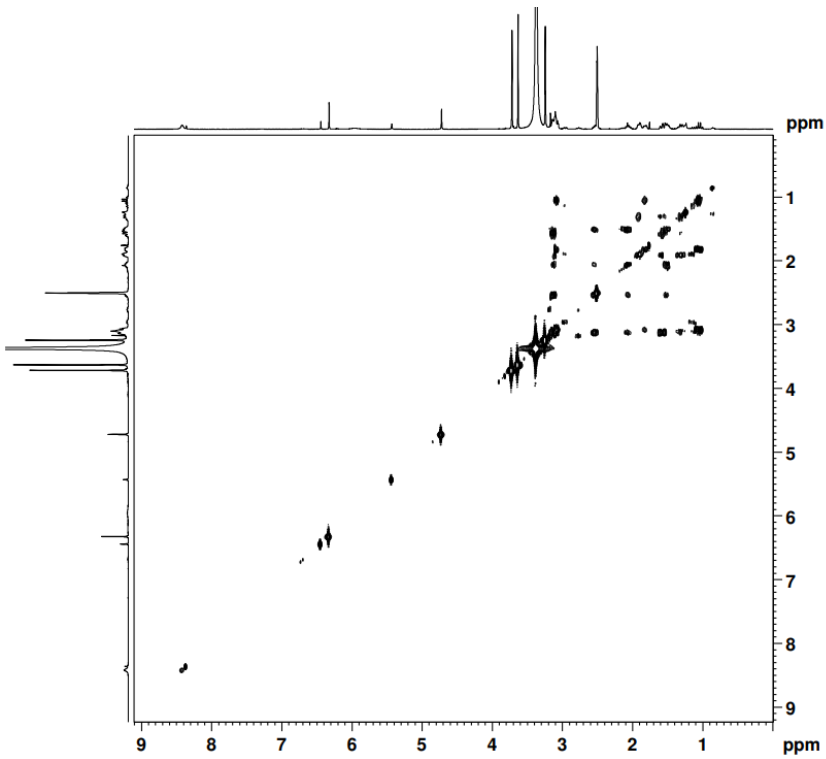


Figure S4. COSY spectrum of compounds **1** and **2** in $(CD_3)_2SO$

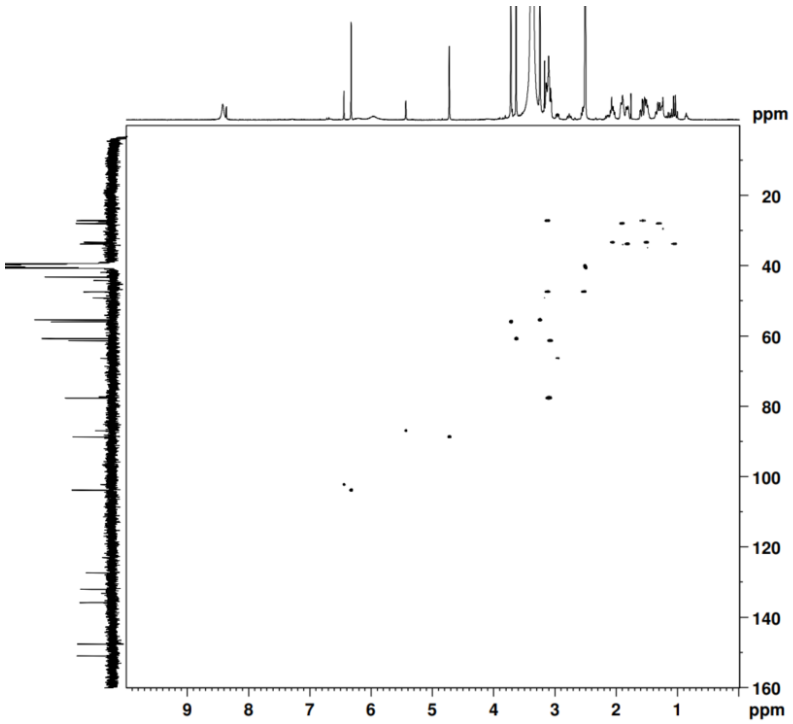


Figure S5. HSQC spectrum of compounds **1** and **2** in $(CD_3)_2SO$

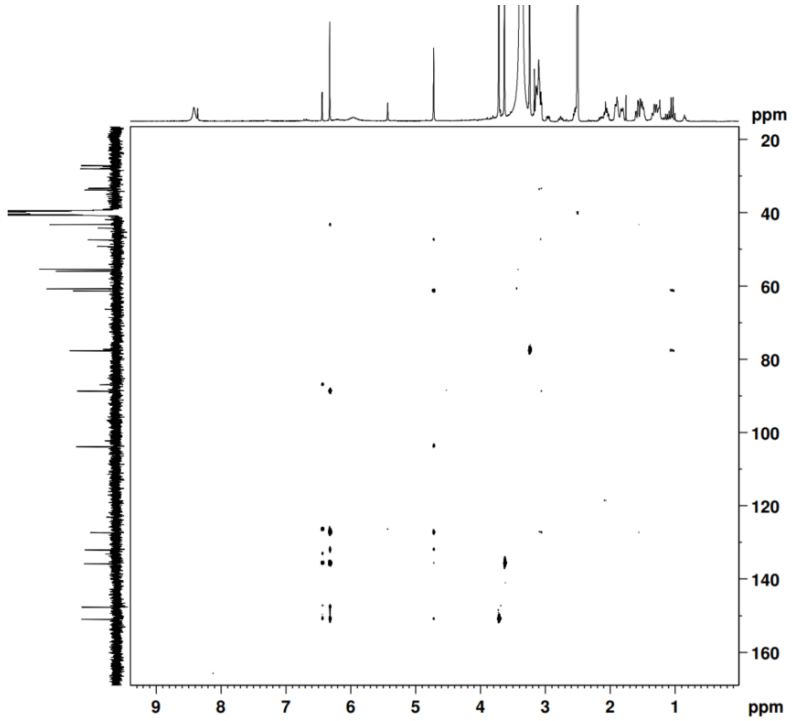


Figure S6. HMBC spectrum of compounds **1** and **2** in $(CD_3)_2SO$

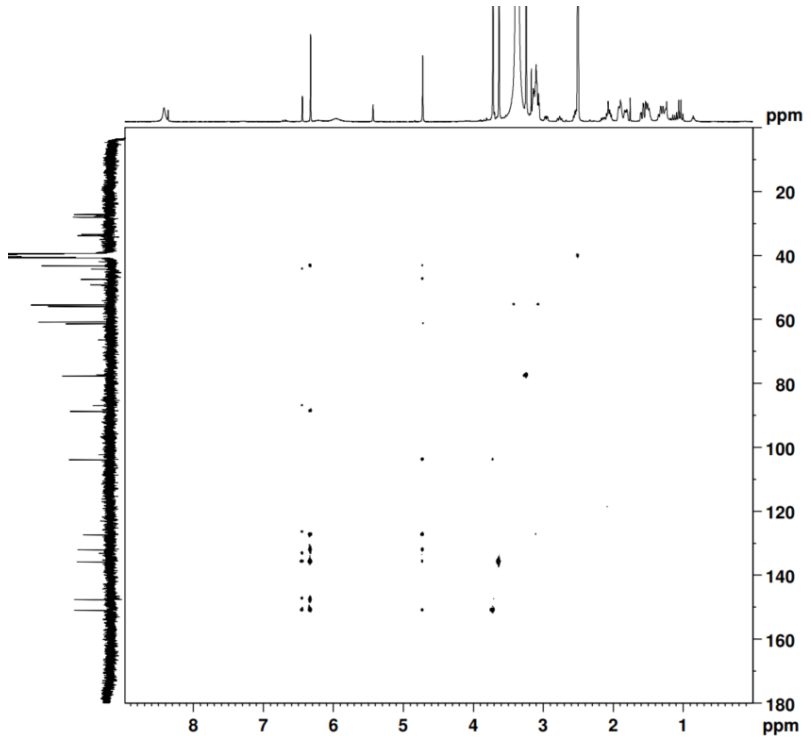


Figure S7. Long-range HMBC spectrum of compounds **1** and **2** in $(CD_3)_2SO$

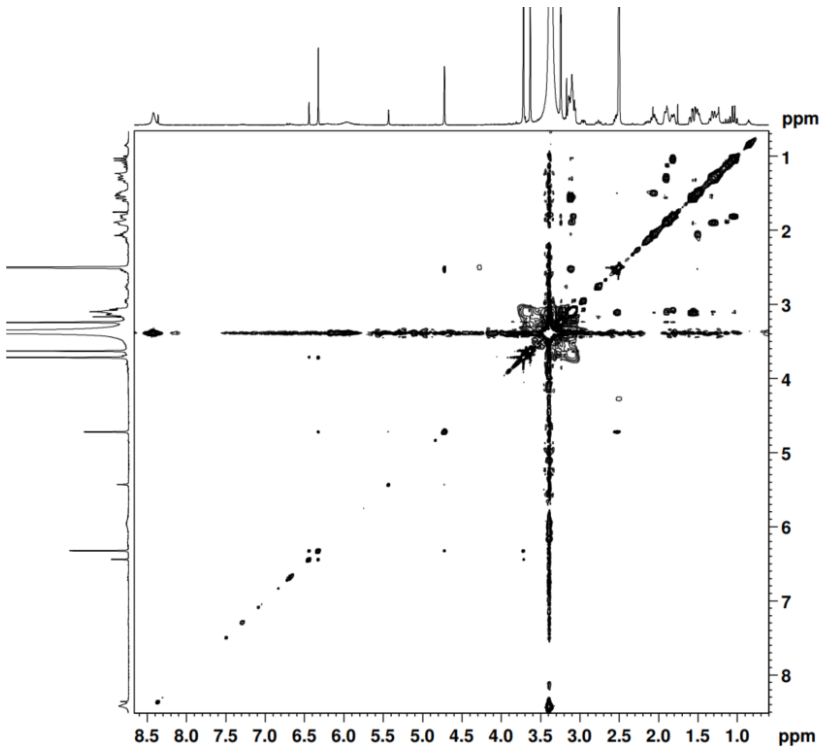


Figure S8. NOESY spectrum of compounds **1** and **2** in $(CD_3)_2SO$

Table S1. Anti-SARS-CoV-2 and cytotoxicity screening results on Vero E6-EGFP cells

Compounds		**IC_{50} (µM)**		**CC_{50} (µM)**	**av. IC_{50} (µM)**	**av. CC_{50} (µM)**	**SI**
1	=	54	>	100	45	>100	>2
	=	37	>	100			
2 and 3	>	100	>	100	>100	>100	X1
	>	100	>	100			
4 and 5	=	44	>	100	44	>100	>2
	=	45	>	100			
6		n.a.	=	1.28	n.a.	1.21	n.a.
		n.a.	=	1.14			
7	=	69	>	100	54	>100	>2
	=	39	>	100			
8	=	37	>	100	39	>100	>3
	=	42	>	100			
9	>	100	>	100	>100	>100	X1
	>	100	>	100			
10		n.a.	=	0.14	n.a.	0.13	n.a.
		n.a.	=	0.12			
11	=	99	>	100	77	>100	>1
	=	54	>	100			

n.a. = not applicable

3.6. References

1. Newman, D.J.; Cragg, G.M. Natural Products as Sources of New Drugs over the Nearly Four Decades from 01/1981 to 09/2019. *J. Nat. Prod.* **2020**, *83*, 770–803.
2. Lodewyk, M.W.; Siebert, M.R.; Tantillo, D.J. Computational Prediction of 1H and 13C Chemical Shifts: A Useful Tool for Natural Product, Mechanistic, and Synthetic Organic Chemistry. *Chem. Rev.* **2012**, *112*, 1839–1862.
3. Merrill, A.T.; Tantillo, D.J. Solvent Optimization and Conformational Flexibility Effects on 1H and 13C NMR Scaling Factors. *Magn. Reson. Chem.* **2020**, *58*, 576–583.
4. Smith, S.G.; Goodman, J.M. Assigning Stereochemistry to Single Diastereoisomers by GIAO NMR Calculation: The DP4 Probability. *J. Am. Chem. Soc.* **2010**, *132*, 12946–12959.
5. Ermanis, K.; Parkes, K.E.B.; Agback, T.; Goodman, J.M. Doubling the Power of DP4 for Computational Structure Elucidation. *Org. Biomol. Chem.* **2017**, *15*, 8998–9007.
6. Cooper, J.K.; Li, K.; Aubé, J.; Coppage, D.A.; Konopelski, J.P. Application of the DP4 Probability Method to Flexible Cyclic Peptides with Multiple Independent Stereocenters: The True Structure of Cyclocinamide A. *Org. Lett.* **2018**, *20*, 4314–4317.
7. Sarotti, A.M. In Silico Reassignment of (+)-Diplopyrone by NMR Calculations: Use of a DP4/J-DP4/DP4+/DIP Tandem to Revise Both Relative and Absolute Configuration. *J. Org. Chem.* **2020**, *85*, 11566–11570.
8. De Gussem, E.; Tehrani, K.A.; Herrebout, W.A.; Bultinck, P.; Johannessen, C. Comparative Study of the Vibrational Optical Activity Techniques in Structure Elucidation: The Case of Galantamine. *ACS Omega* **2019**, *4*, 14133–14139.
9. Grimblat, N.; Zanardi, M.M.; Sarotti, A.M. Beyond DP4: An Improved Probability for the Stereochemical Assignment of Isomeric Compounds Using Quantum Chemical Calculations of NMR Shifts. *J. Org. Chem.* **2015**, *80*, 12526–12534.
10. Zanardi, M.M.; Sarotti, A.M. Sensitivity Analysis of DP4+ with the Probability Distribution Terms: Development of a Universal and Customizable Method. *J. Org. Chem.* **2021**, *86*, 8544–8548.

11. Marcarino, M.O.; Cicetti, S.; Zanardi, M.M.; Sarotti, A.M. A Critical Review on the Use of DP4+ in the Structural Elucidation of Natural Products: The Good, the Bad and the Ugly. A Practical Guide. *Nat. Prod. Rep.* **2021**, *39*, 58–76.
12. Alexander, H.; Goodman, J.M. The DP5 probability, quantification and visualisation of structural uncertainty in single molecules. *Chem. Sci.* **2022**, *13*, 3507–3518.
13. Howarth, A.; Ermanis, K.; Goodman, J.M. DP4-AI Automated NMR Data Analysis: Straight from Spectrometer to Structure. *Chem. Sci*. **2020**, *11*, 4351–4359.
14. Willoughby, P.H.; Jansma, M.J.; Hoye, T.R. A Guide to Small-Molecule Structure Assignment through Computation of (1 H and 13 C) NMR Chemical Shifts. *Nat. Protoc.* **2014**, *9*, 643–660.
15. Navarro-Vazquez, A.; Gil, R.R.; Blinov, K. Computer-Assisted 3D Structure Elucidation (CASE-3D) of Natural Products Combining Isotropic and Anisotropic NMR Parameters. *J. Nat. Prod.* **2018**, *81*, 203–210.
16. Troche-Pesqueira, E.; Anklin, C.; Gil, R.R.; Navarro-Vázquez, A. Computer-Assisted 3D Structure Elucidation of Natural Products Using Residual Dipolar Couplings. *Angew. Chemie Int. Ed.* **2017**, *56*, 3660–3664.
17. Soong, R.; Pautler, B.G.; Moser, A.; Jenne, A.; Lysak, D.H.; Adamo, A.; Simpson, A.J. *CASE (Computer-Assisted Structure Elucidation) Study for an Undergraduate Organic Chemistry Class*; ACS Publications: Washington, DC, USA, 2020.
18. Nuzillard, J.; Plainchont, B. Tutorial for the Structure Elucidation of Small Molecules by Means of the LSD Software. *Magn. Reson. Chem.* **2018**, *56*, 458–468.
19. Lodewyk, M.W.; Tantillo, D.J. Prediction of the Structure of Nobilisitine A Using Computed NMR Chemical Shifts. *J. Nat. Prod.* **2011**, *74*, 1339–1343.
20. Burns, D.C.; Mazzola, E.P.; Reynolds, W.F. The Role of Computer-Assisted Structure Elucidation (CASE) Programs in the Structure Elucidation of Complex Natural Products. *Nat. Prod. Rep.* **2019**, *36*, 919–933.
21. Berova, N.; Polavarapu, P.L.; Nakanishi, K.; Woody, R.W. *Comprehensive Chiroptical Spectroscopy: Applications in Stereochemical Analysis of Synthetic Compounds, Natural Products, and Biomolecules*; John Wiley & Sons: Hoboken, NJ, USA, 2012; Volume 2.

22. Superchi, S.; Rosini, C.; Mazzeo, G.; Giorgio, E. Determination of Molecular Absolute Configuration: Guidelines for Selecting a Suitable Chiroptical Approach. *Compr. chiroptical Spectrosc.* **2012**, *2*, 421–448.
23. Bohle, F.; Seibert, J.; Grimme, S. Automated Quantum Chemistry-Based Calculation of Optical Rotation for Large Flexible Molecules. *J. Org. Chem.* **2021**, *86*, 15522–15531.
24. Khatri Chhetri, B.; Lavoie, S.; Sweeney-Jones, A.M.; Mojib, N.; Raghavan, V.; Gagaring, K.; Dale, B.; McNamara, C.W.; Soapi, K.; Quave, C.L.; et al. Peyssonnosides A-B, Unusual Diterpene Glycosides with a Sterically Encumbered Cyclopropane Motif: Structure Elucidation Using an Integrated Spectroscopic and Computational Workflow. *J. Org. Chem.* **2019**, *84*, 8531–8541.
25. Batista, J.M.; Blanch, E.W.; Da Silva Bolzani, V. Recent Advances in the Use of Vibrational Chiroptical Spectroscopic Methods for Stereochemical Characterization of Natural Products. *Nat. Prod. Rep.* **2015**, *32*, 1280–1302.
26. Ren, J.; Zhao, D.; Wu, S.-J.; Wang, J.; Jia, Y.-J.; Li, W.-X.; Zhu, H.-J.; Cao, F.; Li, W.; Pittman, C.U. Reassigning the Stereochemistry of Bioactive Cepharanthine Using Calculated versus Experimental Chiroptical Spectroscopies. *Tetrahedron* **2019**, *75*, 1194–1202.
27. Pettit, G.R.; Eastham, S.A.; Melody, N.; Orr, B.; Herald, D.L.; McGregor, J.; Knight, J.C.; Doubek, D.L.; Pettit, G.R.; Garner, L.C.; et al. Isolation and Structural Modification of 7-Deoxynarciclasine and 7-Deoxy-Trans-Dihydronarciclasine,1. *J. Nat. Prod.* **2006**, *69*, 7–13.
28. Chhetri, B.K.; Lavoie, S.; Sweeney-Jones, A.M.; Kubanek, J. Recent Trends in the Structural Revision of Natural Products. *Nat. Prod. Rep.* **2018**, *35*, 514–531.
29. Yoo, H.-D.; Nam, S.-J.; Chin, Y.-W.; Kim, M.-S. Misassigned Natural Products and Their Revised Structures. *Arch. Pharm. Res.* **2016**, *39*, 143–153.
30. Paul, D.; Kundu, A.; Saha, S.; Goswami, R.K. Total Synthesis: The Structural Confirmation of Natural Products. *Chem. Commun.* **2021**, *57*, 3307–3322.
31. Le, H.T.N.; Van Roy, E.; Dendooven, E.; Peeters, L.; Theunis, M.; Foubert, K.; Pieters, L.; Tuenter, E. Alkaloids from Lepidium Meyenii (Maca), Structural Revision of Macaridine and UPLC-MS/MS Feature-Based Molecular Networking. *Phytochemistry* **2021**, *190*, 112863.

32. Batista, A.N.L.; Angrisani, B.R.P.; Lima, M.E.D.; Da Silva, S.M.P.; Schettini, V.H.; Chagas, H.A.; dos Santos, F.M., Jr.; Batista, J.M., Jr.; Valverde, A.L. Absolute Configuration Reassignment of Natural Products: An Overview of the Last Decade. *J. Braz. Chem. Soc.* **2021**, *32*, 1499–1518.
33. Baldwin, S.W.; Debenham, J.S. Total Syntheses of (−)-Haemanthidine, (+)-Pretazettine, and (+)-Tazettine. *Org. Lett.* **2000**, *2*, 99–102.
34. Shawky, E.; Hammoda, H.; Abou Donia, A.; Toaima, S.; Kinoshita, E.; Takayama, H. Phytochemical and Biological Investigation of Narcissus Pseudonarcissus Cultivated in Egypt. *Rec. Pharm. Biomed. Sci.* **2018**, *2*, 26–34.
35. Shitara, N.; Hirasawa, Y.; Hasumi, S.; Sasaki, T.; Matsumoto, M.; Wong, C.P.; Kaneda, T.; Asakawa, Y.; Morita, H. Four New Amaryllidaceae Alkaloids from Zephyranthes Candida. *J. Nat. Med.* **2014**, *68*, 610–614.
36. Wagner, J.; Pham, H.L.; Döpke, W. Alkaloids from Hippeastrum Equestre Herb. —5. Circular Dichroism Studies. *Tetrahedron* **1996**, *52*, 6591–6600.
37. Bessa, C.D.P.B.; de Andrade, J.P.; de Oliveira, R.S.; Domingos, E.; Santos, H.; Romão, W.; Bastida, J.; Borges, W.S. Identification of Alkaloids from Hippeastrum Aulicum (Ker Gawl.) Herb.(Amaryllidaceae) Using CGC-MS and Ambient Ionization Mass Spectrometry (PS-MS and LS-MS). *J. Braz. Chem. Soc.* **2017**, *28*, 819–830.
38. Kepplinger, B.; Mardiana, L.; Cowell, J.; Morton-Laing, S.; Dashti, Y.; Wills, C.; Marrs, E.C.L.; Perry, J.D.; Gray, J.; Goodfellow, M.; et al. Discovery, Isolation, Heterologous Expression and Mode-of Action Studies of the Antibiotic Polyketide Tatiomicin from Amycolatopsis sp. DEM30355. *Sci. Rep.* **2022**, *12*, 15579.
39. Polavarapu, P.; Donahue, E.; Shanmugam, G.; Scalmani, G.; Hawkins, E.; Rizzo, C.; Ibnusaud, I.; Thomas, G.; Habel, D.; Sebastian, D. A single chiroptical spectroscopic method may not be able to establish the absolute configurations of diastereomers: Dimethylesters of hibiscus and garcinia acids. *J. Phys. Chem. A* **2011**, *115*, 5665–5673.
40. Prasad, P.L.; Covington, C.L. Comparison of experimental and calculated chiroptical spectra for chiral molecular structure determination. *Chirality* **2014**, *26*, 539–552.
41. Polavarapu Prasad, L. Molecular structure determination using chiroptical spectroscopy: Where we may go wrong? *Chirality* **2012**, *24*, 909–920.

42. Polavarapu Prasad, L.; Santoro, E. Vibrational optical activity for structural characterization of natural products. *Nat. Prod. Rep.* **2020**, *37*, 1661–1699.
43. Pescitelli, G.; Bruhn, T. Good Computational Practice in the Assignment of Absolute Configurations by TDDFT Calculations of ECD Spectra. *Chirality* **2016**, *28*, 466–474.
44. Yu, J.; Cao, Y.; Song, H.; Wang, X.; Yao, S. Calculations of Optical Rotation: Influence of Molecular Structure. *J. Serbian Chem. Soc.* **2012**, *77*, 887–898.
45. Stephens, P.J.; Devlin, F.J.; Cheeseman, J.R.; Frisch, M.J. Calculation of Optical Rotation Using Density Functional Theory. *J. Phys. Chem. A* **2001**, *105*, 5356–5371.

CHAPTER 4

Amaryllidaceae alkaloids from *Hymenocallis littoralis*

Publication:

Le Ngoc-Thao-Hien, Steven De Jonghe, Kristien Erven, Tom Vermeyen, Aliou M. Baldé, Wouter A. Herrebout, Johan Neyts, Christophe Pannecouque, Luc Pieters, and Emmy Tuenter. 2023b. "Anti-SARS-CoV-2 Activity and Cytotoxicity of Amaryllidaceae Alkaloids from *Hymenocallis littoralis*". Molecules 28, no. 7: 3222.

Abstract:

The Amaryllidaceae species are well-known as a rich source of bioactive compounds in nature. Although *Hymenocallis littoralis* has been studied for decades, its polar components were rarely explored. The current phytochemical investigation of Amaryllidaceae alkaloids from *H. littoralis* led to the identification of three previously undescribed compounds: *O*-demethyl-norlycoramine (1), (−)-2-*epi*-pseudolycorine (2) and (+)-2-*epi*-pseudolycorine (3), together with eight known compounds: 6α-hydroxyhippeastidine (4), 6β-hydroxyhippeastidine (5), lycorine (6), 2-*epi*-lycorine (7), zephyranthine (8), ungeremine (9), pancratistatin (10) and 9-*O*-demethyl-7-*O*-methyllycorenine (11). Among the eight previously reported compounds, five were isolated from *H. littoralis* for the first time (compounds 4, 5, 7, 8, and 9). Compounds 1, 4, 5, 7, 8, and 11 exhibited weak anti-SARS-CoV-2 activity (EC_{50} = 40–77 µM) at non-cytotoxic concentrations. Assessment of cytotoxicity on the Vero-E6 cell line revealed lycorine and pancratistatin as cytotoxic substances with CC_{50} values of 1.2 µM and 0.13 µM, respectively. The preliminary structure-activity relationship for the lycorine-type alkaloids in this study was further investigated, and as a result ring C appears to play a crucial role in their anti-SARS-CoV-2 activity.

Keywords: Amaryllidaceae alkaloids; *Hymenocallis littoralis*; lycorine-type; galanthamine-type; DFT calculation; SARS-CoV-2

4.1. Introduction

Amaryllidaceae alkaloids (AAs) belong to the large group of isoquinoline alkaloids, and they are almost exclusively found in the Amaryllidaceae family [1,2]. These alkaloids show a wide variety of pharmacological activities, such as anticancer, antiplasmodial, anti-inflammatory, antioxidant, etc. One of them, galanthamine, is a worldwide registered drug for the treatment of cognitive decline in mild to moderate Alzheimer's disease and various other memory impairments [3,4]. To date, over 600 Amaryllidaceae alkaloids have been reported up to the end of 2018, and the number is still increasing [5].

Hymenocallis littoralis (Jacq.) Salisb. or 'spider lily' (Amaryllidaceae) is widely distributed by the sea and in swamps in tropical, sub-tropical and temperate climates throughout the world [6]. Alongside its medicinal value, it is also known for its horticultural and ornamental appeal [7]. This plant species has been the subject of research for decades and up to now, to our knowledge, 37 alkaloids have been isolated from *H. littoralis*, including four lycorine-type alkaloids (lycorine, littoraline, diacetyllycorine, norpluviine), four lycorenine-type alkaloids (lycorenine, homolycorine, hippeastrine, 6-*O*-methyllycorenine), seven haemanthamine- and crinine-type alkaloids (haemanthamine, vittatine, 11-hydroxyvittatine, crinine, *O*-acetyldihydrocrinine, 8-*O*-demethylmaritidine, bowdensine), five pretazettine-type alkaloids (tazettine, pretazettine, macronine, tazettamide, hymenolitatine G), five phenanthridone-type alkaloids (pancratistatin, narciclasine, lycoricidine, 7-deoxy-*trans*-dihydronarciclasine, isocarbostyril), two phenanthridine-type alkaloids (trisphaeridine, 5,6-dihydrobicolorine), one galanthamine-type alkaloid (lycoramine), five plicamine-type (6-oxo-5,6-dihydroplicane; 5,6-dihydroplicane, hymenolitatines A-C), two secoplicamine-type (hymenolitatines D and E), one belladine-type (hymenolitatine F), and an unclassified alkaloid (hymenolitatine) [7–13]. The most recent study of *H. littoralis* was published by the authors concerning the application of DFT calculations for structure elucidation of an epimeric mixture of 6α-hydroxyhippeastidine and 6β-hydroxyhippeastidine [14].

A brief overview of some pharmacological studies on *H. littoralis* is summarized hereafter. Lin et al. (1995) first reported that the alkaloidal extract of *H. littoralis*

showed in vitro cytotoxic activity and further investigations of the phytochemical constituents proved that two out of their 14 isolated compounds (lycorine and haemanthamine) exhibited cytotoxicity against 11 cultured cell lines [10]. Afterwards, several studies conducted in the period 2005-2016 showed that the phenanthridone-type alkaloids of this plant species, including narciclasine, lycoricidine, and especially pancratistatin, exhibited potent selective toxicity against human tumor cells, as well as antiparasitic activity [15–17]. In 2014, Ji et al. studied the anti-tumor activity of a mixture of three alkaloids isolated from *H. littoralis* (5,6-dihydrobicolorine, 7-deoxy-*trans*-dihydronarciclasine and littoraline), pointing out a possible mode of action for apoptosis induction through the Fas-signaling pathway [18]. In 2016, Chen et al. reported a moderate in vitro cytotoxic activity of hymenolitatine against four cell lines [12]. Lastly, antiproliferation was further assessed by Ma et al. revealing jonquailine, 6-*O*-methylpretazettine, and hymenolitatine F as moderate antiproliferative agents [13].

In the present study, three unreported Amaryllidaceae alkaloids were identified, including two lycorine-type and one galanthamine-type alkaloids, alongside eight known alkaloids from *H. littoralis*. The chemical structures of all compounds were elucidated based on extensive 1D- and 2D-NMR spectroscopy and HRMS (High Resolution Mass Spectrometry) data. Propositions of relative configurations were based on *J*-coupling and nuclear Overhauser effects (NOE). Comparison between experimental and calculated optical rotation has benefited the determination of absolute configuration recently [19], and this technique was also applied in this study. At last, the anti-SARS-CoV-2 activity and cytotoxicity were assessed in vitro for all purified compounds to shed some light on the applicability of Amaryllidaceae alkaloids in SARS-CoV-2 treatment. Results obtained were also compared with recently published data of related compounds.

4.2. Results and Discussion

4.2.1. Structure Elucidation

The current phytochemical study of *H. littoralis* resulted in three unreported compounds, which were all isolated from the *n*-BuOH extract, alongside eight known compounds (see **Figure 1**).

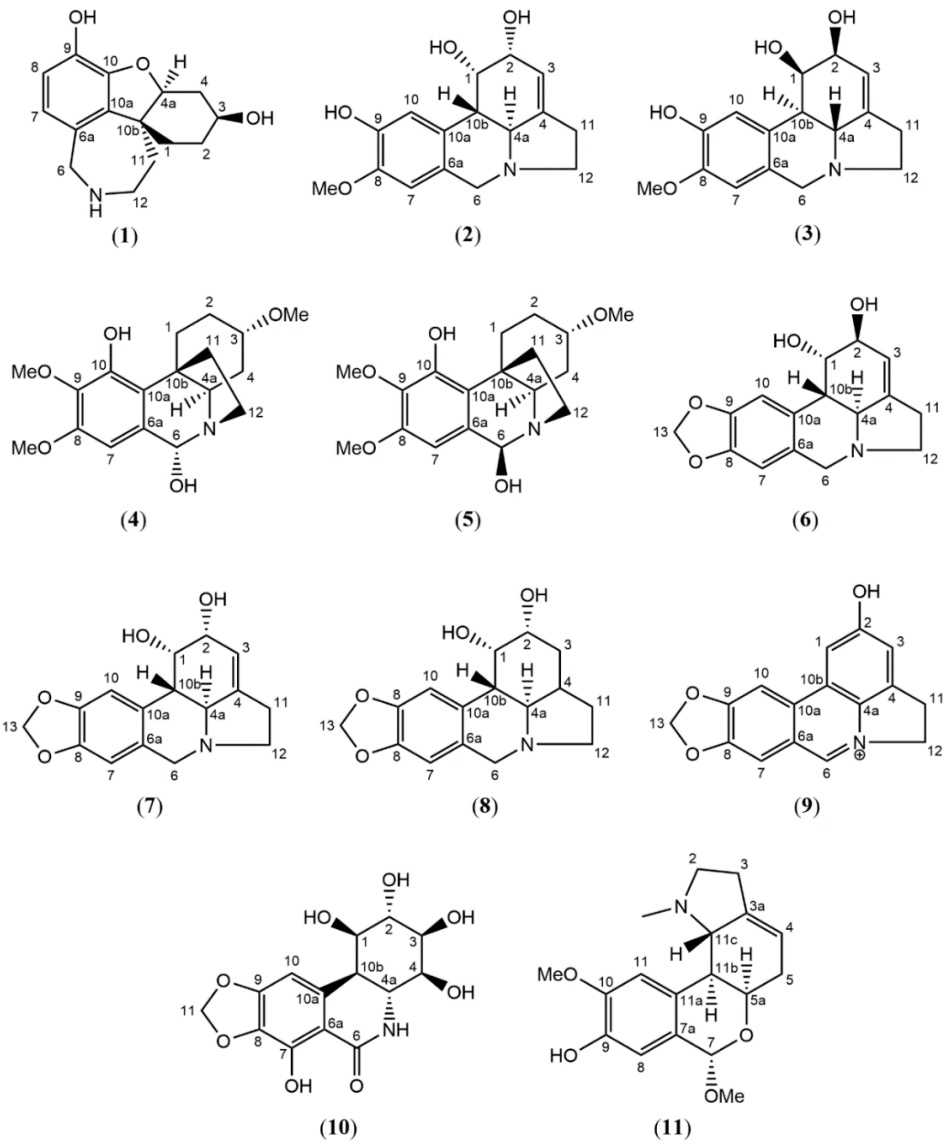


Figure 1. Structures of compounds **1**–**11** isolated from the bulbs of *Hymenocallis littoralis.*

One of the common ^{1}H NMR spectroscopic patterns for Amaryllidaceae alkaloids is the presence of a methylenedioxy proton resonance around 5.95 ppm (-O-CH_2-O- signal) [20,21]. In the absence of this characteristic signal, substituents such as hydroxy- or methoxy groups are prevalent instead of the methylenedioxy moiety. Using this NMR pattern as a starting point, five compounds (**6**–**10**) were found to possess the methylenedioxy moiety, and five compounds (**2**–**5**, **11**) were found to

bear other substituents (Figure 1). The detailed structure elucidation procedure of the three unreported compounds (**1**–**3**) is described hereafter, and the NMR data of known compounds can be consulted in the Supporting Information.

Compound **1** was obtained as an off-white and amorphous solid. HRMS-based elemental analysis suggested a chemical formula of $C_{15}H_{19}NO_3$ (*m*/*z* 262.1446 $[M+H]^+$, calculated for $C_{15}H_{20}NO_3$). Its ^1H NMR spectrum showed two ortho-oriented aromatic hydrogens (δ_H 6.51 and δ_H 6.55, *J* = 8.0 Hz), which are typical for the galanthamine-type skeleton [21]. Indeed, the ^1H and ^{13}C NMR spectra were similar to those reported for galanthamine, but two clear differences could be observed: (1)—Compound **1** does not have two olefinic protons around 6 ppm like galanthamine, implying the absence of the double bond in the structure of galanthamine, and (2)—Compound **1** does not show any signals typical for -OCH_3 or -NCH_3 groups, indicating the absence of the methoxy and the N-methyl groups in the structure of galanthamine. Therefore, compound **1** was identified as *O*-demethyl-norlycoramine (or 1,2-dihydro-*O*-demethyl-*N*-demethylgalanthamine). Correlations of the 2D NMR spectra were in agreement with the proposed structure (Figure 2). Four methylene groups resonated in the upfield region (seven protons in the range of 1.4–2.0 ppm). In COSY, correlations between H-1 and H-2, H-2 and H-3, H-3 and H-4, and H-4 and H-4a confirmed that they are part of the same spin system (see Figure 2). Similarly, another spin system was observed for H-11 and H-12 on COSY. These COSY correlations were crucial to correctly assign the positions of the overlapping proton signals (see Table 1). Assignment of H-7 was confirmed by the NOE correlations between H-7 and H-6. Key HMBC correlations listed as follows aided in confirming the 2D structure (Figure 2): C-10b with H-2, H-4a, H-12; C-1 with H-4a; C-12 with H-6; C-7 with H-6; C-9 with H-7; C-10a with H-7. Finally, the molecular formula $C_{15}H_{19}NO_3$ was confirmed by HRMS-based elemental analysis. Remarkably, a search based on the chemical structure of this compound gave an exact match on Scifinder (checked on 24 March 2023), but not a single reference reporting this compound could be found. Thus, to the best of our knowledge, compound **1** was reported for the first time as a natural product in this study. Biosynthetically, since the substituent of C-9 is a phenolic hydroxyl instead of a methoxy, compound **1** could originate directly from a para-ortho'

oxidative phenol coupling of norbelladine [5,22]. Enzymes responsible for methylation might not be involved in its biosynthetic pathway (Figure 3).

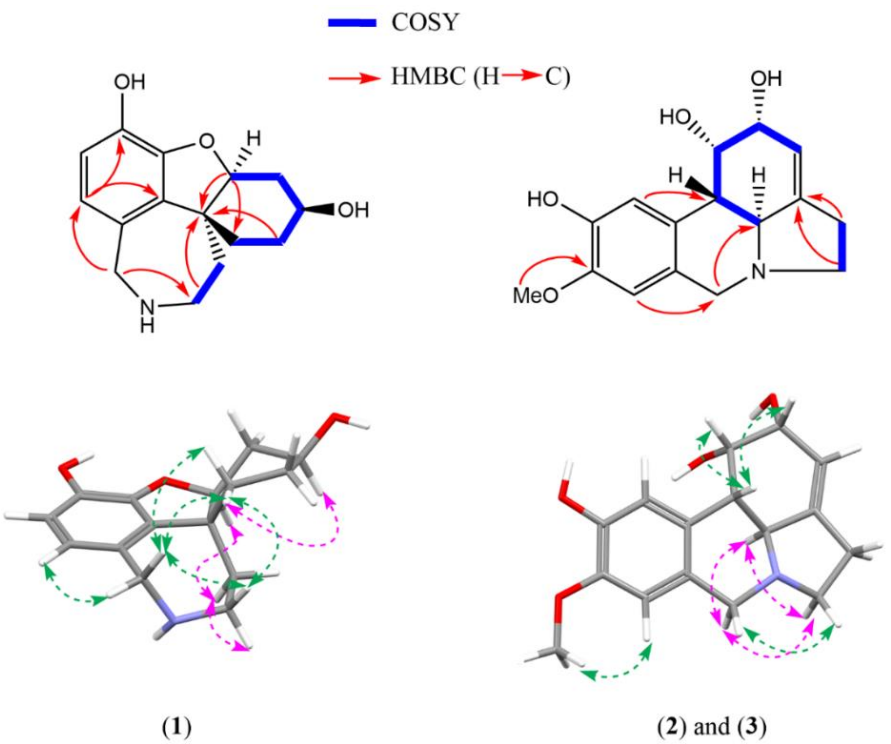


Figure 2. Key HMBC and COSY correlations for compounds **1**–**3** and selected NOESY correlations for compounds **1**–**3**. Green arrows: NOESY correlations above the plane (β-orientation); Pink arrows: NOESY correlations behind the plane (α-orientation).

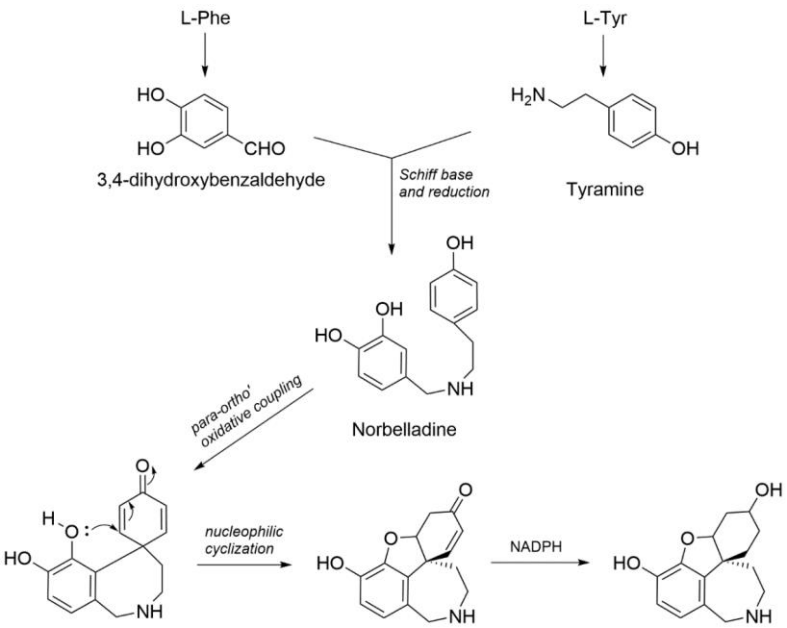


Figure 3. Proposed biosynthetic scheme of compound **1.**

Table 1. Experimental and computed NMR spectroscopic data (400 MHz, CD_3OD) for *O*-demethyl-norlycoramine (**1**).

	Experimental		**Calculated**	
Position	**δ_C, type**	**δ_H (*J* in Hz)**	**δ_C**	**δ_H**
1	24.4, CH_2	1.67 m	25.1	1.53
1	24.4, CH_2	1.85 *	25.1	2.11
2	26.9, CH_2	1.81 *	26.9	2.23
2	26.9, CH_2	1.45 m	26.9	1.15
3	64.7, CH	4.09 m	64.2	3.91
4	31.2, CH_2	2.00 *	29.2	2.08
4	31.2, CH_2	2.32 m	29.2	2.33
4a	88.9, CH	4.31 t (3.2)	90.2	4.50
6	52.7, CH_2	3.83 d (15.0)	50.3	3.71
6	52.7, CH_2	3.97 d (15.0)	50.3	3.95
6a	128.9, C	-	135.1	-
7	120.6, CH	6.51 d (8.0)	120.0	6.52
8	114.9, CH	6.55 d (8.0)	112.2	6.57

9	141.1, C	-	140.5	-
10	146.0, C	-	144.5	-
10a	135.5, C	-	134.5	-
10b	47.2, C	-	49.2	-
11	36.5, CH_2	1.75 *	39.7	1.71
11	36.5, CH_2	1.98 *	39.7	1.83
12	46.5, CH_2	3.26 m	44.4	3.20
12	46.5, CH_2	3.19 m	44.4	3.19
CMAE			1.8	0.13
Max. outlier			6.2	0.42

*overlapping signals

Owing to the stereocenter at C-10b, assigning α- and β-orientation was implausible, like compounds **2** and **3** (see Figure 2). The NOESY spectrum of compound **1** showed strong cross-peaks between δ_H-1.67, δ_H-1.85, δ_H-3.97 and δ_H-3.19, indicating their spatial proximity. Next, the NOE effect observed between H-4a (δ_H-4.31) and δ_H-1.75 implied that they are spatially close to each other. Another important NOE effect was observed between δ_H-1.67 and H-3 (δ_H-4.09), indicating δ_H-1.67, δ_H-4.09, δ_H-4.31 and δ_H-1.75, and δ_H-3.26 are co-facial (see Table 1 and Figure 2). Finally, the configuration of the 3-OH was assigned on the same side with δ_H-1.67 due to a weak NOESY cross-peak between them. Combining all data together, the relative configuration of compound **1** was proposed as (3*S*,4a*S*,10b*R*)-*O*-demethyl-norlycoramine or its enantiomer.

As a follow-up, DP4+ probability was assessed to compare the closeness between the experimental and calculated chemical shifts of all possible diastereomers. The computed diastereomers were chosen as follows: For compound **1,** the presence of three stereocenters infers eight possible configurations (four pairs of enantiomers). As the DP4+ method can only be used to assign a relative configuration, due to the incapacity of NMR to determine absolute configuration, the chemical shift values of only one member of every pair of enantiomers needed to be computed. As shown in Table 2, the highest probability was found for diastereomer (4a*S*,3*S*,10b*R*), which is in line with the configuration deduced from NOE correlations. Therefore, the relative configuration of compound **1** was

confirmed, and the absolute configuration is either (4a*S*,3*S*,10b*R*) or (4a*R*,3*R*,10b*S*). It is important to state that the probability of 73% alone is not conclusive and validation by interatomic distances obtained through the NOESY experiment was required. This practice was recommended by the authors of a recent review on the topic of DP4+ probability [23]. As listed in Table 1, calculated chemical shifts of the diastereomer (4a*S*,3*S*,10b*R*) showed a strong resemblance with the experimental ones. In addition, this ascertained that the proposed planar structure was correct.

Table 2. DP4+ probabilities of compounds **1**–**3** and computed *J*-couplings of compounds **2** and **3** estimated using CD_3OD as solvation model. For compound 1, stereocenters' order is (4a,3,10b). For compounds **2** and **3**, stereocenters' order is (1,2,4a,10b).

Compound 1		**Compounds 2 and 3**		
Diastereomer	**Probability (%)**	**Diastereomer**	**Probability (%)**	$J_{H\text{-}4a/H\text{-}10b}$
SRR	26.78	SRRR	0	10.69
SRS	0	SRRS	61.05	5.52
SSR	73.22	SRSR	0	7.36
SSS	0	SRSS	0	10.56
		SSRR	0	10.25
		SSRS	10.15	5.78
		SSSR	0	7.99
		SSSS	28.80	10.85

Next, specific optical rotations of the two configurations (4a*S*,3*S*,10b*R*) and (4a*R*,3*R*,10b*S*) were computed. Since a pair of enantiomers will have opposite optical activity, the experimental optical rotation was expected to match with only one of the two. According to Table 3, the configuration (4a*S*,3*S*,10b*R*) is levorotatory, and the configuration (4a*R*,3*R*,10b*S*) is dextrorotatory. Apart from the sign, there is a noticeable difference between the magnitudes of the experimental and the computed OR of the configuration (4a*S*,3*S*,10b*R*), which could be due to impurities in the sample (see ^{1}H spectrum in Supplementary Materials). However, since the purity of the sample is approximately 90%, it can be

assumed that compound **1** determines the overall sign of the sample. Therefore, the absolute configuration of compound **1** is defined as (4a*S*,3*S*,10b*R*).

Table 3. Experimental and computed optical rotation values for compounds **1**–**3** ($deg \cdot mL \cdot g^{-1} \cdot dm^{-1}$).

	Compound 1		**Compounds 2 and 3**	
Experimental	**−63.6**		**−8.0**	
B3LYP/6-31++G(d,p)//	SSR	−48.8	SRSS	−230.4
6-311++G(3df,2dp)	RRS	48.7	RSRR	230.6
B3LYP/6-31++G(d,p)//	SSR	−46.8	SRSS	−227.7
aug-cc-pVTZ	RRS	46.6	RSRR	227.9

Compounds **2** and **3** were obtained as a yellowish solid mixture. A chemical formula of $C_{16}H_{19}NO_4$ was deduced from the *m/z* 290.1392 $[M+H]^+$ (calculated for $C_{16}H_{20}NO_4$). The NMR spectra were initially recorded in CD_3OD and showed the general characteristics of lycorine-type alkaloids [21]. In the downfield region, two singlets corresponding to the para-oriented aromatic protons appeared at δ_H greater than 6.0 ppm (H-10/6.88 ppm and H-7/6.76 ppm), together with a broad singlet at δ_H 5-6 ppm (H-3/5.59 ppm), which represented the olefinic proton attached to C-3. Furthermore, similar to many lycorine-type alkaloids, compounds **2** and **3** also showed two doublets corresponding to an AB system of the benzylic methylene at C-6 (H-6α and H-6β, *J* = 14.1 Hz) [27]. Taking into consideration the fact that compounds **2** and **3** do not have a methylenedioxy moiety as discussed above, and the fact that a single methoxy-group (δ_H 3.84 and δ_C 56.5) is present, compounds **2** and **3** appear to have one hydroxy- and one methoxy-group in positions 8 and 9, instead of the methylenedioxy moiety. The position of the methoxy-group was deduced from a NOESY cross-peak between δ_H 3.84 and H-7 (δ_H 6.76) and an HMBC correlation between δ_H 3.84 and δ_C 147.8 (Figure 2). Another two key HMBC correlations were between H-7 and C-6 and between H-10 and C-10b. Furthermore, the COSY spectrum showed strong cross-peaks between three couples of protons: H-2 and H-1, H-1 and H-10b, and H-10b and H-4a (Figure 2). Taking these data together and by comparison with reported NMR data of

lycorine-type alkaloids, it was found that the planar structures of compounds **2** and **3** were identical to the structure of pseudolycorine. This was confirmed by the molecular formula $C_{16}H_{19}NO_4$ obtained by HRMS-based elemental analysis. However, several hydrogens of compounds **2** and **3** at the upfield region (2.50–4.00 ppm) were more deshielded than those of pseudolycorine (Figure S19), which could be indicative for a differing stereochemistry.

DP4+ probability was examined thereafter for compounds **2** and **3** in order to obtain proof for their relative configuration. Nonetheless, no decisive probability (90–100%) was observed. Although the SRRS configuration had the highest probability (61%), the dihedral angle between H-10b and H-4a of this configuration can only produce a $J_{H\text{-}10b/H\text{-}4a}$ of 5.52 Hz (Table 2). Therefore, using the experimental $J_{H\text{-}10b/H\text{-}4a}$ of 11.0 Hz as a constraint, the SRRS configuration was ruled out, and based on the calculated *J*-couplings, SRRR, SRSS, SSRR or SSSS could be the possible configuration (Table 2). Since SSSS is the known configuration of pseudolycorine, only three possibilities remained: SRRR, SRSS and SSRR. Although these three configurations possessed a DP4+ probability of 0% (Table 2), the use of $J_{H\text{-}10b/H\text{-}4a}$ as a restriction in this case is of more reliable since it is well-known that a *trans*-conjunction between two fused rings of lycorine-type skeleton exhibits a typical *J*-coupling of 10–11 Hz.[27] The DP4+ probability, however, only takes into account experimental and calculated chemical shifts as inputs, and the coupling constants are not incorporated in the stereochemistry determination by DP4+ [30]. In addition, the accuracy of DP4+ probability can be affected by many reasons and unsuccessful examples were also reported by the authors of this method [23, 24].

Further analysis of NOE interactions also inferred that compounds **2** and **3** are diastereomers of pseudolycorine. As can be seen in Figure 2, spatial proximities were observed for H-1, H-2 and H-10b in the NOESY spectrum, indicating that these three hydrogens are on the same side of the plane. H-1 and H-2 appeared as two broad singlets with no noticeable splitting, suggesting a *cis*-orientation. In addition, H-10b and H-4a shared a *J*-coupling of 11.1 Hz, typical for the *trans*-orientation between these two hydrogens reported for the lycorine-type scaffold as mentioned above [21]. Therefore, combining NOE correlations and *J*-coupling information, it can be deduced that H-1, H-2 and H-10b were on the same side and

were on the opposite side of H-4a. Among the three remaining possibilities as discussed above, the SRSS is the sole configuration that can show all these NOESY correlations. Hence, the relative configuration of compounds **2** and **3** was established as (1*S*,2*R*,4a*S*,10b*S*)-pseudolycorine, or its enantiomer. The structures of compounds **2** and **3** were elucidated with the configuration of one chiral center (at position H-2) differing from the configuration reported for pseudolycorine, and thus, correspond to 2-*epi*-pseudolycorine. To be noticed, C-4 was not observed in CD_3OD, whereas it could be observed in $(CD_3)_2SO$ in both HMBC and ^{13}C-NMR spectra. Table 4 shows NMR data of compounds **2** and **3** in CD_3OD and $(CD_3)_2SO$.

Table 4. Experimental and computed NMR spectroscopic data (400 MHz, CD_3OD and $(CD_3)_2SO$ for (±)-2-*epi*-pseudolycorine (**2** and **3**).

	In CD_3OD				**In $(CD_3)_2SO$**	
	Experimental		**Calculated**		**Experimental**	
Position	**δ_C, type**	**δ_H (*J* in Hz)**	**δ_C**	**δ_H**	**δ_C**	**δ_H**
1	71.6, CH	4.49 s	75.1	3.94	70.8, CH	4.21 s
2	72.9, CH	4.19 *	71.2	4.31	72.2, CH	3.96 s
3	119.9, CH	5.59 s	126.7	5.62	118.9, CH	5.35 s
4	n.o.	-		-	142.4, C	-
4a	62.6, CH	3.06 d (11.0)	59.9	2.96	61.4, CH	2.58 d (10.6)
6α	57.0, CH_2	4.19 *	56.4	4.03	57.0, CH_2	3.99 d (13.8)
6β	57.0, CH_2	3.71 d (14.1)	56.4	3.57	57.0, CH_2	3.30 d (13.8)
6a	127.2, C	-	129.3	-	127.5, C	-
7	111.7, CH	6.76 s	107.8	6.73	111.6, CH	6.63 s
8	147.8, C	-	145.2	-	146.1, C	-
9	146.8, C	-	145.4	-	145.2, C	-
10	112.7, CH	6.88 s	114.2	7.07	112.4, CH	6.70 s
10a	128.8, C	-	128.5	-	129.1, C	-
10b	40.5, CH	2.76 d (11.1)	41.6	3.05	40.2, CH	2.49 *
11	29.3, CH_2	2.68 * (2H)	31.3	2.70	28.5, CH_2	2.46 * (2H)
12α	54.8, CH_2	3.40 *	53.1	3.38	53.9, CH_2	3.20 t (8.2)
12β	54.8, CH_2	2.67 *	53.1	2.61	53.9, CH_2	2.19 q (8.4)
8-OMe	56.5, CH_3	3.84 s	54.5	3.97	56.2, CH_3	3.71 s

CMAE	2.3	0.13
Max. outlier	6.8	0.55

*overlapping signals. n.o. = not observed.

As reported in Table 3, the experimental optical rotation of 2-*epi*-pseudolycorine (compounds **2** and **3**) was only −8.0 and was significantly different from the computed values, ranging from 227.9 to 230.6, if dextrorotatory, and from −230.4 to −227.7, if levorotatory. Based on these data, it was not possible to deduce the absolute configuration of compounds **2** and **3**. At this point, it was hypothesized that the sample consisted of a racemic mixture. The fact that the measured value is not 0, but −8, could be caused by the presence of some impurities and/or the measurement or weighing errors due to the small amount of material. In order to determine whether the sample consisted of a racemate, an enantioseparation was performed, and the obtained HPLC chromatogram of the mixture of compounds **2** and **3** indeed showed two peaks with almost equal peak areas (Figure S20). Hence, the mixture of compounds **2** and **3** was confirmed to be a racemate. Compounds **2** and **3**, occurring as an enantiomeric mixture, were reported for the first time and the name 2-*epi*-pseudolycorine was adopted. Vibrational circular dichroism (VCD) spectrum of compounds **2** and **3** in $(CD_3)_2SO$ was devoid of any signal, confirming the existence of a racemic mixture (Figure S23). Finally, the assignment of (−)-2-*epi*-pseudolycorine (**2**) to the configuration (1*S*, 2*R*, 4a*S*, 10b*S*) and (+)-2-*epi*-pseudolycorine (**3**) to the configuration (1*R*, 2*S*, 4a*R*, 10b*R*) was made according to their simulated optical rotations (see Table 3).

4.2.2. Anti-SARS-CoV-2 Activity and Cytotoxicity

Compounds **1**, **4**, **5**, **7**, **8,** and **11** (*O*-demethyl-norlycoramine, 6α-hydroxyhippeastidine, 6β-hydroxyhippeastidine, 2-*epi*-lycorine, zephyranthine, and 9-*O*-demethyl-7-*O*-methyllycorenine, respectively) exhibited weak inhibition of SARS-CoV-2 (EC_{50} = 40-77 μM), at non-cytotoxic concentrations (CC_{50} > 100 μM). Compounds **1**, **4**, **5**, and **8** were the most potent ones among all tested compounds, displaying EC_{50} values of 45 μM, 44 μM, and 39 μM, respectively. Compounds **6** and **10** (lycorine and pancratistatin, respectively) lacked selective antiviral activity, since they were cytotoxic for the Vero-E6 cells as evidenced by CC_{50} values of 1.2

μM and 0.13 μM, respectively. Thus, our finding for lycorine is in contrast with the results of Zhang et al., Jin et al. and Ren et al., published in the last two years. Zhang reported lycorine as having potent anti-SARS-CoV-2 activity with an EC_{50} of 0.18 μM and no cytotoxicity (CC_{50} > 40 μM) [25]. Similarly, lycorine displayed a highly potent activity (EC_{50} = 0.878 ± 0.022 μM in Vero cells, CC_{50} > 10 μM on HEK293) according to Jin et al. [26]. Also in 2021, Ren et al. again reported a high potency of lycorine (EC_{50} = 0.439 ± 0.122 μM) in the Vero-E6 cell and varying cytotoxicity values in Vero-E6 (CC_{50} > 1000 μM), Huh-7 (CC_{50} = 0.834 ± 0.0630 μM), HEK293T (CC_{50} = 1.044 ± 0.0734 μM) [27]. Although these antiviral assays also run in Vero-E6 cells, the RT-PCR read-out was after 18-24 h, whereas we used a high-content imaging read-out after 5 days, which might explain the increased cytotoxicity in our assay. Compounds **2**, **3,** and **9**, ((±)-2-*epi*-pseudolycorine and ungeremine, respectively) lacked antiviral activity, as well as cytotoxicity. Detailed results can be found as Supplementary Materials.

4.3. Conclusion

The phytochemical investigation of the bulbs of *H. littoralis* with advanced spectroscopic techniques led to the isolation of eleven Amaryllidaceae alkaloids. Three previously undescribed alkaloids, namely *O*-demethyl-norlycoramine (**1**), (−)-2-*epi*-pseudolycorine (**2**), (+)-2-*epi*-pseudolycorine (**3**), together with eight known alkaloids (**4**–**11**). In addition, 2-epi-lycorine, ungeremine and zephyranthine were reported for the first time in *H. littoralis*.

In vitro anti-SARS-CoV-2 screening revealed weak anti-SARS-CoV-2 activity of *O*-demethyl-norlycoramine, 2-*epi*-lycorine, 6α-hydroxyhippeastidine, 6β-hydroxyhippeastidine, zephyranthine, and 9-*O*-demethyl-7-*O*-methyllycorenine (EC_{50} = 40–77 μM) at non-cytotoxic concentrations (CC_{50} > 100 μM). Lycorine and pancratistatin exhibited cytotoxicity with a CC_{50} of 1.2 μM and 0.13 μM, respectively. Investigation of the structure-activity relationship for the lycorine-type alkaloids in this study suggested a crucial role of ring C: the meaningful ligand-host interactions appeared to be related to ring C; the spatial flexibility of ring C was required to exhibit anti-SARS-CoV-2 activity; the stereochemistry of C-2 on ring C might determine the activity and cytotoxicity. In view of the limited scope of this study, the chirality effect of C-2 was explored, but that of C-1 remains unknown.

Therefore, further research on a bigger library of lycorine-type analogs with a wide chemical diversity is necessary. Besides, the potential of galanthamine-type and crinine-type scaffolds as anti-SARS-CoV-2 agents remains a compelling research topic.

4.4. Materials and Methods

4.4.1. Isolation and Purification

The fresh bulbs of *Hymenocallis littoralis* (Jacq.) Salisb. were collected in April 2017 in Dubréka (Guinea-Conakry) and provided by Prof. A. M. Baldé (University Gamal Abdel Nasser of Conakry, Guinea). The voucher specimen (No. 2aHK2) is kept at the Institute for Research and Development of Medicinal and Food Plants of Guinea, Du-bréka. After cleaning and lyophilization, the total weight of the plant material was 3.3 kg.

Freeze-dried, powdered bulbs were macerated and percolated with approximately 30 L of 80% methanol. The extract was filtered and dried under reduced pressure. Then, liquid-liquid partition was applied in order to obtain subfractions. In short, the crude extract was suspended in water and was acidified with 10% HCl to a pH < 3 before performing liquid-liquid partition with DCM (I). Next, the pH of the acidified phase was increased to a pH ≥ 9 by adding NH_4OH (25%), followed by a subsequent liquid-liquid partition with DCM (II), EtOAc and *n*-BuOH. In this way, four fractions (DCM (I), DCM (II), EtOAc and *n*-BuOH) were obtained and were examined by TLC with the mobile phase DCM-MeOH-NH_4OH (90:10:1). Observation of the TLC plates after spraying with the Dragendorff + $NaNO_2$ 10% reagents indicated that all fractions contained alkaloids, and the majority of alkaloids were present in the DCM (I) fraction.

The DCM (I) (3.1 g) and DCM (II) extracts (3.0 g) were fractionated by flash chromatography. After TLC investigation, silica gel was chosen as the suitable stationary phase. The percentage of NH_4OH was fixed at 0.1% in the solvent systems. Other parameters were set as follows: (1)—Detectors: ELSD and UV (254 and 270 nm), (2)—Flow rate: 25 mL/min. The step-wise gradient used for fractionating the DCM (I) and DCM (II) extracts was: 0–5 min (0% MeOH), 5–15 min (5% MeOH), 15-30 min (10% MeOH), 30–40 min (15% MeOH), 40–45 min (20%

MeOH), 45–50 min (25% MeOH), 55–60 min (70% MeOH), and the gradient was held when a peak was eluted. As for the fractionation of the EtOAc extract (2.2 g), column chromatography was used with a sample/stationary phase (silica gel) ratio of 1/50. Isocratic elution was applied, using DCM-MeOH-NH_4OH (90/10/1, *v*/*v*) as the solvent system, by virtue of the result of TLC analysis. Finally, MeOH was used to elute any remaining compounds of interest from the column. The *n*-BuOH fraction (4.6 g) was also fractionated by column chromatography on a silica gel column. On the basis of TLC investigation, a three-step gradient was chosen to elute compounds from the column: DCM-MeOH-H_2O-NH_4OH (80:20:10:1, *v*/*v*), DCM-MeOH-H_2O-NH_4OH (60:40:10:1, *v*/*v*), and DCM-MeOH-H_2O-NH_4OH (40:60:10:1, *v*/*v*). Finally, MeOH was used to elute any remaining compounds of interest from the column.

Then, purification was carried out using semi-preparative HPLC-DAD-MS, silica gel column chromatography and recrystallization. Compounds **2** and **3** (2.4 mg), **4** (4.0 mg), **5** (4.5 mg), **8** (3.0 mg) and **11** (6.0 mg) were isolated by semi-preparative HPLC, of which compounds **2, 3,** and **8** originated from the *n*-BuOH extract, compounds **4** and **5** from the DCM (I) extract and compound **11** from the DCM (II) extract.

With regard to the recrystallization, in the current work, impure crystals were first collected and were subsequently rinsed with DCM, DCM-MeOH (99:1, *v*/*v*) and DCM-MeOH (95:5, *v*/*v*) to remove impurities. Next, the crystals were dissolved in a minimal volume of DCM-MeOH (50:50, *v*/*v*). Finally, the solution obtained was left to dry in open air, resulting in the formation of solid crystals of the desired constituent. In this way, 1.0 g and 18.0 mg were purified of compounds **6** and **10**, respectively. Compound **6** was found to be the most abundant alkaloid and was present in the DCM (II) and EtOAc extracts, and compound **10** was only purified from the EtOAc extract.

Compounds **1** (5.8 mg), **7** (1.5 mg), and **9** (1.5) were purified by silica gel column chromatography. Compound **1** was isolated from the *n*-BuOH extract and compounds **7** and **9** from the DCM (II) extract. The solvent systems used were deduced from TLC analysis. In the case of compound **1**, the elution was performed with DCM-MeOH-H_2O-NH_4OH (80:20:5:1, *v*/*v*), followed by DCM-MeOH-H_2O-NH_4OH (60:40:5:1, *v*/*v*). In the case of compounds **7** and **9**, the elution was carried

out with DCM-MeOH-NH_4OH (90:10:1, *v*/*v*), followed by DCM-MeOH-NH_4OH (70:30:1, *v*/*v*).

Chiral HPLC separation was performed for compounds **2** and **3** on a Daicel Chiralpak IB column (250 × 4.6 mm) (Chiral Technologies Europe, Iilkirch Cedex, France), eluted with solvent A—MeOH:DEA (100:0.1, *v*/*v*) and solvent B—EtOH:DEA (100:0.1, *v*/*v*) using the following gradient: 0–10 min (50%B), 10–36 min (50–80% B); flow rate 1.0 mL/min.

O-demethyl-norlycoramine (**1**)

White amorphous powder (5.8 mg); UV λ_{max} 210, 254 nm; $[\alpha]_D^{25}$ −63.6 (*c* 0.9, MeOH) and $[\alpha]_D^{25}$ −60.2 (*c* 0.10, MeOH); ^1H and ^{13}C NMR in DMSO-d_6 (400 and 100 MHz): see Table 1; Positive HRESIMS *m/z* 262.1446 $[M+H]^+$ (calcd for $C_{15}H_{20}NO_3$, 262.1443)

(±)-2-*epi*-pseudolycorine (**2** and **3**)

Yellowish amorphous powder (2.4 mg); UV λ_{max} 210, 245, 300 nm; $[\alpha]_D^{25}$ −8.0 (*c* 0.4, MeOH); ^1H and ^{13}C NMR in DMSO-d_6 (400 and 100 MHz): see Table 4; Positive HRESIMS *m/z* 290.1392 $[M+H]^+$ (calcd for $C_{16}H_{20}NO_4$, 290.1392)

4.4.2. Computational Details

Conformational analysis was performed by PCMODEL (version 10.0) using the Monte Carlo algorithm and the MMFF94 force field. All conformers within an energy window of 10 kcal.mol^{-1} were selected. Specific optical rotations were simulated for the sodium D-line wavelength (589.3 nm), as experimental optical rotations were measured at the same wavelength. The obtained conformers were first geometrically optimized in the gas phase. After dereplication, Boltzmann weighting was applied using the sum of electronic and thermal free energies at 298.15 K, and only conformers with energies within an energy window of 2.5 kcal.mol^{-1} from the global minimum were considered as the contributing ones.

For DP4+ probability, the B3LYP/6-31G(d)//mPW1PW91/6-311+G(d,p) level was applied to simulate shielding tensors [30]. On the other hand, two theory levels

were used to compute specific optical rotation: B3LYP/6-31++G(d,p)//6-311++G(3df,2dp) and B3LYP/6-31++G(d,p)//aug-cc-pVTZ, as recommended by Yu (2012) and Stephens (2001), respectively [31,32]. The polarizable continuum model using methanol as solvent was also performed to improve the accuracy of shielding tensors and OR calculations. Coupling constants were simulated in gas-phase at the B3LYP/6-31G(d)//B3LYP/6-31G(d,p) level reported by Bally and Rablen which considers only the Fermi contact as the main contributor [33]. All the above quantum chemical calculations were performed by Gaussian16. Detailed information about the results of DP4+ probability and 3D coordinates of computed conformers can be found as Supplementary Materials.

4.5. Supporting Information – Chapter 4

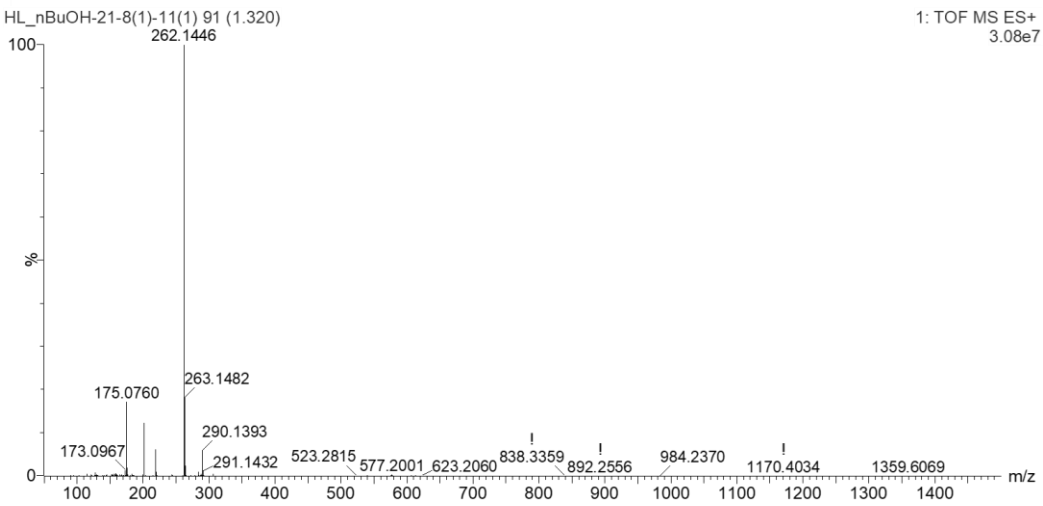


Figure S1. HRMS spectrum of compound **1**

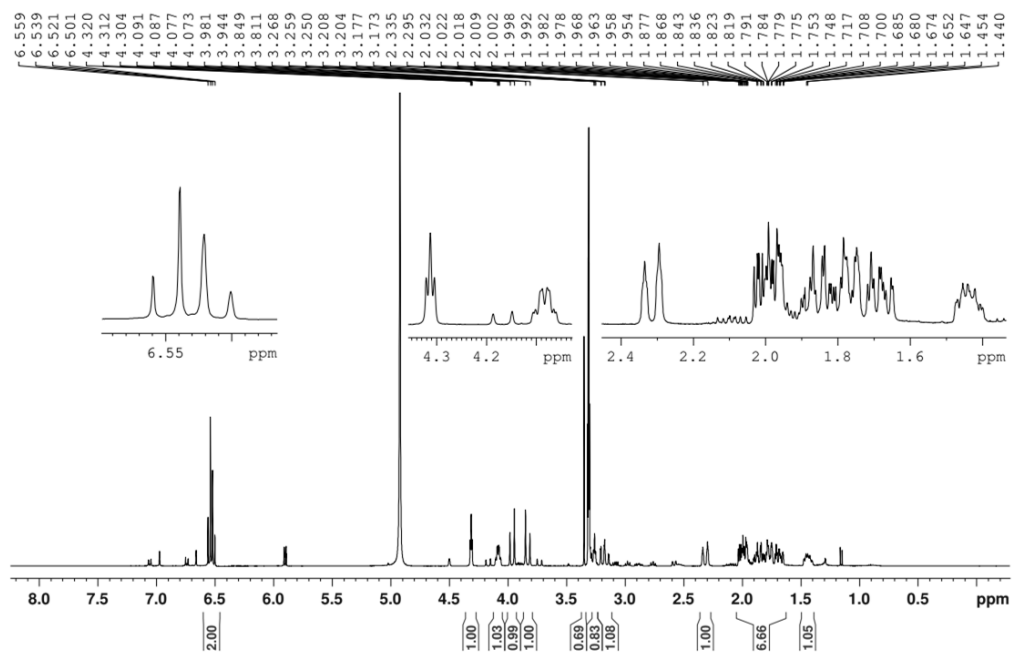


Figure S2. ^{1}H spectrum of compound **1** in CD_3OD

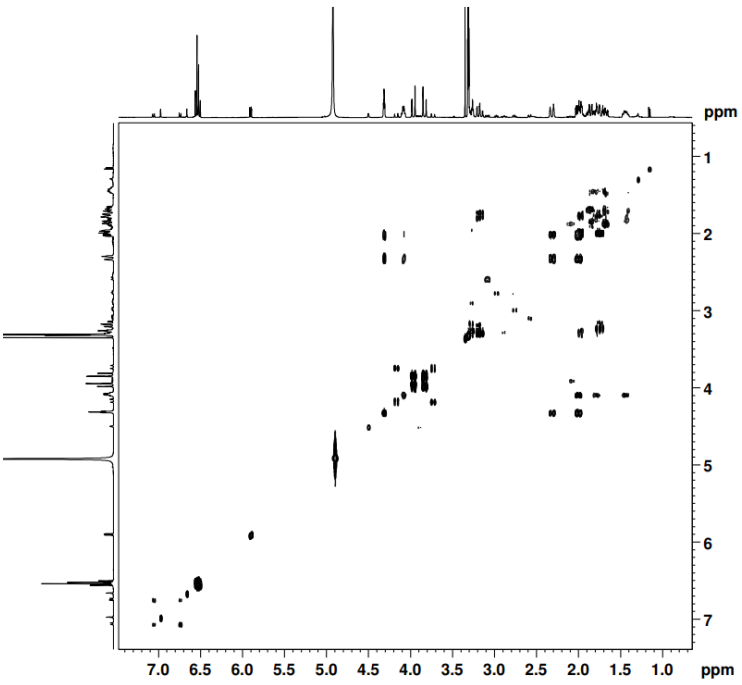


Figure S2. COSY spectrum of compound **1** in CD_3OD

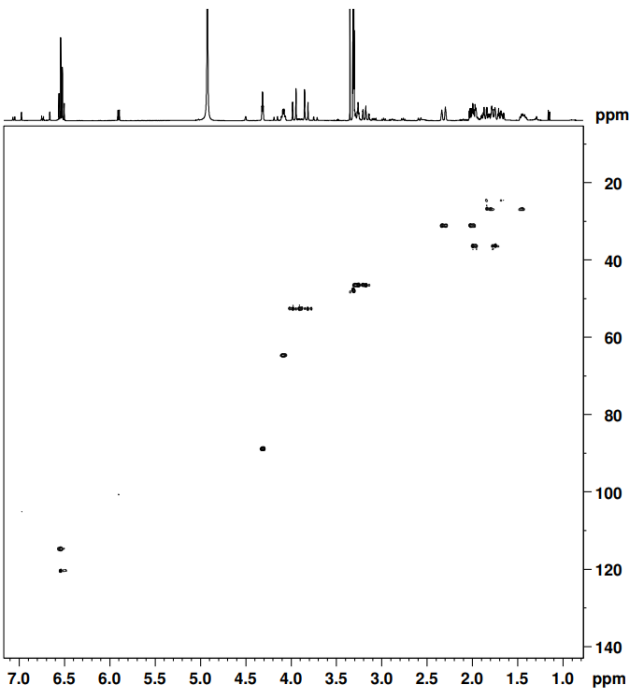


Figure S4. HSQC spectrum of compound **1** in CD_3OD

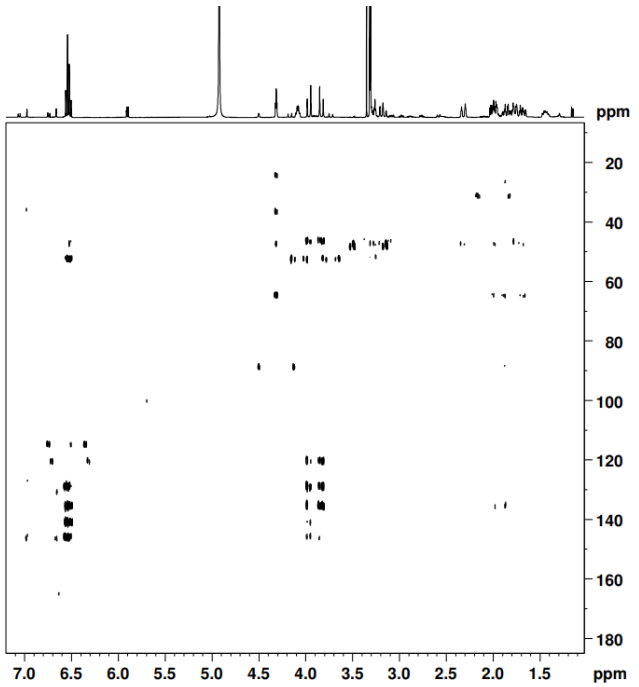


Figure S5. HMBC spectrum of compound **1** in CD_3OD

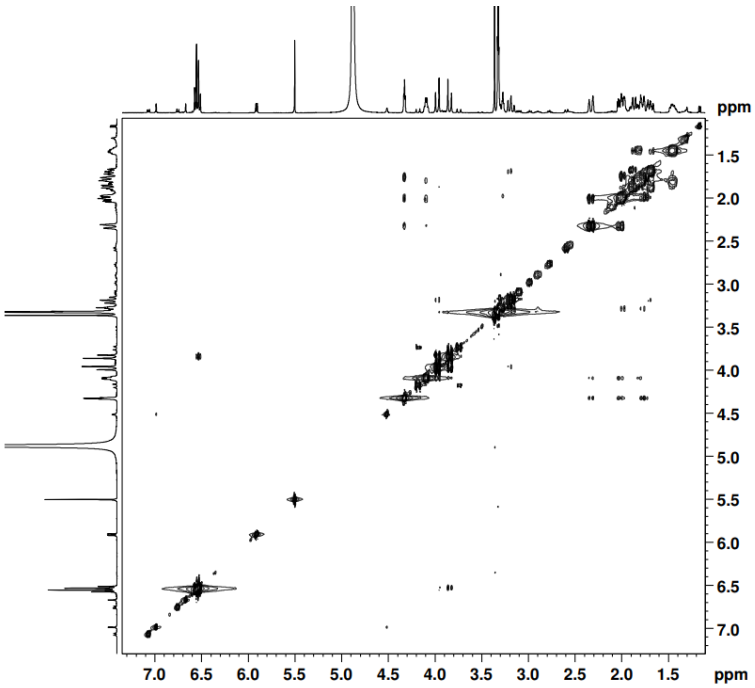


Figure S6. NOESY spectrum of compound **1** in CD_3OD

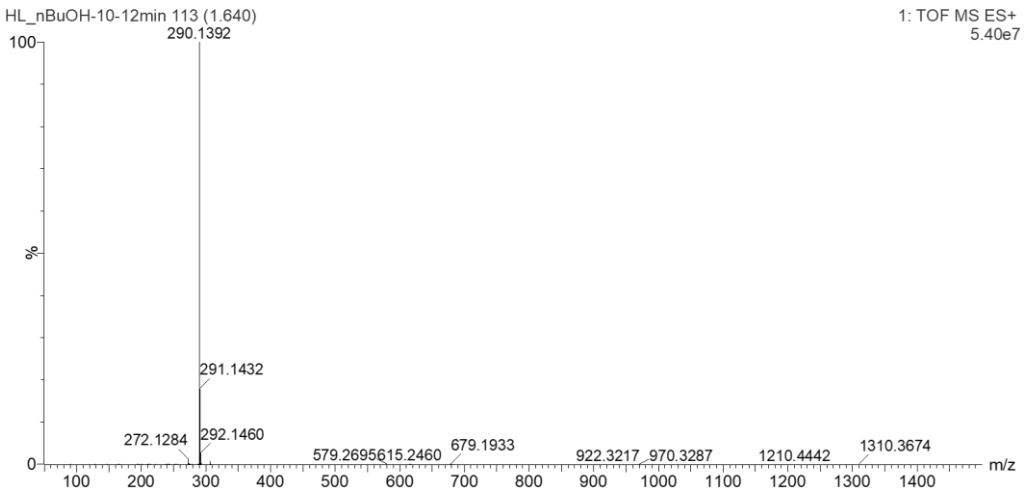


Figure S7. HRMS spectrum of compounds **2** and **3**

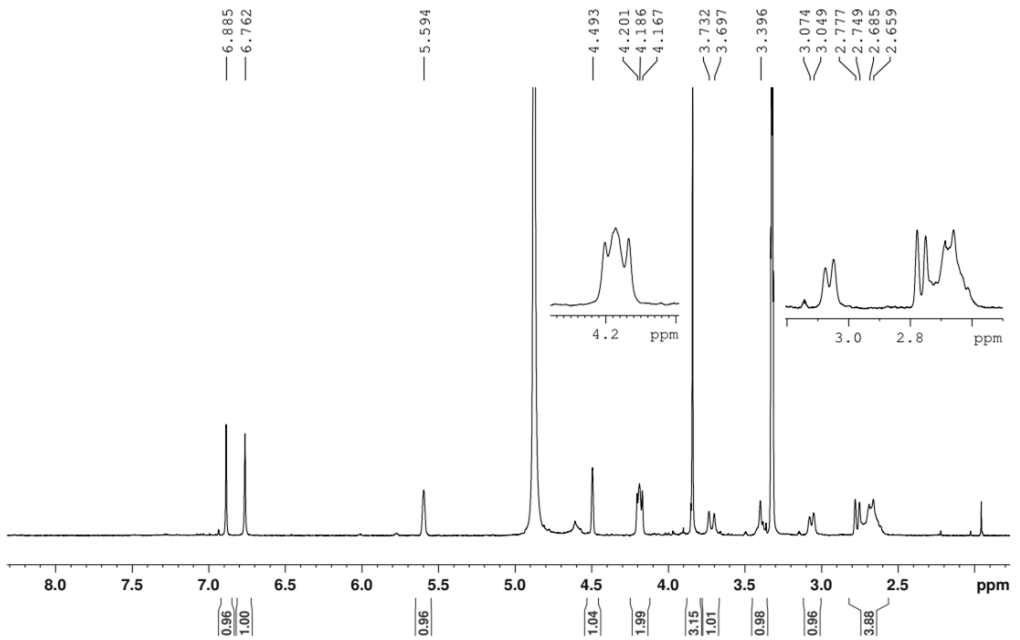


Figure S8. ^{1}H spectrum of compounds **2** and **3** in CD_3OD

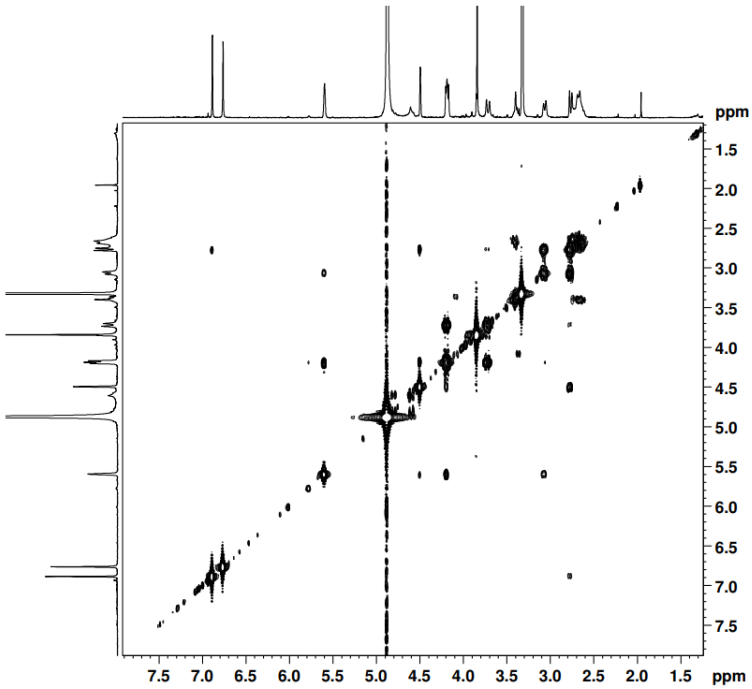


Figure S9. COSY spectrum of compounds **2** and **3** in CD_3OD

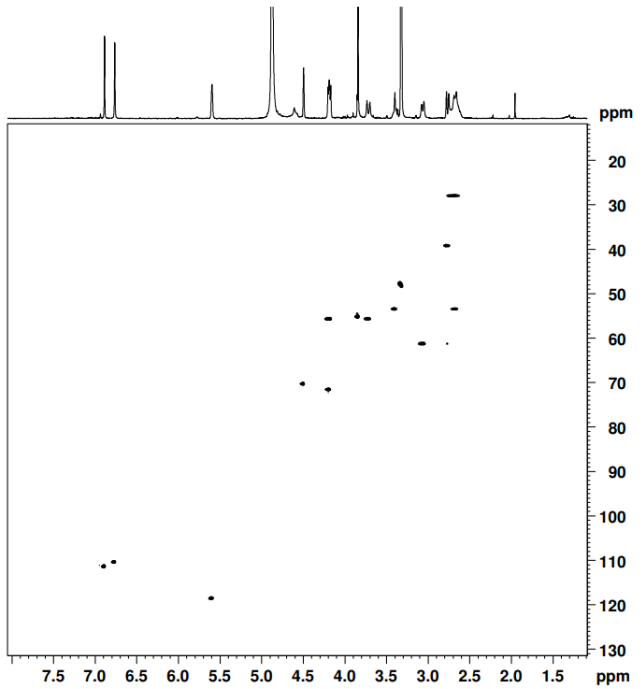


Figure S10. HSQC spectrum of compounds **2** and **3** in CD_3OD

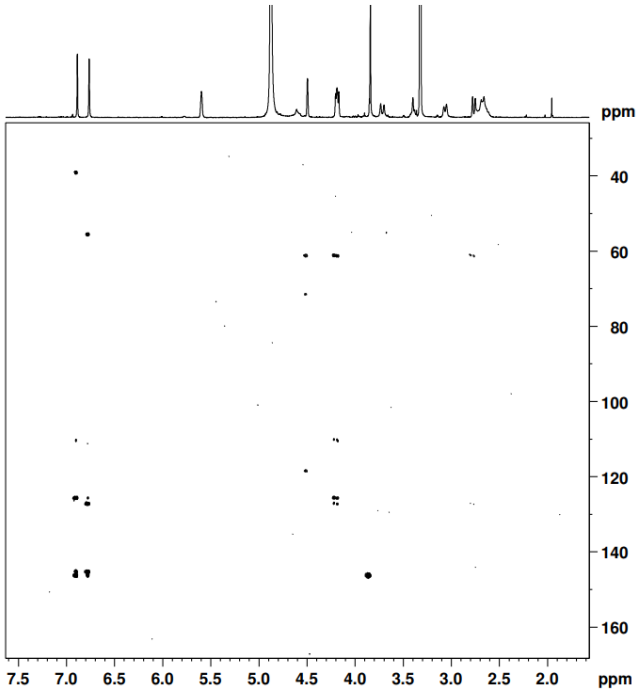


Figure S11. HMBC spectrum of compounds **2** and **3** in CD_3OD

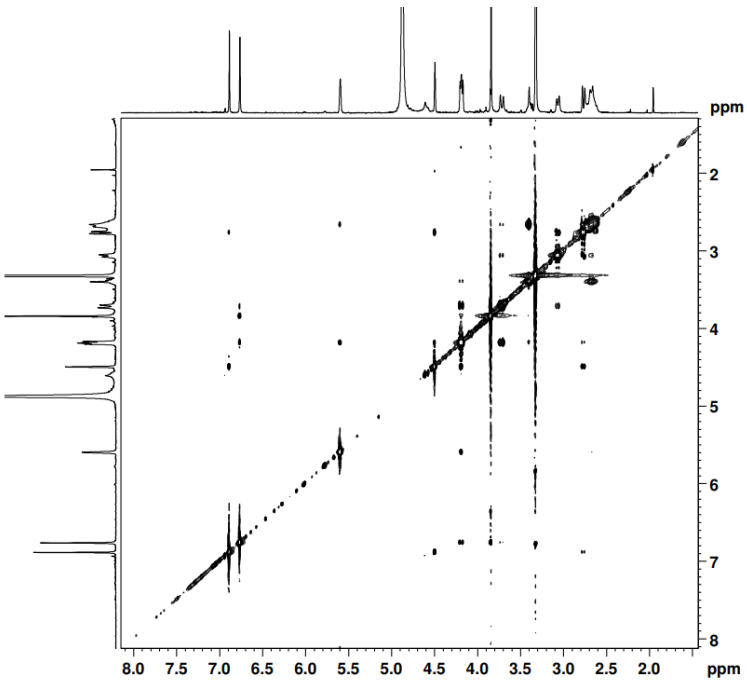


Figure S12. NOESY spectrum of compounds **2** and **3** in CD_3OD

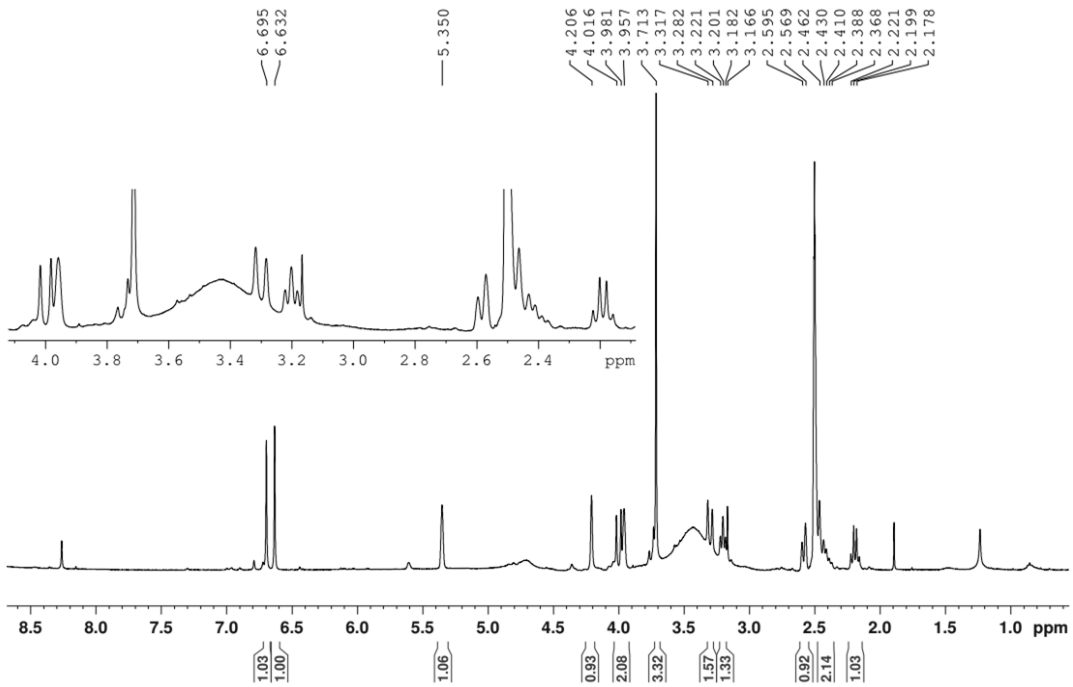


Figure S13. ^{1}H spectrum of compounds **2** and **3** in $(CD_3)_2SO$

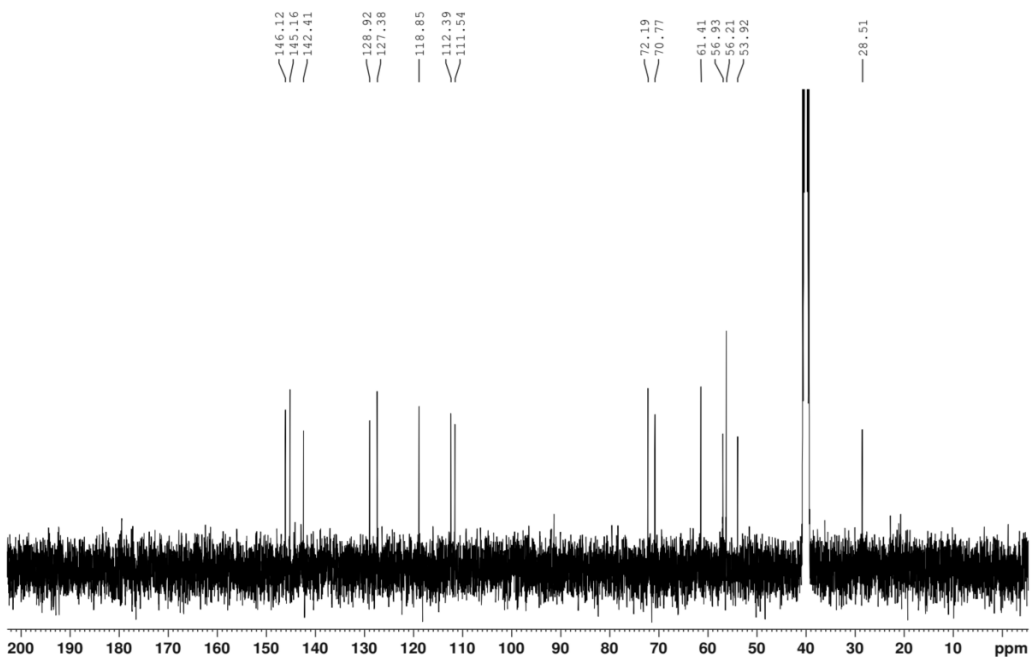


Figure S14. ^{13}C spectrum of compounds **2** and **3** in $(CD_3)_2SO$

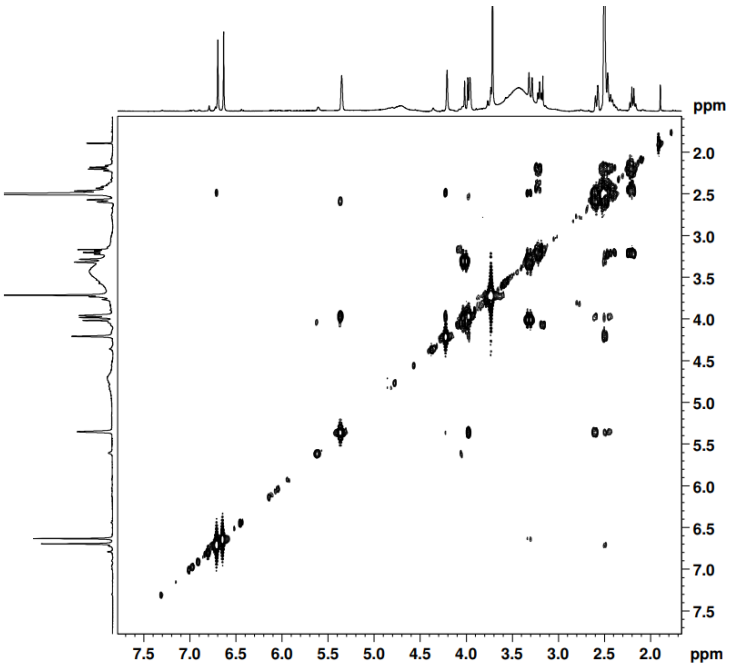


Figure S15. COSY spectrum of compounds **2** and **3** in $(CD_3)_2SO$

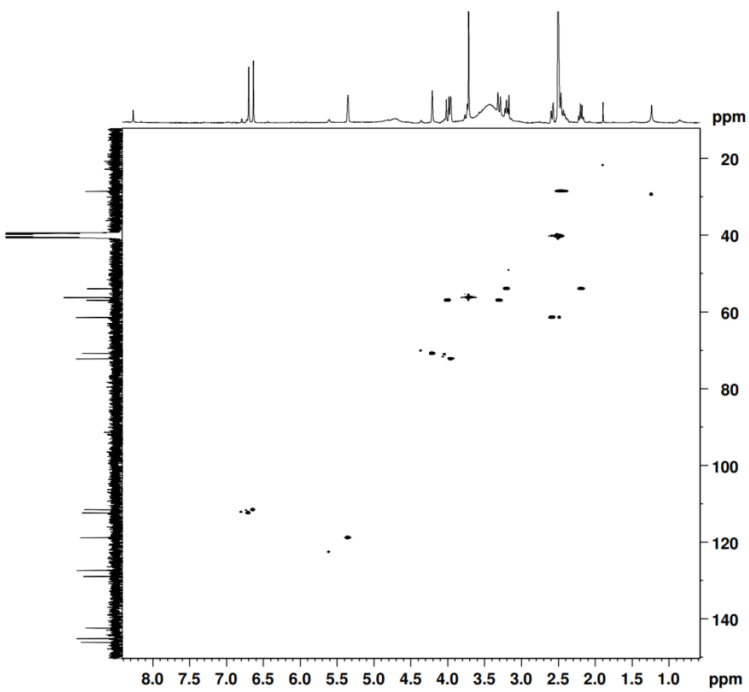


Figure S16. HSQC spectrum of compounds **2** and **3** in $(CD_3)_2SO$

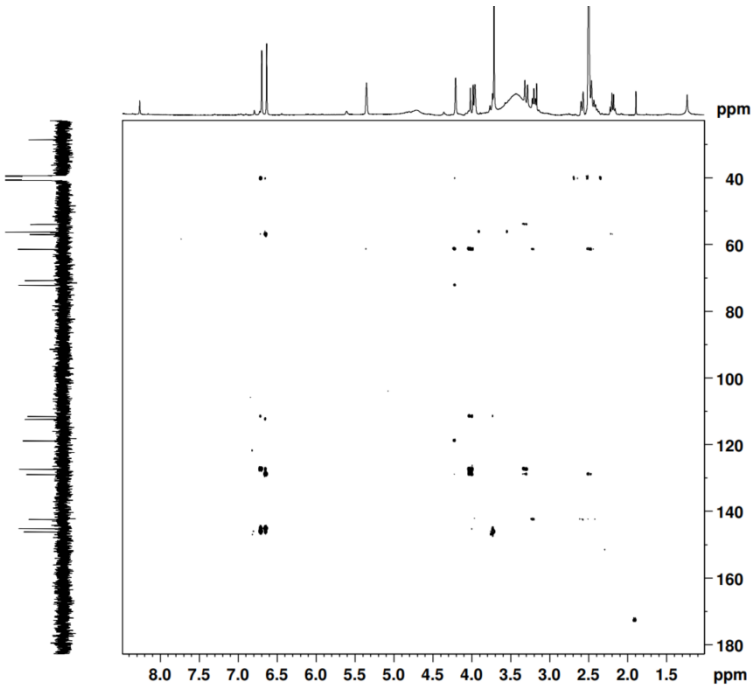


Figure S17. HMBC spectrum of compounds **2** and **3** in $(CD_3)_2SO$

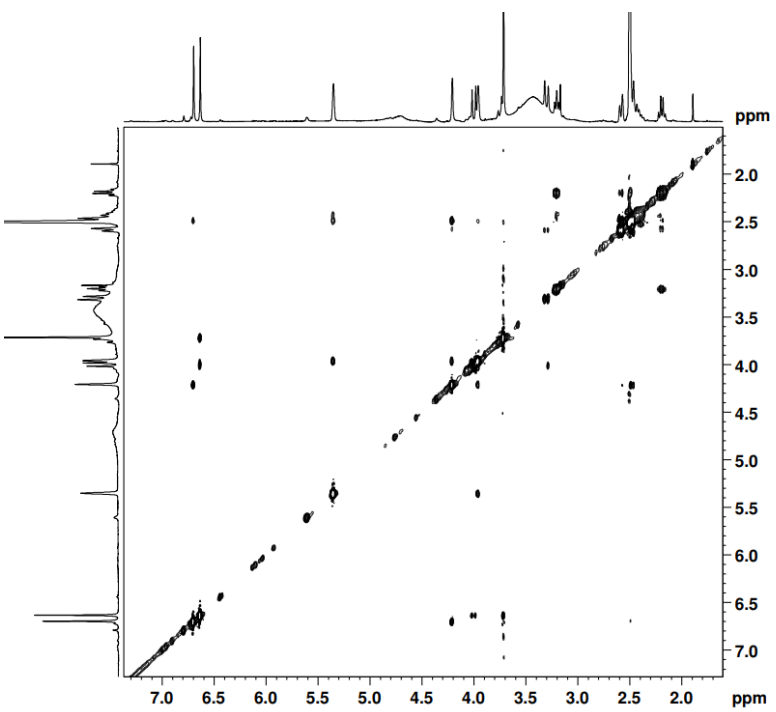


Figure S18. NOESY spectrum of compounds **2** and **3** in $(CD_3)_2SO$

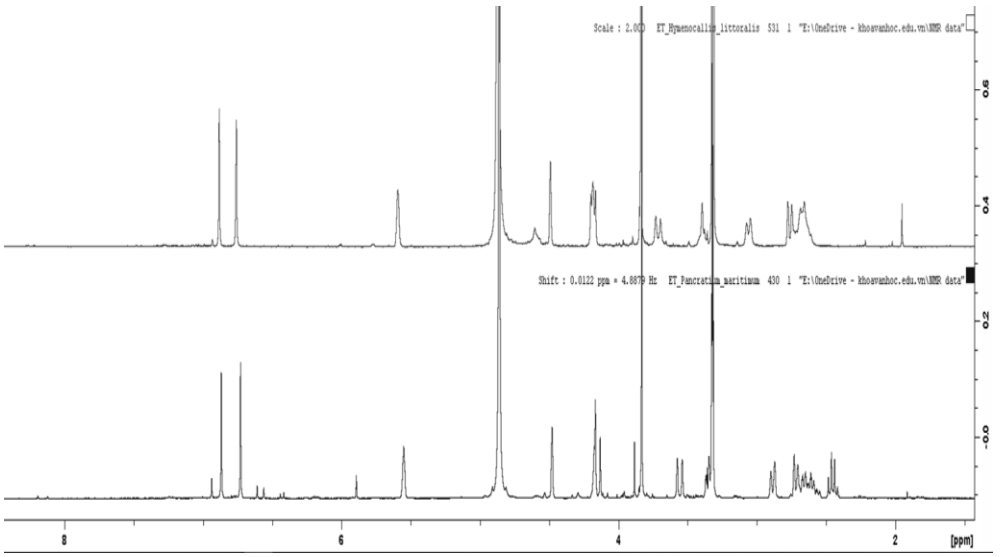


Figure S19. ^{1}H spectra of compounds **2** and **3** (upper) and pseudolycorine (lower) in CD_3OD

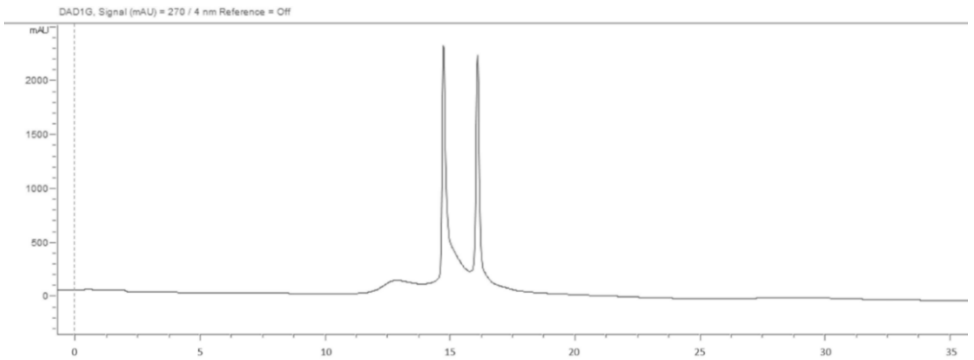


Figure S20. HPLC-UV chromatogram (270 nm) of compounds **2** and **3** after enantioseparation

4.6. References

1. Aniszewski, T. *Alkaloids: Chemistry, Biology, Ecology, and Applications*, 2nd ed.; Elsevier: Helsinki, Finland, 2015.
2. Jin, Z. Amaryllidaceae and Sceletium Alkaloids. *Nat. Prod. Rep.* **2005**, *22*, 111–126.
3. Takos, A.M.; Rook, F. Towards a Molecular Understanding of the Biosynthesis of Amaryllidaceae Alkaloids in Support of Their Expanding Medical Use. *Int. J. Mol. Sci.* **2013**, 11713–11741.
4. Marco, L.; Carreiras, C. Galanthamine, a Natural Product for the Treatment of Alzheimer's Disease. *Recent Pat. CNS Drug Discov.* **2006**, *1*, 105–111.
5. Berkov, S.; Osorio, E.; Viladomat, F.; Bastida, J. *Chemodiversity, Chemotaxonomy and Chemoecology of Amaryllidaceae Alkaloids*, 1st ed.; Elsevier Inc.: Amsterdam, The Netherlands, 2020; Volume 83.
6. National Plant Data Center NRCS, U. The Plants Database: *Hymenocallis littoralis* (Jacq.) Salisb. (accessed on 22 March 2021).
7. Jaume, B.; Strahil, B.; Laura, T. Chemical and Biological Aspects of Amaryllidaceae Alkaloids. *Recent Adv. Pharm. Sci.* **2011**, *3*, 66–100
8. Singh, G.; Saxena, R.K. Chemistry and Medicinal Properties of *Hymenocallis littoralis*. *Int. J. Sci. Res*. **2017**, *6*, 2016–2018.
9. Ding, Y.; Qu, D.; Zhang, K.M.; Cang, X.X.; Kou, Z.N.; Xiao, W.; Zhu, J.B. Phytochemical and Biological Investigations of Amaryllidaceae Alkaloids: A Review. *J. Asian Nat. Prod. Res.* **2017**, *19*, 53–100.
10. Chen, N.; Ji, Y.; Zhang, W.; Xu, Y.; Yan, X.; Sun, Y.; Song, H.; Xu, C.; Cai, L.; Zheng, H.; et al. Chemical Constituents from *Hymenocallis littoralis*. *Lett. Org. Chem.* **2016**, *13*, 536–539.
11. Ji, Y.B. Research Progress on Chemical Compositions of *Hymenocallis littoralis*. *Med. Sci. Bioeng.* **2015**, 759–762.
12. Lin, L.Z.; Hu, S.F.; Chai, H.B.; Pengsuparp, T.; Pezzuto, J.M.; Cordell, G.A.; Ruangrungsi, N. Lycorine Alkaloids from *Hymenocallis littoralis*. *Phytochemistry* **1995**, *40*, 1295–1298.
13. Ma, W.; Wang, S.; Wang, Y.; Zeng, J.; Xu, J.; He, X. Antiproliferative Amaryllidaceae Alkaloids from the Bulbs of *Hymenocallis littoralis* (Jacq.) Salisb. *Phytochemistry* **2022**, *197*, 113112.

14. Le, N.T.H.; Vermeyen, T.; Aerts, R.; Herrebout, W.A.; Pieters, L.; Tuenter, E. Epimeric Mixture Analysis and Absolute Configuration Determination Using an Integrated Spectroscopic and Computational Approach—A Case Study of Two Epimers of 6-Hydroxyhippeastidine from *Hymenocallis littoralis*. *Molecules* **2023**, *28*, 214.
15. Ingrassia, L.; Lefranc, F.; Mathieu, V.; Darro, F. Amaryllidaceae Isocarbostyril Alkaloids and Their Derivatives as Promising Antitumor Agents. *Transl. Oncol.* **2008**, *1*, 1–13.
16. Fürst, R. Narciclasine–an Amaryllidaceae Alkaloid with Potent Antitumor and Anti-Inflammatory Properties. *Planta Med.* **2016**, *82*, 1389–1394.
17. Ndibwami, A. A General Synthesis of Phenanthridinone Alkaloids. *Synlett* **2007**, *2007*, 1940–1945.
18. Ji, Y.B.; Chen, N.; Zhu, H.W.; Ling, N.; Li, W.L.; Song, D.X.; Gao, S.Y.; Zhang, W.C.; Ma, N.N. Alkaloids from Beach Spider Lily (*Hymenocallis littoralis*) Induce Apoptosis of HepG-2 Cells by the Fas-Signaling Pathway. *Asian Pacific J. Cancer Prev.* **2014**, *15*, 9319–9325.
19. Welsch, J.T.; Smalley, T.B.; Matlack, J.K.; Avalon, N.E.; Binning, J.M.; Johnson, M.P.; Allcock, A.L.; Baker, B.J.; Tuaimenals, B.-H. Merosesquiterpenes from the Irish Deep-Sea Soft Coral Duva Florida with Bioactivity against Cervical Cancer Cell Lines. *J. Nat. Prod.* **2022**, *86*, 182–190.
20. Cordell, G.A. *The Alkaloids: Chemistry and Biology, Volume 63*, 1st ed.; Academic Press: Chicago, IL, USA, 2006.
21. Cedrón, J.C.; Del Arco-Aguilar, M.; Estévez-Braun, A.; Ravelo, Á.G. *Chapter 1—Chemistry and Biology of Pancratium Alkaloids*; Cordell, 1st Ed.; Academic Press: Cambridge, MA, USA, 2010; Volume 68, pp 1–37.
22. Dewick, P.M. *Medicinal Natural Products: A Biosynthetic Approach*; John Wiley & Sons: Hoboken, NJ, USA, 2002.
23. Marcarino, M.O.; Cicetti, S.; Zanardi, M.M.; Sarotti, A.M. A Critical Review on the Use of DP4+ in the Structural Elucidation of Natural Products: The Good, the Bad and the Ugly. A Practical Guide. *Nat. Prod. Rep.* **2022**, *39*, 58–76.
24. Grimblat, N.; Gavín, J.A.; Hernández Daranas, A.; Sarotti, A.M. Combining the Power of J Coupling and DP4 Analysis on Stereochemical Assignments: The J-DP4 Methods. *Org. Lett.* **2019**, *21*, 4003–4007. https://doi.org/10.1021/acs.orglett.9b01193.

25. Zhang, Y.N.; Zhang, Q.Y.; Li, X.D.; Xiong, J.; Xiao, S.Q.; Wang, Z.; Zhang, Z.R.; Deng, C.L.; Yang, X.L.; Wei, H.P.; et al. Gemcitabine, Lycorine and Oxysophoridine Inhibit Novel Coronavirus (SARS-CoV-2) in Cell Culture. *Emerg. Microbes Infect.* **2020**, *9*, 1170–1173.
26. Jin, Y.H.; Min, J.S.; Jeon, S.; Lee, J.; Kim, S.; Park, T.; Park, D.; Jang, M.S.; Park, C.M.; Song, J.H.; et al. Lycorine, a Non-Nucleoside RNA Dependent RNA Polymerase Inhibitor, as Potential Treatment for Emerging Coronavirus Infections. *Phytomedicine* **2021**, *86*, 153440.
27. Ren, P.; Shang, W.; Yin, W.; Ge, H.; Wang, L.; Zhang, X.; Li, B.; Li, H.; Xu, Y.; Xu, E.H.; et al. A Multi-Targeting Drug Design Strategy for Identifying Potent Anti-SARS-CoV-2 Inhibitors. *Acta Pharmacol. Sin.* **2022**, *43*, 483–493.
28. Roy, M.; Long, L.; Xiaojuan, X.; Peifu, F.; Mao, Y.; Jing, L. Lycorine: A Prospective Natural Lead for Anticancer Drug Discovery. *Biomed. Pharmacother.* **2018**, *107*, 615–624.
29. Wang, P.; Li, L.F.; Wang, Q.Y.; Shang, L.Q.; Shi, P.Y.; Yin, Z. Anti-Dengue-Virus Activity and Structure-Activity Relationship Studies of Lycorine Derivatives. *ChemMedChem* **2014**, *9*, 1522–1533.
30. Grimblat, N.; Zanardi, M.M.; Sarotti, A.M. Beyond DP4: An Improved Probability for the Stereochemical Assignment of Isomeric Compounds Using Quantum Chemical Calculations of NMR Shifts. *J. Org. Chem.* **2015**, *80*, 12526–12534.
31. Yu, J.; Cao, Y.; Song, H.; Wang, X.; Yao, S. Calculations of Optical Rotation: Influence of Molecular Structure. *J. Serbian Chem. Soc.* **2012**, *77*, 56–63.
32. Stephens, P.J.; Devlin, F.J.; Cheeseman, J.R.; Frisch, M.J. Calculation of Optical Rotation Using Density Functional Theory. *J. Phys. Chem. A* **2001**, *105*, 5356–5371.
33. Bally, T.; Rablen, P.R. Quantum-Chemical Simulation of 1H NMR Spectra. 2. Comparison of DFT-Based Procedures for Computing Proton-Proton Coupling Constants in Organic Molecules. *J. Org. Chem.* **2011**, *76*, 4818–4830

CHAPTER 5

Amaryllidaceae alkaloids from *Pancratium maritimum*

Publication:

Le Ngoc-Thao-Hien, Steven De Jonghe, Kristien Erven, Tom Vermeyen, Wouter A. Herrebout, Johan Neyts, Christophe Pannecouque, Luc Pieters, and Emmy Tuenter. "A new alkaloid from *Pancratium maritimum* - Structure elucidation using computer-assisted structure elucidation (CASE) and evaluation of anti-SARS-CoV-2 activity". Phytochemistry letters, 2023, 58, 1-7.

Abstract: *Pancratium maritimum* L. (sea daffodil) is one of the most studied plant species in the Amaryllidaceae family. Recently, due to the advancement of new probabilistic methods and spectroscopic techniques, the potential for phytochemical investigations on isomeric compounds has significantly improved. In this study, application of those techniques in combination with conventional phytochemical analysis led to the isolation and identification of a new alkaloid named 2β,10bα-dihydroxy-9-*O*-demethylhomolycorine from *P. maritimum*, together with 17 known ones. Assessment of anti-SARS-CoV-2 activity and cytotoxicity on a Vero E6 cell line for all compounds revealed four compounds with weak anti-SARS-CoV-2 potency at non-cytotoxic concentrations, namely of 9-*O*-demethylhomolycorine (EC_{50} = 44 μM), galanthamine (EC_{50} = 70 μM), 1-*O*-acetyl-norpluviine (EC_{50} = 80 μM) and 1-*O*-acetyl-10-*O*-methylpseudolycorine (EC_{50} = 47 μM). On the other hand, cytotoxicity was observed for 2β,10bα-dihydroxy-9-*O*-demethylhomolycorine (CC_{50} = 13.28 μM), haemanthamine (CC_{50} = 0.76 μM), 6α- and 6β-haemanthidine (CC_{50} = 5.43 μM), 11-hydroxyvittatine (CC_{50} = 10.57 μM) and pseudolycorine (CC_{50} = 3.41 μM).

Keywords: Amaryllidaceae alkaloids, *Pancratium maritimum*, DFT calculation, VCD, SARS-CoV-2

5.1. Introduction

The genus *Pancratium* Dill. ex L. (Amaryllidaceae) consists of approximately 60 plant species according to the World Checklist of Vascular Plants (Govaerts, 2022). The name *Pancratium* is derived from the Greek word "pagkration" which means "almighty", perhaps due to the medicinal properties that species from this genus possess (Cedrón et al., 2010). *Pancratium maritimum* L. (sea daffodil, sea lily, sand lily) is widely distributed along the coastline of the Mediterranean region, from the Black Sea to part of the Atlantic sea, adapting well to the stressful conditions of sand dune environments (Rhizopoulou and Pouris, 2018). *P. maritimum* has been thoroughly studied and consequently, over half of alkaloids isolated from the *Pancratium* genus were identified in this plant species (Cedrón et al., 2010). To the best of our knowledge, 42 alkaloids were isolated heretofore from *P. maritimum*, including all known skeleton-types of Amaryllidaceae alkaloids: nine lycorine-types (lycorine, pseudolycorine, hippadine, ungeremine, ungiminorine, ungiminorine *N*-oxide, zefbetaine, *α*-dihydrocaramine and galanthane), five homolycorine-types (lycorenine, homolycorine, hippeastrine, 8-*O*-demethylhomolycorine and 9-*O*-demethylhomolycorine), four galanthamine-type (galanthamine, *N*-formylnorgalanthamine, habranthine and pancrimatine C), six haemanthamine-types (haemanthamine, two 6-epimers of haemanthidine, 6-*O*-methylhaemanthidine, *O*-demethylhaemanthamine and 8-*O*-demethylmaritidine), five crinine-types (crinine, 3*β*,11*α*-dihydroxy-1,2-dehydrocrinane, 8-hydroxy-9-methoxycrinine, crinine-3-one and buphanisine), two tazettine-types (tazettine and 6*α*-deoxytazettine), five phenanthridone-types (pancratistatin, narciclasine-4-O-*β*-D-glucopyranoside, *N*-methyl-8,9-methylenedioxy-6-phenanthridone, pancrimatine A and pancrimatine B), two phenanthridine-types (trispheridine and *N*-methyl-8,9-methylenedioxy phenanthridine), one montanine-type (montanine), one norbelladine-type (4'-*O, N*-dimethylnorbelladine), and two compounds with a rare scaffold (gralicine and norismine) (Cedrón et al., 2010; Youssef and Frahm, 1998; Ibrahim et al., 2013; Abou-Donia et al., 1992; Berkov et al., 2004; Youssef et al., 2022).

Nowadays, natural product chemists are paying more attention to 3D structure determination for three main reasons: (1) – isomers are a tempting source of

undiscovered natural products; (2) – isomers tend to have different biological activities, which is important for discovering potent, novel hits and leads, and (3) – many incorrect stereochemical assignments were reported in the last decades due to over-reliance on conventional spectroscopic tools (NMR, HRMS), which provide insufficient 3D structural information in some cases (Chhetri et al., 2018). Over the past few years, computer-assisted 3D structure elucidation (CASE-3D) has become a powerful tool to support the structural determination process alongside spectroscopic techniques, from 2D to 3D structures, which reduces the occurrence of misassignment (Ermanis et al., 2019; Zanardi et al., 2021; Marcarino et al., 2022). The power of integrating computational calculation into the traditional process of structure elucidation was proven with many recent successful examples (Chhetri et al., 2018; Sarotti, 2020; Le et al.; 2021; Le et al., 2023a). In this study, a re-investigation of the alkaloidal composition of *P. maritimum* resulted in the isolation and complete characterization of eighteen alkaloids, of which one is reported for the first time using the combination of NMR, HRMS and CASE-3D. Furthermore, the absolute configuration of this novel structure was confirmed by means of vibrational circular dichroism (VCD) spectroscopy and infrared (IR) spectroscopy.

Important bioactivities of alkaloids from *P. maritimum* include anticancer, antiplasmodial, anti-inflammatory, antimicrobial and antioxidant activities, which have been explored for decades (Cedrón et al., 2010; Leporini et al., 2018; Cimmino et al., 2017). Antiviral activity of Amaryllidaceae alkaloids also has drawn the attention of many researchers, especially concerning activity against the dengue, HIV and hepatitis viruses (Nair et al., 2023; Masi et al., 2022). Since 2019, the severe acute respiratory syndrome coronavirus-2 (SARS-CoV-2) has emerged, and continues to threaten human beings' lives. On the one hand, large scale vaccination helped us regain our normal lives, but the goal to permanently eradicate the virus remains challenging, since the virus keeps evolving and new resistant variants keep arising. On the other hand, the use of medicinal plants as complementary and/or alternative medicines still prevails in many places worldwide (Nair et al., 2023; Raman et al., 2022; Christy et al., 2021; Isidoro et al., 2022). However, this practice is primarily based on indigenous and traditional knowledge of plants used for treating respiratory diseases (Owen et al., 2022). In the course of 2020-2022, many natural products were reported to exhibit anti-

SARS-CoV-2 activity, originating from various sources (plants, marine sources, microbes) (Avalon et al, 2022; Li et al., 2021). For example, lycorine, the most abundant Amaryllidaceae alkaloid, was named as a potent candidate by several studies (Zhang et al., 2020; Jin et al., 2021; Ren et al., 2022). Thus, the application of Amaryllidaceae alkaloids for the treatment of SARS-CoV-2 infection became a compelling research topic, not only because of their antiviral potential, but also considering the widespread and highly accessible resources of Amaryllidaceae alkaloids. In the current work, anti-SARS-CoV-2 activity was examined for all isolated compounds from *P. maritimum*, and structure-activity relationships (SAR) were investigated.

5.2. Results and Discussion

5.2.1. Structure Elucidation

The current phytochemical investigation resulted in the isolation and characterization of eighteen alkaloids possessing a wide range of polarities (see **Fig. 1**). Compounds **2**, **3**, **5**, **9**, **10**, **12 – 15** and **18** were purified from the DCM extract; compounds **1**, **4**, **6 – 8**, **10**, **16**, **17** from the EtOAc extract. Compound **1** was described in this study for the first time and its structural elucidation process will be elaborated on hereafter.

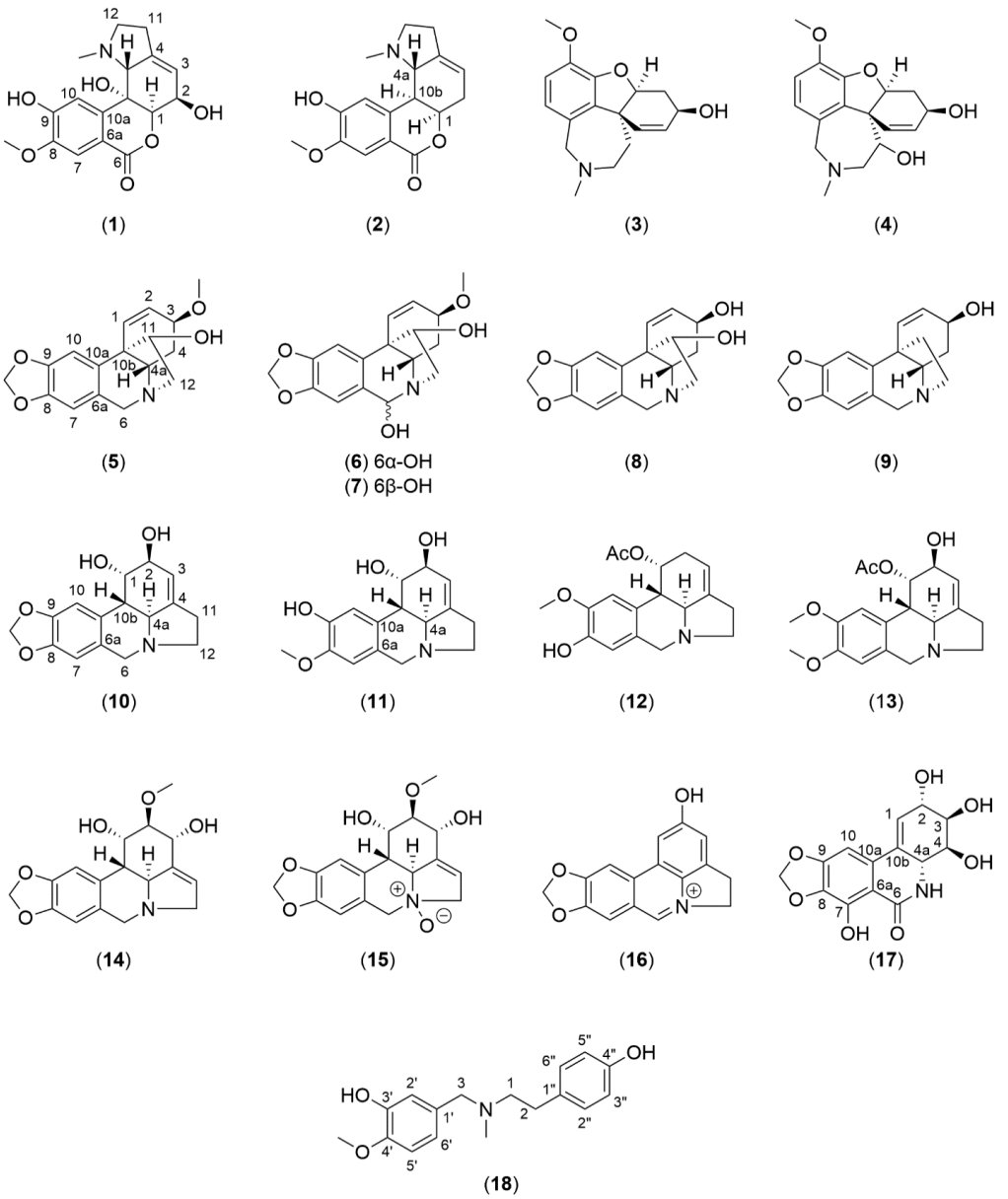


Fig. 1. Structures of 18 isolated alkaloids from *Pancratium maritimum*

Compound **1** was isolated as white powder (11.0 mg) and compound **2** as yellowish amorphous solid (5.0 mg). Their accurate masses suggested chemical formulas of $C_{17}H_{19}NO_6$ (*m/z* 334.1290 $[M+H]^+$) and $C_{17}H_{20}NO_4$ (*m/z* 302.1395 $[M+H]^+$), respectively. Both compounds **1** and **2** possessed a carbon signal at 164.1 ppm (Table 1), which was likely to be involved in either a lactone or a lactam ring. In

other words, the two plausible Amaryllidaceae scaffolds were either the homolycorine- or the narciclasine-type, respectively. However, due to the presence of a typical N-Me moiety (δ_H = 1.82, δ_C = 44.1) which does not suit with the NH moiety of the lactam ring of the narciclasine-type, the homolycorine-type scaffold was selected for further NMR analysis. Compared to NMR data of homolycorine reported in literature, compound **2** only lacked one methoxy group and therefore, based on its 2D NMR spectra, compound **2** was easily assigned as 9-*O*-demethylhomolycorine (Kihara et al., 1991). The molecular weight of compound **1** was 32 mass units higher compared to compound **2**, indicating a possible addition of two oxygens, which might be due to two additional hydroxy groups. This was confirmed by the ^{13}C NMR data, which revealed that compound **1** indeed possessed two more oxygenated carbons (occurring around 70 ppm in ^{13}C NMR) than compound **2**.

Next, separate moieties were elucidated prior to connecting them. Two methylenes were observed in the HSQC spectrum: δ_H-3.05/δ_H-2.21 and δ_H-2.44/δ_H-2.38. They also correlated with each other in the COSY spectrum (**Fig. 2**). Another spin system was identified by COSY cross-peaks between δ_H-5.57, δ_H-4.10 and δ_H-4.35, and the three hydrogens were all methine groups. Two singlets, occurring at 7.16 and 7.31 ppm, must belong to the aromatic ring. An N-Me, as aforementioned, and an O-Me (δ_H = 3.82, δ_C = 56.1) were also identified (**Table 1**).

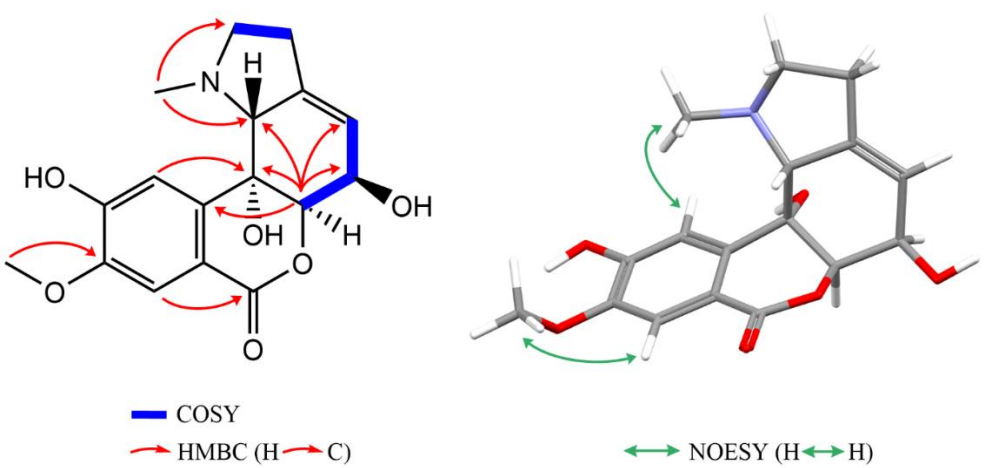


Fig. 2. Key COSY, HMBC and NOESY correlations observed in compound **1**

Table 1. Experimental and calculated ^{1}H and ^{13}C-NMR data of compound **1** in $(CD_3)_2SO$

Position	Experimental		DP4		DP4+	
	δ_H (mult., *J* in Hz)	δ_C	δ_H	δ_C	δ_H	δ_C
1	4.35 (s)	83.9	4.47	81.1	4.26	80.8
2	4.1*	68.0	4.35	67.1	4.23	66.8
3	5.57 (m)	119.6	5.68	122.6	5.63	122.7
4		142.5		145.4		145.4
4a	2.6 (brs)	70.2	3.01	69.3	2.93	69.3
6		164.1		163.8		163.7
6a		113.6		114.7		114.6
7	7.31 (s)	112.3	7.35	111.5	7.42	111.8
8		147.9		145.1		145.3
9		152.2		151.6		151.6
10	7.16 (s)	112.7	7.07	110.0	7.12	110.2
10a		142.5		144.8		144.6
10b		67.9		71.1		70.6
11	2.38*	28.4	1.95	32.8	1.97	32.6
	2.44*	28.4	2.40	32.8	2.46	32.6
12	2.21 (q, 8.7)	56.0	2.40	55.0	2.38	55.1
	3.05 (td, 8.6, 3.4)	56.0	2.93	55.0	2.95	55.1
N-Me	1.82 (s)	44.1	1.85	42.4	1.87	42.6
8-OMe	3.82 (s)	56.1	3.65	53.7	3.70	54.1
2-OH	4.80 (d, 7.2)					
10b-OH	4.96 (brs)					
9-OH	10.3 (brs)					
RMSD			0.15	2.0	0.12	1.9
E_{Max}			0.43	4.4	0.41	4.2

*overlapping signals

Starting from the benzene ring, it is known for the homolycorine-type skeleton that due to the presence of a lactone moiety, the signal of H-7 is more downfield than

the one corresponding to H-10. Thus, δ_H-7.31 was assigned to H-7 and δ_H-7.16 to H-10. The HMBC correlation between δ_H-7.31 and δ_C-164.1 further confirmed the assignment (**Fig. 2**). Attachment of the O-Me group to the benzene ring was deduced from the HMBC correlation between δ_H-3.82 and δ_C-147.9, and the NOESY cross-peak between δ_H-3.82 and δ_H-7.31 (H-7) proved its position at C-8. The NOE effect observed between H-10 and N-Me again confirmed a homolycorine-type skeleton, since only this subclass can provide spatial proximity between these hydrogens. Compared to compound **2**, an exact match was found for the NMR data of the nitrogen-containing ring, implying that this moiety was identical in compound **1** (C-4, C-4a, C-11 and C-12). Therefore, it was obvious that the main structural discrepancies should be found in the two rings in the middle of the homolycorine scaffold. The olefinic hydrogen and carbon (δ_H-5.57 and δ_C-119.6) were also similar to compound **2**. Given the fact that compound **2** has a four-membered spin system formed by three methine groups and a methylene, and that compound **1** has a three-membered spin system formed by three methine groups, the two additional hydroxy groups must be located at C-10b and C-2 (**Fig. 2**). Combining all data, the structure of compound **1** was proposed as shown in Figures 1 and 3. In 2015, Carvalho et al. reported 2α-10bα-dihydroxy-9-*O*-demethylhomolycorine possessing the same 2D structure, but H-4a of this compound was found at 4.24 ppm in ^{1}H-NMR, while that of compound **1** was 2.60 ppm (Kaline et al., 2015). Hence, compound **1** must be a diastereomer of the reported compound and was characterized for the first time.

As the NOESY experiment did not provide sufficient information to deduce the relative configuration of compound **1**, ^{1}H and ^{13}C chemical shifts were computed to perform DP4 and DP4+ probabilistic methods (**Table 1**). **Table 2** displays the probabilities of eight possible diastereomers, and in fact both DP4 and DP4+ led to an identical result (the RRRR configuration). As shown in **Table 1**, the root-mean-square-deviations (RMSD) obtained were maximum 2.0 ppm for carbons and 0.15 ppm for hydrogens, which is considered as a perfect match between experimental and calculated values. The maximum outlier observed (E_{max}) was that of C-11, but C-11 is not a stereocenter and is located two bonds away from the nearest stereocenter (C-4a); therefore, it does not significantly contribute to the stereochemical information. After combining all ^{1}H and ^{13}C data, the relative

configuration (1*R*, 2*R*, 4a*R*, 10b*R*) which possesses a probability of 100% was proposed for compound **1**. VCD analysis was performed to verify the propositions from DP4 and DP4+ probabilities. Experimental VCD spectrum were measured in $(CD_3)_2SO$ (**Fig. 3**). A careful analysis of the experimental data and the results obtained with the DFT calculations was performed, with attention for the robustness of the calculated spectra (see Supporting Information). From this, it became clear that RRRR matches best to the experimental spectrum (compared to SSSS). Therefore, the absolute configuration of compound **1** was confirmed to be (1*R*, 2*R*, 4a*R*, 10b*R*), and the name 2β,10bα-dihydroxy-9-*O*-demethylhomolycorine was adopted.

Table 2. DP4 and DP4+ probabilities (%). Stereocenter order is (1, 2, 4a, 10b).

Diastereomer	**DP4**			**DP4+**		
	^{1}H	^{13}C	^{1}H and ^{13}C	^{1}H	^{13}C	^{1}H and ^{13}C
RRRR	99.89	4.06	100.00	8.26	0.02	100.00
RRRS	0.00	0.08	0.00	0.00	0.00	0.00
RRSR	0.00	0.01	0.00	0.00	0.00	0.00
RRSS	0.00	0.06	0.00	0.00	0.00	0.00
RSRR	0.00	95.64	0.00	0.00	99.98	0.00
RSRS	0.00	0.12	0.00	0.00	0.00	0.00
RSSR	0.11	0.00	0.00	91.74	0.00	0.00
RSSS	0.00	0.03	0.00	0.00	0.00	0.00

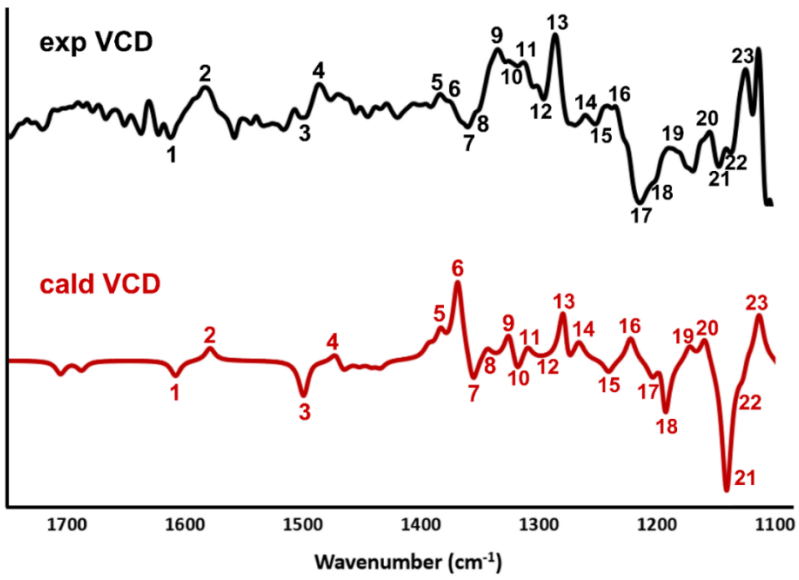


Fig. 3. Experimental and calculated VCD spectra of 2β,10bα-dihydroxy-9-*O*-demethylhomolycorine (**1)**

NMR data of compounds **3** – **18** matched with those of compounds reported in literature and were identified as galanthamine (**3**) (Zotta et al., 1971), habranthine (**4**) (Wildman et al., 1968), haemanthamine (**5**) (Zotta et al., 1971), 6α-haemanthidine (**6**) and 6β-haemanthidine (**7**) (Zhang et al., 2006), 11-hydroxyvittatine (**8**) and vittatine (**9**) (Viladomat et al., 1995), lycorine (**10**), pseudolycorine (**11**), 1-*O*-acetyl-norpluviine (**12**) (Campbell et al., 2000), 1-*O*-acetyl-10-*O*-methylpseudolycorine (**13**) (Sarikaya et al., 2013), ungiminorine (**14**) (Richomme et al., 1989), ungiminorine *N*-oxide (**15**) and ungeremine (**16**) (Suau et al., 1988), narciclasine (**17**) (Southgate et al., 2017), and 4'-*O*-methylnorbelladine (**18**) (Ghosal et al., 1985). Among those, three compounds, i.e. 1-*O*-acetyl-norpluviine, 1-*O*-acetyl-10-*O*-methylpseudolycorine and 4'-*O*-methylnorbelladine were isolated from *P. maritimum* for the first time. Detailed experimental NMR and MS data can be found as Supporting Information.

2.2.2. Anti-SARS-CoV-2 Activity and Cytotoxicity

All analogues were evaluated for antiviral activity in SARS-CoV-2 infected VeroE6 cells. GS-441524, the parent nucleoside of remdesivir was included as positive control, displaying an EC_{50} value of 0.81 µM and a CC_{50} value of 72.38 µM. 9-*O*-demethylhomolycorine (**2**), galanthamine (**3**), 1-*O*-acetyl-norpluviine (**12**), and 1-*O*-acetyl-10-*O*-methylpseudolycorine (**13**) exhibited weak inhibition of SARS-CoV-2 (EC_{50} = 44-80 µM) at non-cytotoxic concentrations (CC_{50} > 100 µM). 9-*O*-Demethylhomolycorine (**2**) and 1-*O*-acetyl-10-*O*-methylpseudolycorine (**13**) were the most potent ones among all tested compounds, both displaying EC_{50} values of around 45 µM in duplicate, respectively. Habranthine (**4**), ungiminorine (**14**), ungiminorine *N*-oxide (**15**), ungeremine (**16**), and 4'-*O*-methylnorbelladine (**18**) were devoid of activity and cytotoxicity.

The cytotoxicity of lycorine (**10**) on a Vero E6 cell line was previously reported by the authors in two studies of alkaloids isolated from *Hymenocallis littoralis* and *Scadoxus multiflorus* (Le et al., 2023b; Le et al. 2023c). Apart from them, it was observed in the current investigation, using the same assay, that 2β,10bα-dihydroxy-9-*O*-demethylhomolycorine (**1**), haemanthamine (**5**), 6α- and 6β-haemanthidine (**6** and **7**), 11-hydroxyvittatine (**8**), and pseudolycorine (**11**) lacked selective antiviral activity, since they were cytotoxic to the Vero E6 cells as evidenced by CC_{50} values of 13.28 µM, 0.76 µM, 5.43 µM, 10.57 µM, and 3.41 µM, respectively.

5.3. Conclusion

In this study, a re-investigation of the alkaloidal composition of *P. maritimum* led to the isolation of an undescribed alkaloid (2β,10bα-dihydroxy-9-*O*-demethylhomolycorine) and 17 known ones. Three compounds (1-*O*-acetyl-norpluviine, 1-*O*-acetyl-10-*O*-methylpseudolycorine and 4'-*O*-methylnorbelladine) were isolated from *P. maritimum* for the first time. Assessment of anti-SARS-CoV-2 activity and cytotoxicity on a Vero E6 cell line revealed four compounds with weak anti-SARS-CoV-2 potency at non-cytotoxic concentrations, namely of 9-*O*-demethylhomolycorine (**2**), galanthamine (**3**), 1-*O*-acetyl-norpluviine (**12**), and 1-*O*-acetyl-10-*O*-methylpseudolycorine (**13**) with EC_{50} values of 44 µM, 70 µM, 80

μM, and 47 μM, respectively. Cytotoxicity was observed for 2β,10bα-dihydroxy-9-*O*-demethylhomolycorine (**1**) (CC_{50} = 13.28 μM), haemanthamine (**5**) (CC_{50} = 0.76 μM), 6α- and 6β-haemanthidine (**6** and **7**) (CC_{50} = 5.43 μM), 11-hydroxyvittatine (**8**) (CC_{50} = 10.57 μM), and pseudolycorine (**11**) (CC_{50} = 3.41 μM). Although only weak anti-SARS-CoV-2 activity was observed in the VeroE6 cells for a selected number of derivatives, it might be worthwhile to evaluate their anti-SARS-CoV-2 activity in physiologically more relevant cell types, such as in the lung cancer cell lines Calu-3 and A549.

5.4. Materials and Methods

5.4.1. Isolation and Purification

About 4 kg of fresh bulbs of *P. maritumum* were collected between April and June 2017 in 'Praia do Ancão', Southern Portugal (Algarve, Loulé), GPS coordinates 37°01'59.4"N 8°02'15.4"W. A voucher specimen (no. MBH37), identified by Dr. Luisa Custodio, is kept at the Centre of Marine Sciences, University of Algarve, Faro, Portugal. Freeze-drying yielded about 750 g of residue.

Ultrasonication-assisted extraction with methanol was carried out on roughly 700 g of freeze-dried material (2.5 L x 5). After evaporating solvent under reduced pressure, 80 g of crude extract was obtained and subsequently suspended in water. The general alkaloidal extraction procedure was then applied. Briefly, the water suspension was acidified to pH < 3 by HCl 5%, following by liquid-liquid partitioning with DCM (x3) to remove acidic impurities. Next, the water phase was basified to pH > 9 by NH_4OH 25% before partitioning with DCM (x3) which yielded the first alkaloid extract (4.4 g DCM extract). The third liquid-liquid partition with EtOAc (x3) was performed thereafter yielding 2.0 g EtOAc extract.

The two extracts were further fractionated by flash chromatography using the solvent mixture of DCM and MeOH + 0.5% NH_4OH and normal phase silica gel. As for the DCM extract, the following gradient was used: 0-10 min (0% MeOH), 10-20 min (5% MeOH), 20-35 min (10% MeOH), 35-45 min (20% MeOH), 45-55 min (30% MeOH), 55-65 min (40% MeOH), 65-80 min (80% MeOH); while the gradient was halted during elution of detected compounds. The gradient applied to the EtOAc extract was as follows: 0-5 min (0% MeOH), 5-10 min (5% MeOH), 10-20 min (10%

MeOH), 20-30 min (20% MeOH), 30-40 min (30% MeOH), 50-60 min (40% MeOH), 60-70 min (50% MeOH), 70-80 min (80% MeOH). In the end, 22 fractions were obtained from the DCM extract (DCM-1 to DCM-22), and 26 fractions form the EtOAc extract (EtOAc-1 to EtOAc-22).

Lycorine (**10**) crystalized directly in many fractions as white needles and was washed with methanol to remove it before all fractions were subjected to analytical HPLC. After profiling, promising fractions were selected for semi-preparative separation.

Semi-preparative HPLC-DAD-MS with manual collection was performed for isolation of pure compounds from the fractions DCM-3, DCM-4, DCM-8, DCM-12, DCM-13, EtOAc-6, EtOAc-7, and EtOAc-15. More specifically, the general preparative LC conditions for all fractions were: (1) sample concentration: 1 mg/mL in water; (2) injection volume: 200 - 600 µL; (3) solvent A: H2O + 0.5% NH_4OH and solvent B: MeOH + 0.5% NH_4OH; (4) UV detection: 210 and 254 nm; (6) flow rate: 3 mL/min. Compounds **2**, **12**, **13**, **18** were purified from fractions DCM-3 and DCM-4; compounds **12 – 15** from DCM-8; compounds **3**, **5**, **9**, **10** from DCM-12 and DCM-13; compounds **1**, **4**, **6 – 8** from EtOAc-6 and EtOAc-7; and compounds **10**, **16**, **17** from EtOAc-15.

2β,10bα-dihydroxy-9-*O*-demethylhomolycorine (**1**)

White powder (11.0 mg); UV λ_{max} 210, 254 nm; ^{1}H- and ^{13}C-NMR ($(CD_3)_2SO$, 400 and 100 MHz): see Table 1; Positive HRESIMS *m/z* 334.1290 $[M+H]^+$ (calcd for $C_{17}H_{20}NO_6$, 334.1291)

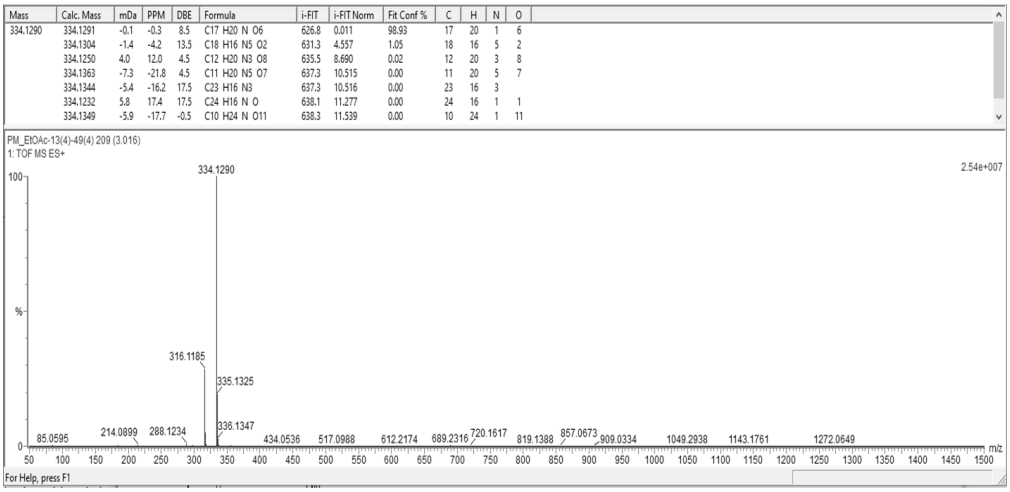
5.4.2. Computational Details

Conformational analysis was performed by PCMODEL (version 10.0) using the MMFF94 force field applying 8 and 7 kcal.mol^{-1} energy windows for two consecutive conformational search cycles. Afterwards, for chemical shift calculations, the level PCM/mPW1PW91/6-311G*//M06-2X/6-31G** was applied for DP4 probability and the level PCM/mPW1PW91/6-311G+(d,p)//B3LYP/6-31G* for DP4+ probability. After dereplication, resulting isotropic shielding values were referenced to TMS and then either scaled by linear regression for DP4 probability

or used directly for DP4+ probability. To compute VCD and IR spectra, the level PCM/B3LYP/6-311G++(2d,2p) was chosen. PCM solvation model was carried out with $(CD_3)_2SO$ to mimic experimental conditions.

Geometry optimization, frequency and shielding tensor quantum mechanical calculations were performed by the Gaussian16 program. Boltzmann distribution was estimated using the sum of electronic and thermal free energies at 298.15 K. Only conformers having energies within 2.5 kcal mol^{-1} from the global minimum were considered as contributing ones and submitted to the GIAO (Gauge-Independent Atomic Orbital) and VCD simulations.

5.5. Supporting Information – Chapter 5



Mass	Calc. Mass	mDa	PPM	DBE	Formula	i-FIT	i-FIT Norm	Fit Conf %	C	H	N	O
334.1290	334.1291	-0.1	-0.3	8.5	C17 H20 N O6	626.8	0.011	98.93	17	20	1	6
	334.1304	-1.4	-4.2	13.5	C18 H16 N5 O2	631.3	4.557	1.05	18	16	5	2
	334.1250	4.0	12.0	4.5	C12 H20 N3 O8	635.5	8.690	0.02	12	20	3	8
	334.1363	-7.3	-21.8	4.5	C11 H20 N5 O7	637.3	10.515	0.00	11	20	5	7
	334.1344	-5.4	-16.2	17.5	C23 H16 N3	637.3	10.516	0.00	23	16	3	
	334.1232	5.8	17.4	17.5	C24 H16 N O	638.1	11.277	0.00	24	16	1	1
	334.1349	-5.9	-17.7	-0.5	C10 H24 N O11	638.3	11.539	0.00	10	24	1	11

Fig. S1. HRMS spectrum and elemental analysis of compound **1**

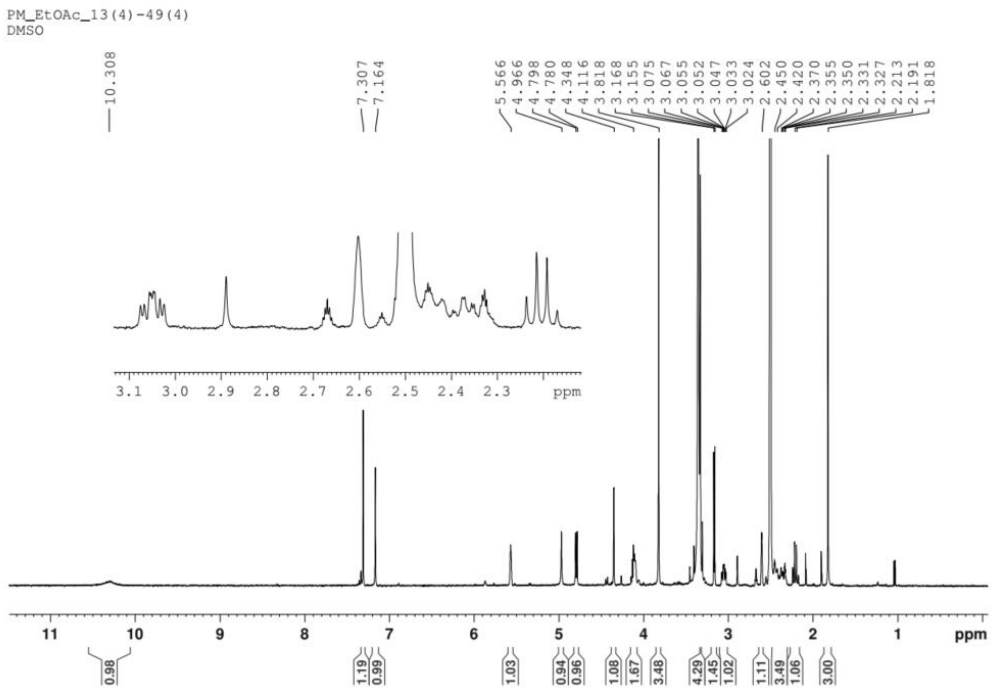


Fig. S2. ^{1}H spectrum of compound (**1**) in DMSO-d_6

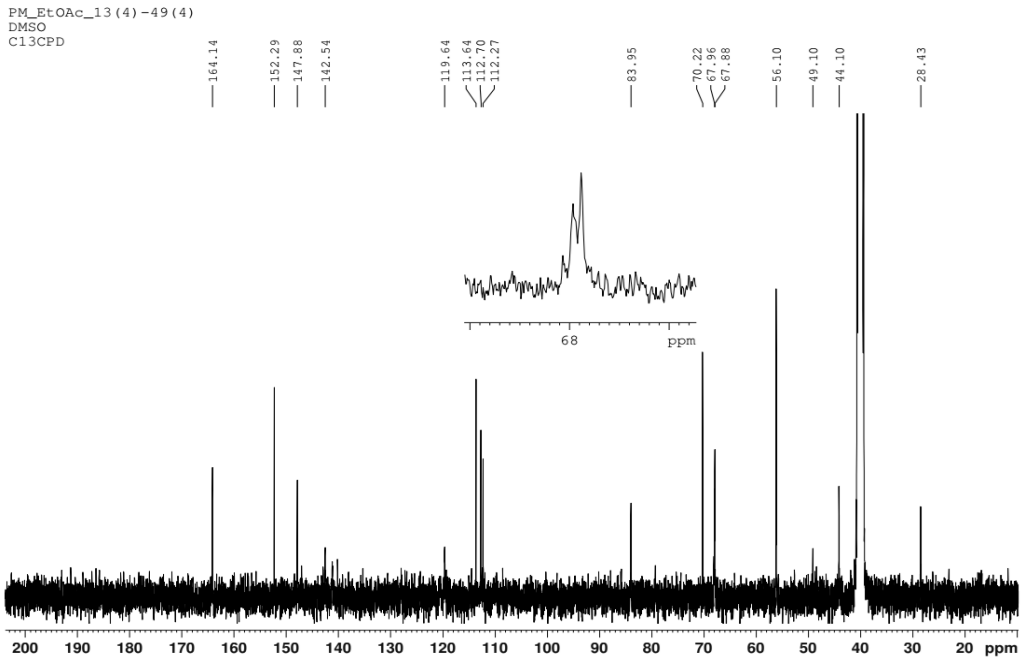


Fig. S3. ^{13}C spectrum of compound (**1**) in DMSO-d_6

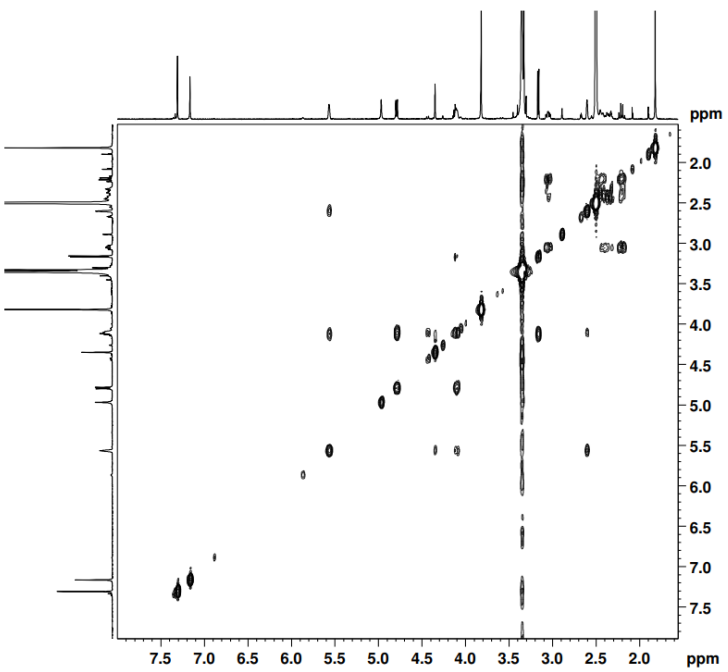


Fig. S4. COSY spectrum of compound (**1**) in DMSO-d_6

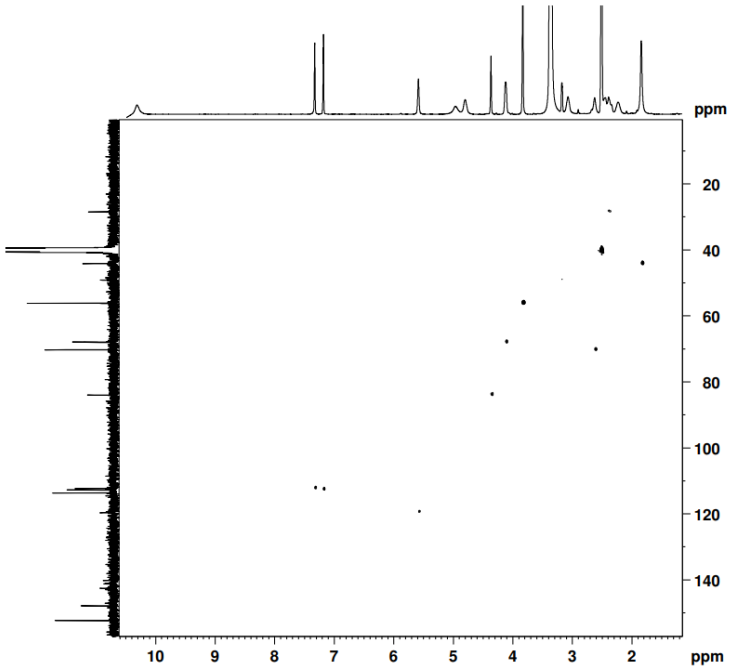


Fig. S5. HSQC spectrum of compound (**1**) in DMSO-d_6

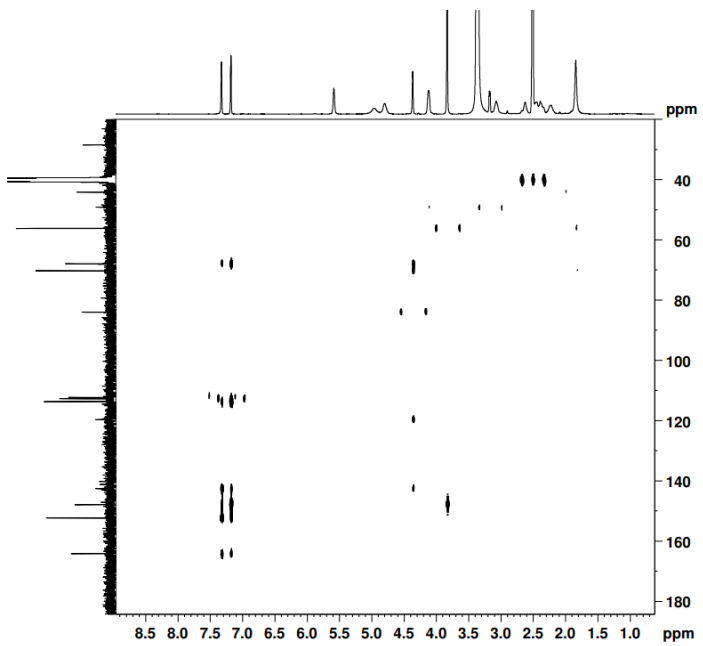


Fig. S6. HMBC spectrum of compound (**1**) in DMSO-d_6

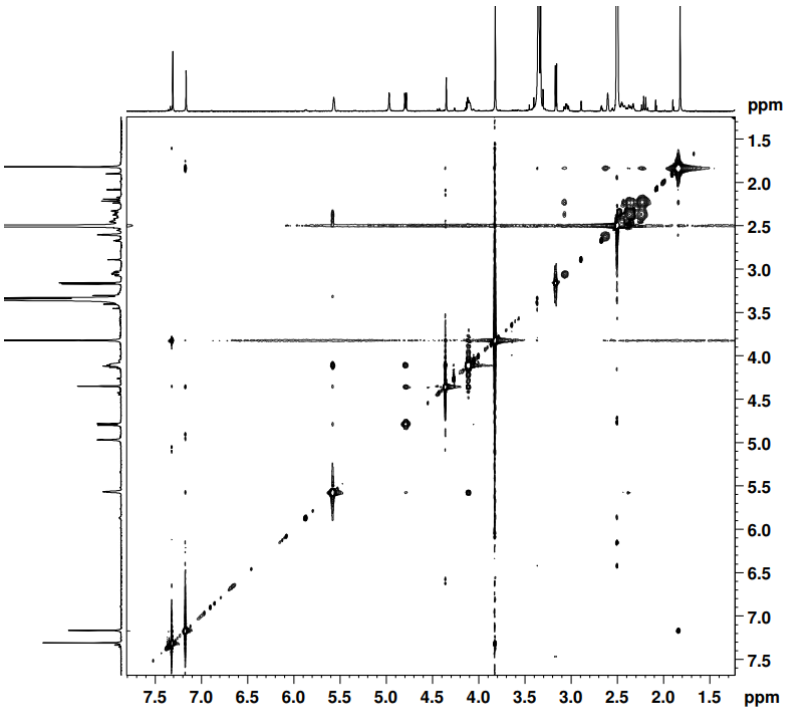


Fig. S7. NOESY spectrum of compound (**1**) in DMSO-d_6

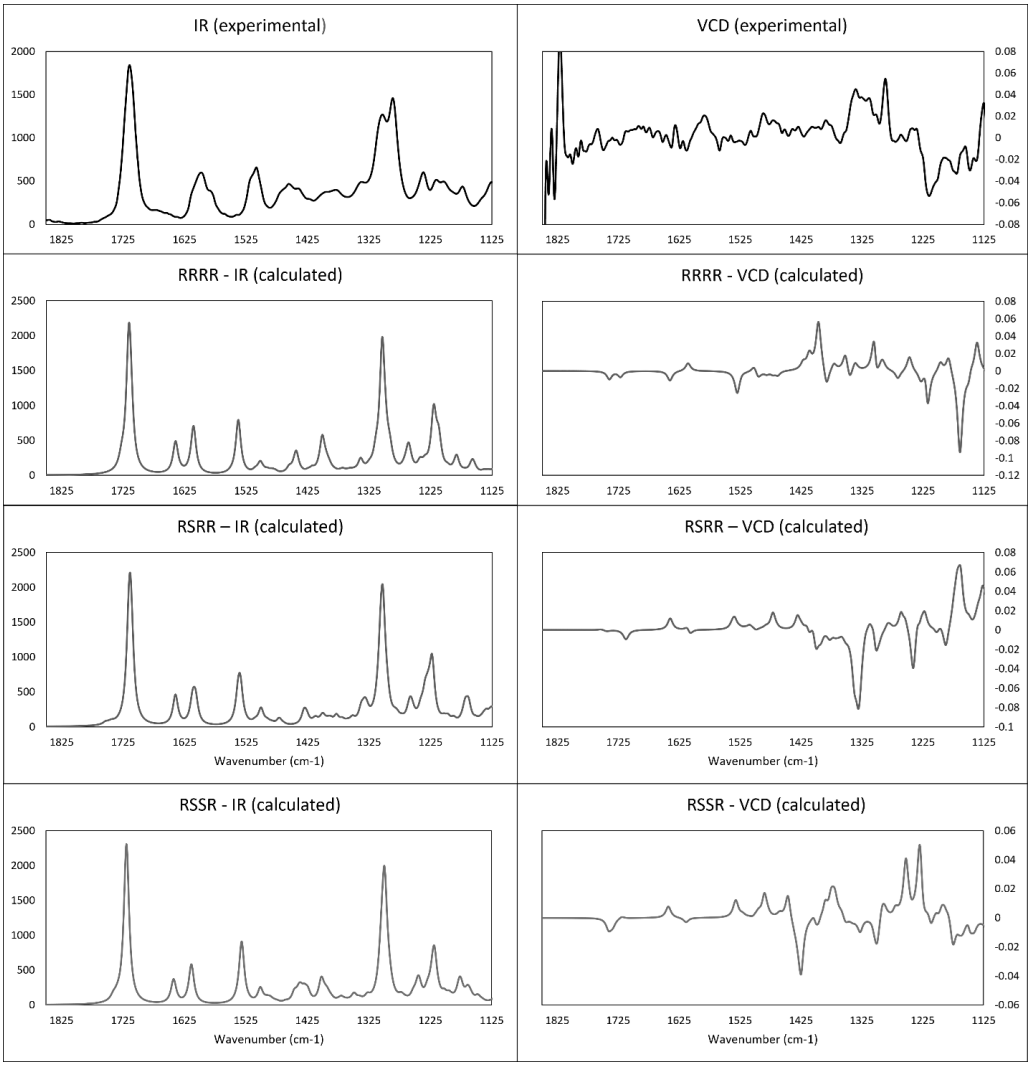


Fig. S8: Experimental VCD and IR Spectra of 2β,10bα-dihydroxy-9-*O*-demethylhomolycorine (**1)** and Computed Spectra of the RRRR, RSRR and RSSR Configurations

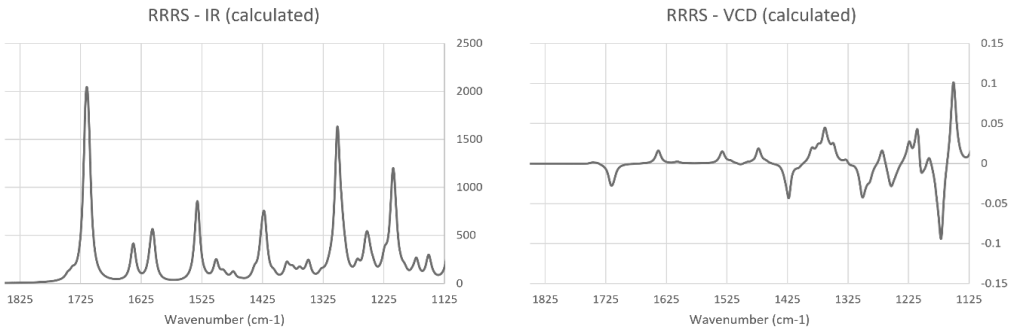


Fig. S9: Computed IR and VCD spectra of the RRRS configuration

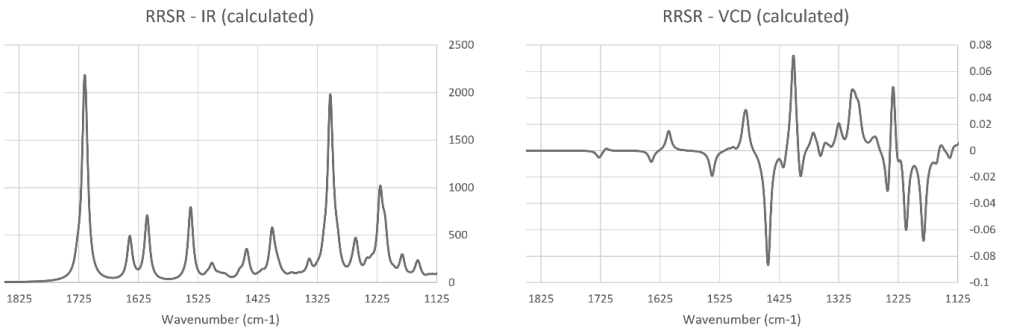


Fig. S10: Computed IR and VCD spectra of the RRSR configuration

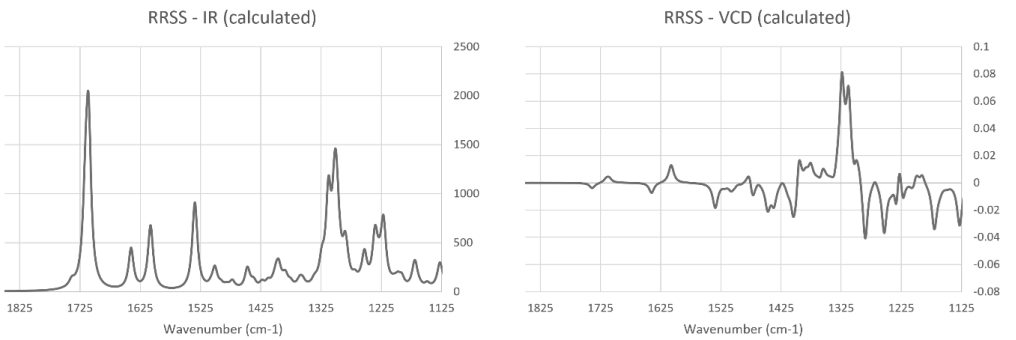


Fig. S11: Computed IR and VCD spectra of the RRSS configuration

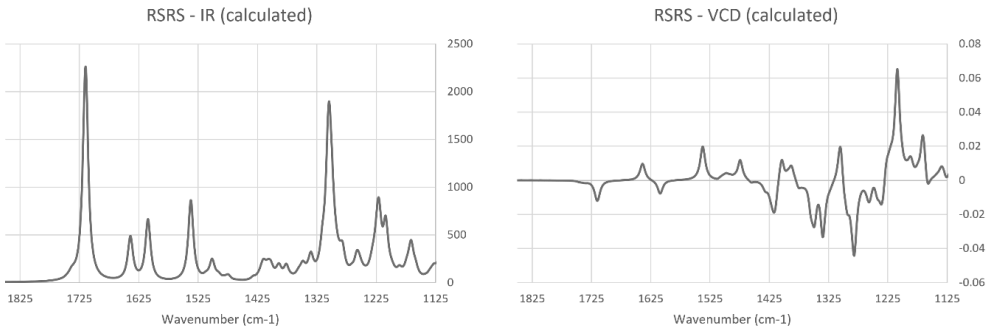


Fig. S12: Computed IR and VCD spectra of the RSRS configuration

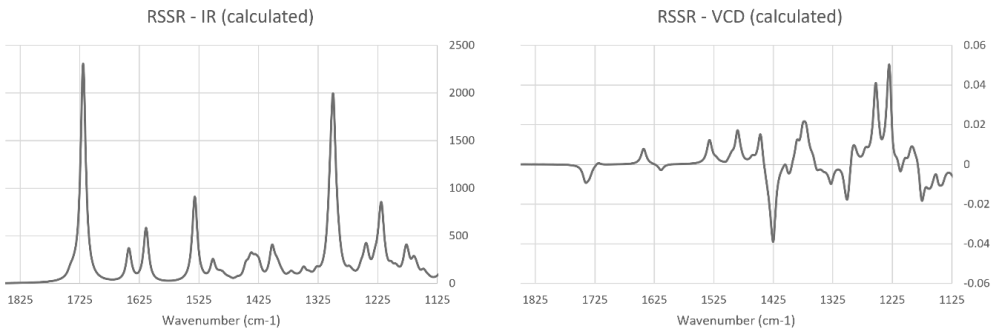


Fig. S13: Computed IR and VCD spectra of the RSSR configuration

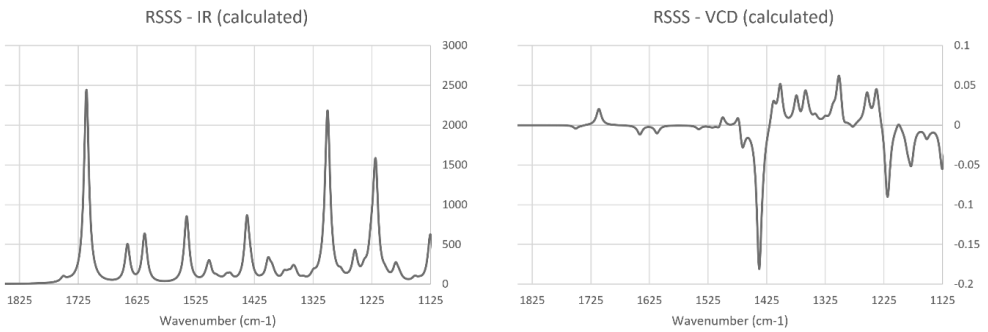


Fig. S14: Computed IR and VCD spectra of the RSSS configuration

Table S1. Anti-SARS-CoV-2 and cytotoxicity screening results on Vero E6-EGFP cells

Compounds		**EC_{50} (µM)**		**CC_{50} (µM)**	**av. EC_{50} (µM)**	**av. CC_{50} (µM)**	**SI**
1	>	n.a.	=	14.87	n.a.	13.27	n.a.
	>	n.a.	=	11.68			
2	=	52	>	100	44	>100	>2
	=	37	>	100			
3	=	95	>	100	70	>100	>1
	=	46	>	100			
4	>	100	>	100	>100	>100	n.a.
	>	100	>	100			
5		n.a.	=	0.90	n.a.	0.75	n.a.
		n.a.	=	0.61			
6 and 7		n.a.	=	5.34	n.a.	5.42	n.a.
		n.a.	=	5.51			
8		n.a.	=	11.93	n.a.	10.56	n.a.
		n.a.	=	9.19			
9	>	100	>	100	>100	>100	n.a.
	>	100	>	100			
11		n.a.	=	3.44	n.a.	3.40	n.a.
		n.a.	=	3.37			
12	=	66	>	100	80	>100	>1
	=	94	>	100			
13	=	47	>	100	47	>100	>2
	=	48	>	100			
18	=	94	>	100	≥97	>100	>1
	>	100	>	100			

n.a. = not applicable

5.6. References

Abou-Donia AH, Abib A-A, El Din AS, Evidente A, Gaber M, Scopa A, 1992. Two betaine-type alkaloids from Egyptian *Pancratium maritimum*. Phytochemistry. 31, 2139–2141

Avalon NE, Nafie J, De Marco Verissimo C, Warrensford LC, Dietrick SG, Pittman AR, Young RM, Kearns FL, Smalley T, Binning JM, Dalton JP, Johnson MP, Woodcock HL, Allcock AL, Baker BJ, 2022. Tuaimenal A, a meroterpene from the Irish deep-sea soft coral duva florida, displays inhibition of the SARS-CoV-2 3CLpro enzyme. J. Nat. Prod. 85, 1315–1323

Berkov S, Evstatieva L, Popov S, 2004. Alkaloids in Bulgarian *Pancratium maritimum* L. Zeitschrift für Naturforsch C. 59, 65–69

Campbell WE, Nair JJ, Gammon DW, Codina C, Bastida J, Viladomat F, Smith PJ, Albrecht CF, 2000. Bioactive alkaloids from *Brunsvigia radulosa*. Phytochemistry. 53, 587–591

Cedrón JC, Del Arco-Aguilar M, Estévez-Braun A, Ravelo ÁG, 2010. Chemistry and biology of *Pancratium* Alkaloids. Alkaloids Chem. Biol. 68, 1–37

Chhetri BK, Lavoie S, Sweeney-Jones AM, Kubanek J, 2018. Recent trends in the structural revision of natural products. Nat. Prod. Rep. 35, 514–531

Christy MP, Uekusa Y, Gerwick L, Gerwick WH, 2021. Natural products with potential to treat RNA virus pathogens including SARS-CoV-2. J. Nat. Prod. 84, 161–182

Cimmino A, Masi M, Evidente M, Superchi S, Evidente A, 2017. Amaryllidaceae alkaloids: Absolute configuration and biological activity. Chirality. 29, 486–499

De Castro O, Di Maio A, Di Febbraro M, Imparato G, Innangi M, Vela E, Menale B, 2016. A multi-faceted approach to analyse the effects of environmental variables on geographic range and genetic structure of a perennial psammophilous geophyte: The case of the sea daffodil *Pancratium maritimum* L. in the mediterranean basin. PLoS One. 11, 1–23

Ghosal S, S. Saini K, Razdan S, 1985. Crinum alkaloids: their chemistry and biology. Phytochemistry. 24, 2141–2156

Govaerts R, 2022. The World Checklist of Vascular Plants (WCVP). R. Bot. Gard.

Ibrahim SRM, Mohamed GA, Shaala LA, Youssef DTA, El Sayed KA, 2013. New alkaloids from *Pancratium maritimum*. Planta. Med. 79, 1480–1484

Ivens T, Van Den Eynde C, Van Acker K, Nijs E, Dams G, Bettens E, Ohagen A, Pauwels R, Hertogs K, 2005. Development of a homogeneous screening assay for automated detection of antiviral agents active against severe acute respiratory syndrome-associated coronavirus. J. Virol. Methods. 129, 56–63

Isidoro C, Chiung-Fang Chang A, Sheen LY, 2022. Natural products as a source of novel drugs for treating SARS-CoV2 infection. J. Tradit. Complement. Med. 12, 1–5

Jin YH, Min JS, Jeon S, Lee J, Kim S, Park T, Park D, Jang MS, Park CM, Song JH, Kim HR, Kwon S, 2021. Lycorine, a non-nucleoside RNA dependent RNA polymerase inhibitor, as potential treatment for emerging coronavirus infections. Phytomedicine. 86, 153440

Jochmans D, Leyssen P, Neyts J, 2012. A novel method for high-throughput screening to quantify antiviral activity against viruses that induce limited CPE. J. Virol. Methods. 183, 176–179

Kihara M, Konishi K, Xu L, Kobayashi S, 1991. Alkaloidal constituents of the flowers of *Lycoris radiata* HERB. (Amaryllidaceae). Jstage. 39, 1849–1853

Kaline R. Carvalho ABS, Maria Conceição M. Torres FCLP, Larissa A. Guimarães DDR, Edilberto R. Silveira LVC-L, Raimundo Braz-Filho and Otília Deusdênia. L. Pessoa, 2015. Cytotoxic alkaloids from *Hippeastrum solandriflorum* Lindl. J. Braz. Chem. Soc. 26, 1976–1980

Le HTN, Van Roy E, Dendooven E, Peeters L, Theunis M, Foubert K, Pieters L, Tuenter E, 2021. Alkaloids from *Lepidium meyenii* (Maca), structural revision of macaridine and UPLC-MS/MS feature-based molecular networking. Phytochemistry. 190, 112863

Le NTH, Vermeyen T, Aerts R, Herrebout A W, Pieters L, Tuenter E, 2023a. Epimeric mixture analysis and absolute configuration determination using an integrated spectroscopic and computational approach – a case study of two epimers of 6-hydroxyhippeastidine. Molecules. 28(1), 214

Le NTH, De Jonghe S, Erven K, Neyts J, Pannecouque C, Vermeyen T, Herrebout A W, Pieters L, Tuenter E, 2023b. Anti-SARS-CoV-2 activity and cytotoxicity of Amaryllidaceae alkaloids from *Hymenocallis littoralis*. Molecules. 28(7), 3222

Leporini M, Catinella G, Bruno M, Falco T, Tundis R, Loizzo MR, 2018. Investigating the antiproliferative and antioxidant properties of *Pancratium*

maritimum L. (Amaryllidaceae) stems, flowers, bulbs, and fruits extracts. Evidence-based Complement. Altern. Med. 2018

Li YT, Yang C, Wu Y, Lv JJ, Feng X, Tian X, Zhou Z, Pan X, Liu S, Tian LW, 2021. Axial chiral Binaphthoquinone and Perylenequinones from the stromata of *Hypocrella bambusae* are SARS-CoV-2 entry inhibitors. J. Nat .Prod. 84, 436–443

Marcarino MO, Cicetti S, Zanardi MM, Sarotti AM, 2022. A critical review on the use of DP4+ in the structural elucidation of natural products: the good, the bad and the ugly. A practical guide. Nat. Prod. Rep. 39, 58–76

Nair JJ, van Staden J, 2023. Antiviral alkaloid principles of the plant family Amaryllidaceae. Phytomedicine. 108, 154480

Masi M, Di Lecce R, Mérindol N, Girard M-P, Berthoux L, Desgagné-Penix I, Calabrò V, Evidente A, 2022. Cytotoxicity and antiviral properties of alkaloids isolated from *Pancratium maritimum*. Toxins (Basel). 14, 262

Raman K, Rajagopal K, Islam F, Dhawan M, Mitra S, Aparna B, Varakumar P, Byran G, Choudhary OP, Emran T Bin, 2022. Role of natural products towards the SARS-CoV-2: A critical review. Ann Med Surg. 80, 104062

Ren P xuan, Shang W juan, Yin W chao, Ge H, Wang L, Zhang X lei, Li B qian, Li H lin, Xu Y chun, Xu EH, Jiang H liang, Zhu L li, Zhang L ke, Bai F, 2022. A multi-targeting drug design strategy for identifying potent anti-SARS-CoV-2 inhibitors. Acta. Pharmacol. Sin. 43, 483–493

Rhizopoulou S, Pouris J, 2018. On *Pancratium maritimum* (sea daffodil, sea lily, sand lily). Hortic Int Journal. 2, 116–118

Richomme P, Pabuccuoglu V, Gozler T, Freyer A, Shamma M, 1989. (-)-siculinine: a lycorine-type alkaloid from *Sternbergia sicula*. 52, 1150–1152

Sarotti AM, 2020. In Silico Reassignment of (+)-Diplopyrone by NMR calculations: use of a DP4/ J-DP4/DP4+/DIP tandem to revise both relative and absolute configuration. J. Org. Chem. 85, 11566–11570

Sarikaya BB, Kaya GI, Onur MA, Bastida J, Somer NU, 2013. Phytochemical investigation of *Galanthus woronowii*. Biochem. Syst. Ecol. 51, 276–279

Southgate EH, Holycross DR, Sarlah D, 2017. Total synthesis of Lycoricidine and Narciclasine by chemical dearomatization of bromobenzene. Angew. Chemie – Int. Ed. 56, 15049–15052

Suau R, Gómez AI, Rico R, Tato MPV, Castedo L, Riguera R, 1988. Alkaloid N-oxides of amaryllidaceae. Phytochemistry. 27, 3285–3287

Wildman WC, Brown CL, 1968. The structure of habranthine. Tetrahedron Lett. 9, 4573–4576

Viladomat F, Bastida J, Codina C, Campbell WE, Mathee S, 1995. Alkaloids from *Boophane flava*. Phytochemistry. 40, 307–311

Youssef DTA, Frahm AW, 1998. Alkaloids of the flowers of *Pancratium maritimum*. Planta. Med. 64, 669–670

Youssef DTA, Shaala LA, Altyar AE, 2022. Cytotoxic phenylpropanoid derivatives and alkaloids from the flowers of *Pancratium maritimum* L. Plants. 11, 1–13

Zanardi MM, Sarotti AM, 2021. Sensitivity Analysis of DP4+ with the probability distribution terms: development of a universal and customizable Method. J. Org. Chem. 86: 8544–8548

Zotta BL, Gatti G, Fuganti C, 1971. 13C nucclear magnetic resonance spectra of Amaryllidaceaea alkaloids. 1–5

Zhang FM, Tu YQ, Liu JD, Fan XH, Shi L, Hu XD, Wang SH, Zhang YQ, 2006. A general approach to crinine-type Amaryllidaceae alkaloids: total syntheses of (±)-haemanthidine, (±)-pretazettine, (±)-tazettine, and (±)-crinamine. Tetrahedron. 62, 9446–9455

CHAPTER 6

Amaryllidaceae alkaloids from *Scadoxus multiflorus*

Publication:

Le Ngoc-Thao-Hien, Steven De Jonghe, Kristien Erven, Tom Vermeyen, Wouter A. Herrebout, Johan Neyts, Christophe Pannecouque, Luc Pieters, and Emmy Tuenter. "Comprehensive study of alkaloids from *Scadoxus multiflorus* by HPLC-PDA-SPE-NMR and evaluation of their anti-SARS-CoV-2 activity". Phytochemistry Letters, 57, 156-162.

Abstract: *Scadoxus multiflorus* (Martyn) Raf. (Amaryllidaceae) is traditionally used for the treatment of many respiratory problems (bronchitis, asthma, pneumonia, etc.). However, the alkaloidal composition of *S. multiflorus* has rarely been explored. In the present study, a full investigation of bulbs of *S. multiflorus* was carried out using HPLC-PDA-SPE-NMR, followed by structure elucidation using extensive spectroscopic techniques coupled with DFT calculation. With respect to the indigenous usage of this species for treating respiratory-related problems, antiviral activity against the severe acute respiratory syndrome-related coronavirus-2 (SARS-CoV-2) was assessed for all isolated compounds. As a result, a previously undescribed alkaloid, 9-*O*-demethyl-11-hydroxylgalanthine (or 2-*O*-methyl-11-hydroxypseudolycorine), together with fourteen known ones were identified. Nine alkaloids, i.e. 2-*O*-methylpseudolycorine, ungiminorine, ungiminorine *N*-oxide, 9-de-*O*-methyl-11β-hydroxygalanthamine, 2-hydroxy-*O,N*-dimethylnorbelladine, *O*-demethylmaritidine, 8-*O*-demethyloxomaritidine, narcissidine, and sanguinine, were reported for the first time in this species. In addition, also lycorine, narciclasine, lycoricidine, montanine and 3-*O*-acetylsanguinine were obtained. Results showed that 9-*O*-demethyl-11-hydroxylgalanthine, 8-*O*-demethyloxomaritidine, 3-*O*-acetylsanguinine and 9-de-*O*-methyl-11β-hydroxygalanthamine exhibited weak antiviral activity at non-cytotoxic concentrations (EC_{50} of 53, 70, 49, 62 μM, respectively). *S. multiflorus* appears to be a promising source of anti-SARS-CoV-2 Amaryllidaceae alkaloid scaffolds.

Keywords: Amaryllidaceae alkaloids; *Scadoxus multiflorus*; SARS-CoV-2; DFT calculation

6.1. Introduction

The genus *Scadoxus* Raf. (Amaryllidaceae), or "blood lily" genus, is native to Africa and comprises thirteen known species enlisted by the World Checklist of Vascular Plants (WCVP) (Govaerts, 2022). However, it is not among the most studied Amaryllidaceae genera (*Crinum*, *Narcissus*, *Galanthus*, *Hippeastrum*, *Lycoris*, *Pancratium* and *Zephyranthes*) (Berkov et al., 2020). Up to 2021, phytochemical records were available for only four species, namely *S. multiflorus*, *S. pseudocaulus*, *S. puniceus* and *S. membranaceus* (Das et al., 2022).

S. multiflorus (Martyn) Raf. (synonym: *Haemanthus multiflorus* Martyn) is the most widespread *Scadoxus* species (Kwembeya, 2021; Das et al., 2022). According to the comprehensive review of Das at el. and WCVP (2022), its geographical distribution covers Africa (Southern Africa, Egypt, Ethiopia, Senegal, Cameroon, Gambia, Somalia, Togo, Kenya, Guinea, Tanzania, Zanzibar, Zambia, Uganda, Benin, Namibia), America (Mexico, Brazil, Ecuador, Panama, Puerto Rico) and Asia (Saudi Arabia, Oman, Yemen, India, Bangladesh, China, Vietnam). However, this plant species is commercially traded in many other countries, due to its ornamental value. It is traditionally used for the treatment of asthma, cough, bronchitis, pneumonia, gastrointestinal problems, as well as mental illness, and for wound healing (Teklehaymanot, 2009; Das et al., 2022). Concerning its phytochemical profile, alkaloids, phenols and isoflavonoids of *S. multiflorus* have been explored for their pharmacological properties, including antimicrobial, antioxidant, antimiotic, AChE inhibitory, BuChE inhibitory, antitumoral, and immunostimulatory activities (Das et al., 2021; Nair and van Staden, 2022).

The severe acute respiratory syndrome-related coronavirus-2 (SARS-CoV-2) emerged in late 2019 and caused the death of millions of people worldwide. Using medicinal plants as complementary and alternative treatment remains a common practice in many places around the globe, alongside vaccination (Nair et al., 2021; Raman et al., 2022; Nair and van Staden, 2023). Nevertheless, this practice is mainly based on the expectation of a broad-spectrum antiviral activity of plants historically used for the treatment of antimicrobial and antiviral diseases, and little is known about their bioactive constituents, as well as their pharmacological targets (Owen et al., 2022).

The alkaloid lycorine is widely spread in the Amaryllidaceae family. In the course of 2020-2022, it was reported as a highly potent anti-SARS-CoV-2 agent with an EC_{50} ranging from 0.18 - 0.87 µM (Zhang et al., 2020; Jin et al., 2021). This has drawn scientific attention to Amaryllidaceae alkaloids as potential natural products for the treatment of SARS-CoV-2 infections (Berkov et al., 2020). In addition, the Amaryllidaceae family is renowned for plant-based antiviral drug discovery (Nair and van Staden, 2023). To the best of our knowledge, thirteen alkaloids were isolated from *S. multiflorus* up to the end of 2022, consisting of one of the lycorine-type (lycorine), three of the crinine- and haemanthamine-type (haemanthamine, haemanthidine, haemultine), three of the galanthamine-type (galanthamine, 3-*O*-acetyl sanguinine, chlidanthine), five of the narciclasine-type (narciclasine, kalbreclasine, kalbretorine, hippadine, pratorimine) and one of the lycorenine-type (hippeastrine). Tentative identification of undulatine and montanine by GC/MS was also reported by Cahlikova (Cahlíková et al., 2011). However, since the last isolation work of alkaloids from this plant species dates back to 1989, the alkaloidal composition was re-investigated in the current study, using the state-of-the-art hyphenated separation technique HPLC-PDA-SPE-NMR, which is recognized for its power to discover minor components (Chen et al., 2011; Tuenter et al., 2017; Lima et al., 2017). Apart from applying conventional approaches, DFT calculation was also utilized to assist the 3D structure elucidation process (Le et al. 2021; Le et al. 2023a). In view of the traditional practice to treat respiratory-related diseases with *S. multiflorus*, and reported antiviral activities of Amaryllidaceae alkaloids, including anti-SARS-CoV-2 activity of lycorine, the anti-SARS-CoV-2 activity was assessed for all isolated compounds, and structure-activity relationships were examined.

6.2. Results and Discussion

6.2.1. Structure Elucidation

The current phytochemical investigation led to the isolation and characterization of fifteen alkaloids of varying polarities. Compounds **2** – **6**, **9** – **12**, **14** – **15** were purified from the DCM extract; compounds **7** and **8** from the EtOAc extract and compounds **1** and **13** from the remaining water phase. Compound **1** was described

in this study for the first time and the structural elucidation process will be elaborated on hereafter.

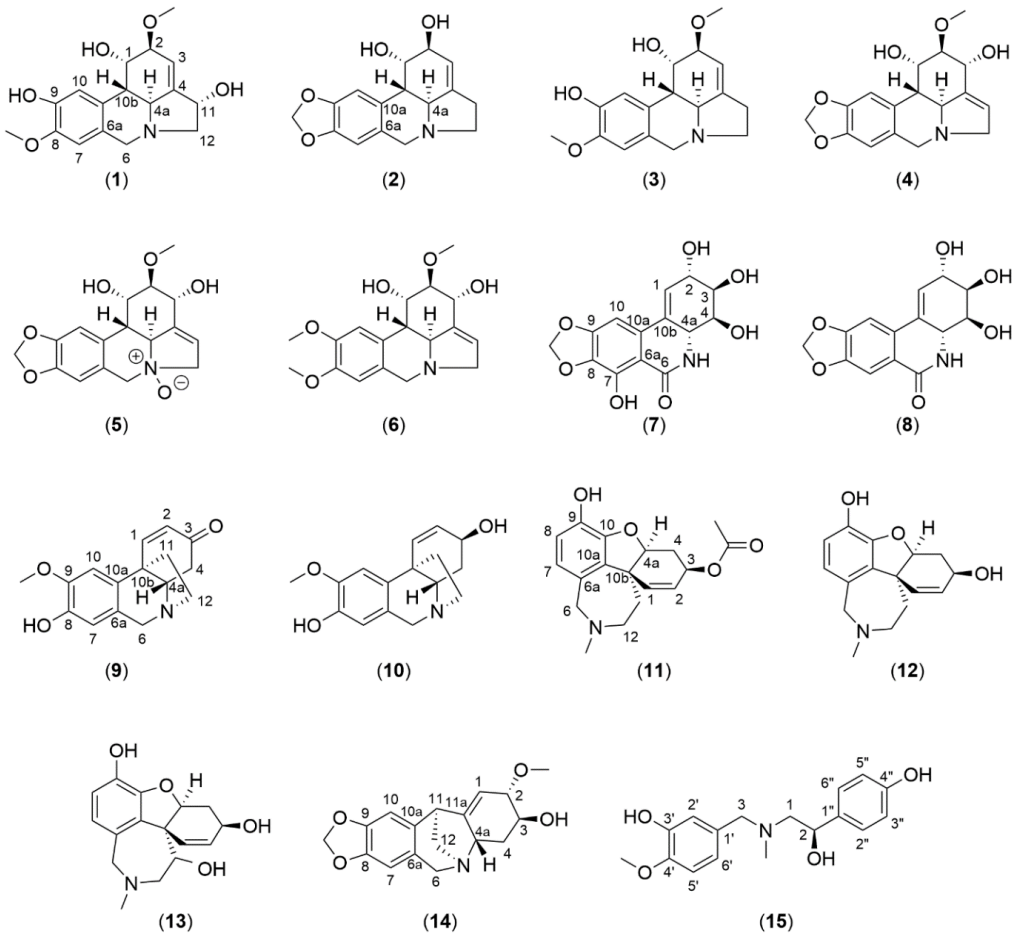


Fig. 1: Alkaloids isolated from *S. multiflorus*: 9-*O*-demethyl-11-hydroxylgalanthine (**1**), lycorine (**2**), 2-*O*-methylpseudolycorine (**3**), ungiminorine (**4**), ungiminorine *N*-oxide (**5**), narcissidine (**6**), narciclasine (**7**), lycoricidine (**8**), 8-*O*-demethyloxomaritidine (**9**), 8-*O*-demethylmaritidine (**10**), 3-*O*-acetylsanguinine (**11**), sanguinine (12), 9-de-*O*-methyl-11β-hydroxygalanthamine (**13**), montanine (**14**) and 2-hydroxy-*O*,*N*-dimethylnorbelladine (**15**).

In total, six lycorine-type alkaloids were purified (**Fig. 1**). Firstly, compound **2** crystalized directly during fractionation of the DCM extract as white needles and was identified as lycorine, the most common Amaryllidaceae alkaloid. Compound

3 was isolated as a yellowish amorphous solid and its ^{1}H-NMR spectrum was similar to that of compound **2**. The most noticeable discrepancy was the absence of the dioxymethylene signal around 6.0 ppm (-O-CH_2-O-) and signals of 2 methoxy groups appearing instead. Analysis of the 2D NMR spectra and comparison with literature confirmed the structure of compound **3** as 2-*O*-methylpseudolycorine (Kobayashi et al., 1977). Compounds **4** – **6** were isolated from the same fraction and their ^{1}H-NMR spectra showed similar patterns. They differed from compounds **2** and **3** in the upfield region, as no shielded signals could be observed between 2.0 - 2.5 ppm. This was explained by the position of the double bond between C-3 and C-4 for compounds **2** and 3 (H-11 and H-12 were the most shielded hydrogens and resonated at 2.0 – 2.5 ppm), and between C-4 and C-11 for compounds **4** – **6** (**Fig. 1**). Comparison of experimental and reported data confirmed compounds **4** – **6** as ungiminorine (Richomme et al., 1989), ungiminorine *N*-oxide (Suau et al., 1988), and narcissidine (Pigni et al., 2012), respectively. Experimental NMR and MS data of compounds **2** – **6** can be found as Supplementary Material.

Compound **1** was purified as a brownish amorphous solid with a chemical formula of $C_{17}H_{21}NO_5$, deduced by ESI-HRMS, showing a $[M+H]^+$ peak at *m/z* 320.1498. In comparison with compounds **2** - **6**, the ^{1}H-NMR spectrum of compound 1 showed a similar pattern, typical of a lycorine-type skeleton (Cedrón et al., 2010) (Table 1). Similar to compounds **3** and **6**, the dioxymethylene signal was not observed, indicating a different substitution at positions 8 and 9. The absence of signals in the upfield region (2.0 - 2.5 ppm) was reminiscent of compounds **4** – **6**, but careful analysis of the 2D NMR spectra revealed that the double bond must be located between C-3 and C-4, like in compounds **2** and **3**. The deshielding effect on H-11 and H-12 was due to substitution, resulting in a downfield shift of their resonances (> 2.5 ppm).

Table 1. ^{1}H- and ^{13}C-NMR spectroscopic data (400 MHz, CD_3OD) of compound 1

Position	**δ_C, type**	**δ_H, multiplicity (*J* in Hz)**	**COSY**	**HMBC (H→C)**	**NOESY**
1	67.1, CH	4.67, brs	H-10b, H-2	C-2, C-3, C-4a	H-10, H-3, 2-OCH_3
2	82.0, CH	3.84*	H-1, H-3	C-3, C-4	-

3	121.8, CH	5.98, brs	H-2, H-4a	C-4a	H-11, H-4a, 2-OCH_3
4	143.9, C	-	-	-	-
4a	60.7, CH	3.71, d (11.0)	H-10b	-	H-6α, H-2, H-12α
6α	55.4, CH_2	4.15, d (13.5)	H-6β	C-4a, C-6a, C-10a, C-7	H-12α, H-4a, H-7
6β	55.4, CH_2	4.03, d (13.6)	H-6α	C-4a, C-6a, C-10a, C-7, C-8	H-12β, H-10b, H-7
6a	125.5, C	-	-	-	-
7	112.1, CH	6.88, s	-	C-6, C-9, C-10a	H-6β, H-6α, 8-OCH_3
8	147.8, C	-	-	-	-
9	147.6, C	-	-	-	-
10	112.9, CH	6.97, s	-	C-8, C-6a, C-10b	H-1, H-10b
10a	129.6, C	-	-	-	-
10b	40.0, CH	2.66, d (11.0)	H-4a, H-1	C-6a, C-4a	H-6β, H-10
11	72.1, CH	4.81, m	H-12α, H-12β	C-4a, C-3	H-3
12α	63.2, CH_2	3.19, dd (11.0, 3.2)	H-11, H-12β	C-11, C-4, C-6	H-6α, H-4a
12β	63.2, CH_2	3.49, dd (11.0, 4.8)	H-11, H-12α	C-11, C-4, C-11	H-6β
8-OCH_3	56.7, CH_3	3.86, s	-	C-8	H-7
2-OCH_3	58.1, CH_3	3.54, s	-	C-2	H-1, H-3

*overlapping signals

In the COSY spectrum, two spin systems were observed for compound **1**: the first one between the ^{1}H-NMR signals at 3.71, 2.66, 4.67, 3.84 and 5.98 ppm; and the second one between the ^{1}H-NMR signals at 4.81, 3.49 and 3.19 ppm (**Fig. 2**, **Table 1**). In the HSQC and DEPT spectra, two methylene (-CH_2-, δ_H 3.49 and 3.19 ppm; δ_H 4.15 and 4.03 ppm) and three unsaturated CH-signals (δ_H 6.97, 6.88 and 5.98 ppm) were observed. It was obvious that the signals at 6.97 and 6.88 ppm were due to aromatic hydrogens, and therefore, the signal at 5.98 ppm must be due to the hydrogen attached to a double bond. Since the signal at 5.98 ppm is part of a 5-membered spin system, as mentioned above, the double bond must be located between C-3 and C-4, like in compounds **2** and **3**. Also five CH-moieties with

hydrogens attached to sp^3 carbons (δ_H 3.71, 2.66, 4.67, 3.84 and 4.81 ppm) and two methoxy groups (δ_H 3.86 ppm and δ_C 56.7 ppm; δ_H 3.54 ppm and δ_C 58.1 ppm) were observed in the HSQC and DEPT spectra.

Finally, HMBC correlations (listed in **Table 1**) were used to connect all moieties (**Fig. 2**). Starting from the aromatic ring, the correlation between the signal at δ_H 6.97 ppm and δ_C 40.0 ppm, inferred that the signal at δ_H 6.97 ppm could be attributed to H-10, since the latter signal (δ_C 40.0 ppm, C-10b) showed an HSQC correlation with the signal at δ_H 2.66 ppm (H-10b), which was included in the 5-membered spin system, as discussed above. An HMBC correlation between the δ_H and δ_C signals at 2.66 ppm and 125.5 ppm (C-6a), respectively, further confirmed the assignment of H-10b. Next, the ^{13}C-NMR signal of a methylene at δ_C 55.4 ppm (C-6) suggested a nitrogenated carbon, and therefore, the HMBC correlation between the signals at δ_H 6.88 ppm and δ_C 55.4 ppm (-CH_2-, δ_H 4.15 and 4.03 ppm) confirmed the assignment of the δ_H signal at 6.88 ppm as H-7. A carbon signal at 67.1 ppm suggested an oxygenated carbon and since its hydrogen (δ_H 4.67 ppm) showed a COSY correlation with H-10b and a NOESY correlation with H-10, it was assigned as C-1. Based on COSY, the δ_C signal at 82.0 ppm was assigned as C-2, since COSY cross-peaks were observed between H-2, H-1 and H-3. The last carbon of the 5-membered spin system was assigned as C-4a because it resonates at 60.7 ppm, inferring an attachment to nitrogen, and because H-4a showed a COSY correlation with H-10b (2.66 ppm). HMBC correlations observed for H-1 with C-2, C-3, C-4a, and for H-4a with C-1 and C-3 confirmed the proposed assignments.

Lastly, the ^{13}C-NMR signals at 63.2 and 72.1 ppm were assigned to carbons attached to a nitrogen and to an oxygen, i.e. C-12 and C-11, respectively. The 3-bond HMBC correlations between H-11 and C-4a, C-3, and between H-12 and C-4 were in agreement with the assignments. Regarding the two methoxy groups, one showed an HMBC correlation with C-2 and the other one with either C-8 or C-9 of the aromatic ring. A NOESY correlation between H-7 and the second methoxy signal aided in the assignment of this substituent at C-8 (**Fig. 2**). In addition, the reported ^{1}H- and ^{13}C-NMR data of 11-hydroxygalanthine support this structural assignment (Pigni et al., 2012). In comparison with 11-hydroxygalanthine, compound **1** lacks the methoxy group in position 9 and the name 9-*O*-demethyl-

11-hydroxylgalanthine (or 2-*O*-methyl-11-hydroxypseudolycorine) was adopted for compound **1**.

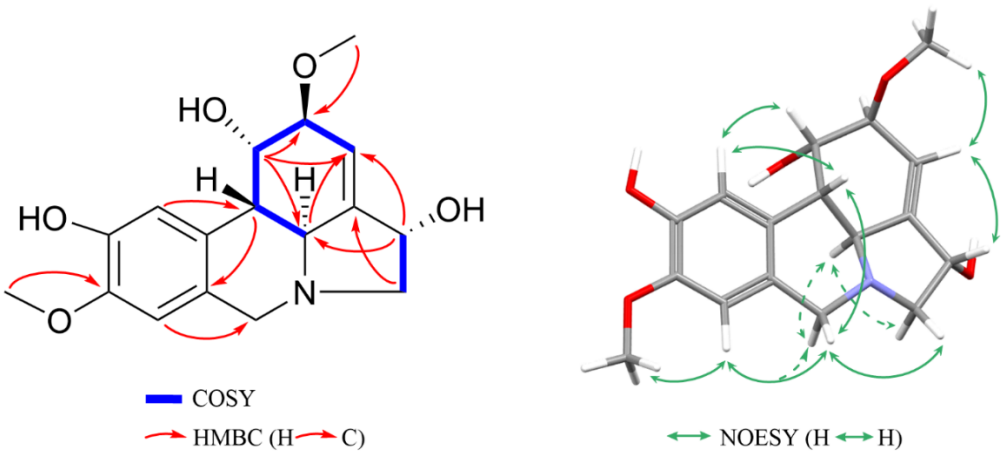


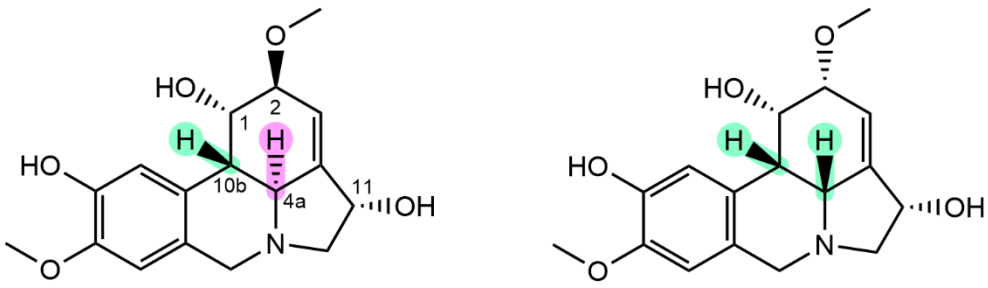
Fig. 2. Key COSY, HMBC and NOESY correlations of compound 1

Coplanar hydrogens were observed by NOESY as demonstrated in **Fig. 2**. Key correlations were observed between H-4a, H-6α, and H-12α, indicating their orientation on the same side of the plane. Similarly, NOESY correlations between H-6β, H-10b, and H-12β also indicate their coplanarity. Spatial proximity was seen for H-3 with H-11 and H-3 with 2-OCH_3. Since no NOESY cross-peaks could be observed for H-2, it was deduced that H-2 and H-10b could be on two opposite sides. Next, the α-orientation of 11-OH was supported by the allylic coupling between by H-11 and H-3. The observed *J*-value of approximately 1.0 Hz (splitting not observed) is consistent with the dihedral angle of about 30° between H-11 and the plane formed by C-11/C-4/C-3, and therefore, according to the Garbisch rules (Garbisch, 1964; Barfield and Chakrabarti, 1969), the β-orientation of H-11 is most likely. If H-11 would have an α-orientation, a 3D model showed that the dihedral angle would be about 75°, which should yield a bigger *J*-coupling (2.0-3.0 Hz – clear splitting). This was previously also observed for 11-hydroxygalanthine (Pigni et al., 2012). So far, the relative configuration (1*S*,2*S*,4a*S*,11*S*,10b*S*) appeared to be the most likely configuration as deduced from NOE effects and the allylic coupling. To obtain additional evidence, implementation of DFT calculation of chemical shifts and *J*-couplings was of use. Computational results as summarized in **Table 2**

pinpointed two possible configurations (**Fig. 3**): (1*S*,2*S*,4a*S*,11*S*,10b*S*) if only ^{1}H NMR data are considered and (1*S*,2*R*,4a*R*,11*S*,10b*S*) if only ^{13}C NMR data are considered, although the overall highest probability when combining both data was found for the SSSSS. At this point, simulation of *J*-coupling ascertained the SSSSS configuration as the best fit since the $J_{H\text{-}4a/H\text{-}10b}$ matched with the corresponding experimental value of 11.0 Hz (**Table 1**), while that of the SRRSS did not. This also inferred that H-4a and H-10b must be in *trans*-orientation to produce a large *J*-coupling of around 10 Hz, whereas a *cis*-orientation would lead to a *J*-coupling of approximately 5 Hz (**Fig. 3**), aligning entirely with reported observations (Cedrón et al., 2010; Le et al., 2023). Hence, combining all data, the relative configuration (1*S*,2*S*,4a*S*,11*S*,10b*S*) was proposed for compound **1**.

Table 2. DP4+ probabilities and computed *J*-couplings of different configurations of compound **1** estimated using methanol as solvation model. Stereocenters order is (1,2,4a,11,10b).

Diastereomer	DP4+ (H data)	DP4+ (C data)	DP4+ (all data)	$J_{H\text{-}4a/H\text{-}10b}$
SRRRS	0.00	0.00	0.00	-
SRRSR	0.00	0.17	0.00	-
SRSRS	0.00	0.00	0.00	-
SRSSR	0.00	0.00	0.00	-
SSRRS	0.00	0.00	0.00	-
SSRSR	0.00	0.00	0.00	-
SSSSR	0.00	0.01	0.00	-
SSSSS	99.76	0.13	100.00	10.9
SRRRR	0.00	0.00	0.00	-
SRRSS	0.00	99.69	0.00	5.8
SRSRR	0.00	0.00	0.00	-
SRSSS	0.02	0.00	0.00	-
SSRRR	0.00	0.00	0.00	-
SSRSS	0.00	0.00	0.00	-
SSSRR	0.00	0.00	0.00	-
SSSRS	0.22	0.00	0.00	-



(1*S*,2*S*,4a*S*,11*S*,10b*S*) **(1*S*,2*R*,4a*R*,11*S*,10b*S*)**

Fig. 3: Two configurations with the highest DP4+ probability and their spatial orientations of H-4a and H-10b: *trans*-orientation in case of (1*S*,2*S*,4a*S*,11*S*,10b*S*) and *cis*-orientation in case of (1*S*,2*R*,4a*R*,11*S*,10b*S*)

Calculation of optical activity for the SSSSS and the RRRRR resulted in the $[\alpha]_D^{25}$ values of -71.4 and 71.4, respectively. This does not align with the experimental value of -6.7 (c 0.66, MeOH), and therefore 9-*O*-demethyl-11-hydroxylgalanthine (**1**) isolated in this study might exist as a racemic mixture. Detailed computed data can be found as Supplementary Information.

^{1}H-, ^{13}C-NMR and MS data of compounds **7** – **15** matched with those of compounds reported in literature and they could be identified as narciclasine (**7**) (Southgate et al., 2017), lycoricidine (**8**) (Yadav et al., 2010), 8-*O*-demethyloxomaritidine (**9**) (Chen et al., 2011), 8-*O*-demethylmaritidine (**10**) (Kihara et al., 1987), 3-*O*-acetylsanguinine (**11**) (Machocho et al., 2004), sanguinine (**12**) (Kobayashi et al., 1976), 9-de-*O*-methyl-11β-hydroxygalanthamine (**13**) (Zhan et al., 2016), montanine (**14**) (Labrana et al., 2002) and 2-hydroxy-*O,N*-dimethylnorbelladine (**15**) (Kitajima et al., 2009), respectively. Experimental NMR and mass spectroscopic data of compounds **7** – **15** can be found as Supplementary Material. Apart from compound **1**, nine known alkaloids 2-*O*-methylpseudolycorine (**3**), ungiminorine (**4**), ungiminorine *N*-oxide (**5**), 9-de-*O*-methyl-11β-hydroxygalanthamine (**13**), 2-hydroxy-*O,N*-dimethylnorbelladine (**15**), 8-*O*-demethylmaritidine (**10**), 8-*O*-demethyloxomaritidine (**9**), narcissidine (**6**), sanguinine (**12**)) were isolated from *S. multiflorus* for the first time (**Fig. 1**).

6.2.2. Anti-SARS-CoV-2 Activity and Cytotoxicity

The fifteen alkaloids were evaluated as potential inhibitors of SARS-CoV-2 replication in VeroE6 cells, using a high-content imaging read-out. In parallel, the cytotoxicity of the compounds for the VeroE6 cells was determined using a 3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2H-tetrazolium (MTS) cell viability assay. Weak inhibition of SARS-CoV-2 replication on Vero E6 cells was observed for 9-*O*-demethyl-11-hydroxylgalanthine (**1**), 8-*O*-demethyloxomaritidine (**9**), 3-*O*-acetylsanguinine (**11**) and 9-de-*O*-methyl-11-hydroxygalanthamine (**13**) at non-cytotoxic concentrations (CC_{50} > 100 µM) with EC_{50} values of 53, 70, 49, and 62 µM, respectively. Three compounds did not display activity, nor cytotoxicity: ungiminorine (**4**), narcissidine (**6**) and 2-hydroxy-*O,N*-dimethylnorbelladine (**15**). 2-*O*-methylpseudolycorine (**3**), ungiminorine *N*-oxide (**5**) and sanguinine (**12**) showed weak activity in one experiment, but were not active in the duplicate.

The cytotoxicity of lycorine (**2**) on a Vero E6 cell line was previously reported by the authors in a study of alkaloids from *Hymenocallis littoralis* (Le et al., 2023). Apart from lycorine, it was observed in the current study that narciclasine (**7**), lycoricidine (**8**), *O*-demethylmaritidine (**10**) and montanine (**14**) exhibited cytotoxicity with CC_{50} values of <0.046, 0.331, 8.800, and 0.550 µM, respectively. Narciclasine was found to be the most cytotoxic compound in this study. However, in order to have a broader view on the anti-SARS-CoV-2 potential of Amaryllidaceae alkaloids, research on a large and chemically diverse library of compounds is required. Detailed results of the *in vitro* anti-SARS-CoV-2 activity and cytotoxicity can be found as Supporting Information.

6.3. Conclusion

For centuries, *S. multiflorus* has been used as folk medicine in many countries, but its phytochemical constituents, especially alkaloids, were underexplored. In this study, the alkaloidal fraction of the bulbs of *S. multiflorus* was investigated, which led to the isolation and characterization of fifteen alkaloids belonging to six skeleton-types: lycorine-type (compounds **1** – **6**), narciclasine-type (compounds **7** – **8**), haemanthamine-type (compounds **9** – **10**), galanthamine-type (**11** – **13**),

montanine-type (compound **14**) and norbelladine-type (compound **15**). 9-*O*-Demethyl-11-hydroxylgalanthine (**1**) (or 2-*O*-methyl-11-hydroxypseudolycorine) was an undescribed lycorine-type alkaloid, and nine alkaloids 2-*O*-methylpseudolycorine (**3**), ungiminorine (**4**), ungiminorine *N*-oxide (**5**), narcissidine (**6**), *O*-demethylmaritidine (**9**), 8-*O*-demethyloxomaritidine (**10**), sanguinine (**12**), 9-de-*O*-methyl-11β-hydroxygalanthamine (**13**) and 2-hydroxy-*O,N*-dimethylnorbelladine (**15**) were reported for the first time in this plant species. Screening of SARS-CoV-2 inhibition revealed several weakly active compounds, i.e. 9-*O*-demethyl-11-hydroxylgalanthine (**1**), 8-*O*-demethyloxomaritidine (**9**), 3-*O*-acetylsanguinine (**11**) and 9-de-*O*-methyl-11β-hydroxygalanthamine (**13**) with EC_{50} values ranging from 49 to 70 µM at non-cytotoxic concentration. Besides, narciclasine (**7**), lycoricidine (**8**), *O*-demethylmaritidine (**10**) and montanine (**14**) displayed cytotoxicity with CC_{50} values of <0.046, 0.331, 8.800, and 0.550 µM, respectively.

6.4. Materials and Methods

6.4.1. Isolation and Purification

Fresh bulbs of *S. multiflorus* (4.2 kg) were collected in July 2017 in Dubréka (Guinea-Conakry) and identified by prof. dr. Aliou Baldé, Centre de Recherche et de Valorisation des Plantes Médicinales de Dubréka, where a voucher specimen of *S. multiflorus* was kept (no. M140HK2). After cutting by a blender, the plant material was freeze-dried to obtain approximately 350 g dried bulbs.

300 g freeze-dried bulbs were macerated in 80% MeOH (2.5 L x 4) with continuous stirring to obtain 80 g crude extract. Liquid-liquid partitioning following the general protocol of alkaloid extraction was carried out thereafter. Briefly, the crude extract was suspended in 900 mL H_2O and was acidified by HCl 10% to pH < 3 before partitioning with DCM (1 L x 3). The acidic water phase was then basified to pH > 9 by NH_4OH 25% which was later on subsequently extracted by DCM (1 L x 4) and EtOAc (1 L x 4) to obtain the first two alkaloidal fractions. The residual water phase was freeze-dried to obtain a polar fraction (WT extract).

Fractionation by flash chromatography was performed for the DCM and EtOAc extracts which resulted in 18 subfractions (DCM-1 to DCM-18) and 9 subfractions

(ET-1 to ET-9), respectively. A commercially packed Claricep flash cartridge containing 40 g irregular silica gel was used for the DCM extract and a 24 g cartridge was used for the EtOAc extract (Agela Technologies, Wilmington, DE, USA). Gradient elution using a solvent mixture of DCM with increasing percentage of MeOH + 0.5% NH_4OH was applied for both flash runs, and the gradient was held while compounds were eluting.

After preliminary HPLC profiling, fractions of interest were selected for further purification. Except for lycorine (**2**), which crystalized directly in fractions DCM-4 and DCM-5, compounds **3** – **6, 14 - 15** were isolated from fraction DCM-4, whereas compounds **9** – **12** were purified from fraction DCM-8. Additional amounts of compounds **4** – **6** were also found in fraction DCM-6. Compounds **7** – **8** were obtained from fraction ET-2.

On the other hand, the WT extract was applied on a Diaion-HP20 open column. The elute was manually collected from subsequent elution with 20% MeOH, 40% MeOH, 60% MeOH, 80% MeOH and 100% MeOH (each 2.0 L) to obtain 10 subfractions of 1 L (WT_Di-F1 to WT_Di-F10). After evaporating solvents under reduced pressure by a rotary evaporator, all fractions were also profiled by the HPLC-PDA-SPE-NMR system. Next, compounds **1** and **13** were purified from fraction WT_Di-F6. All trapping chromatograms and HPLC gradients can be found as Supplementary Material.

9-O-Demethyl-11-hydroxylgalanthine (**1**)

Brownish amorphous solid (6.0 mg); UV λ_{max} 210, 290 nm; $[\alpha]_D^{25}$ -6.7 (c 0.66, MeOH); ^1H and ^{13}C NMR in CD_3OD (400 and 100 MHz): see **Table 1**. Positive HRESIMS *m/z* 320.1498 $[M+H]^+$ (calcd for $C_{17}H_{21}NO_5$, 319.1419).

6.4.2. Computational Details

Conformational analysis was performed by PCMODEL (version 10.0) using the MMFF94 force field. All conformers within an energy window of 10 kcal.mol^{-1} were selected. Geometry optimization and frequency calculation were carried out subsequently by Gaussian16 using methanol as solvation model. After dereplication, the B3LYP/6-31G(d)//mPW1PW91/6-311+G(d,p) level was applied

to simulate shielding tensors which were later on used as inputs for DP4+ probability (Grimblat et al., 2015). Coupling constants were simulated in gas-phase at the B3LYP/6-31G(d)//B3LYP/6-31G(d,p) level (Bally et al., 2011). To compute specific optical rotation: B3LYP/6-31++G(d,p)//aug-cc-pVTZ was used. Detailed protocols for combining experimental and computed data to assist the structure elucidation process were previously elaborated on (Le et al, 2023a; Le et al., 2023b).

6.5. Supporting information – Chapter 6

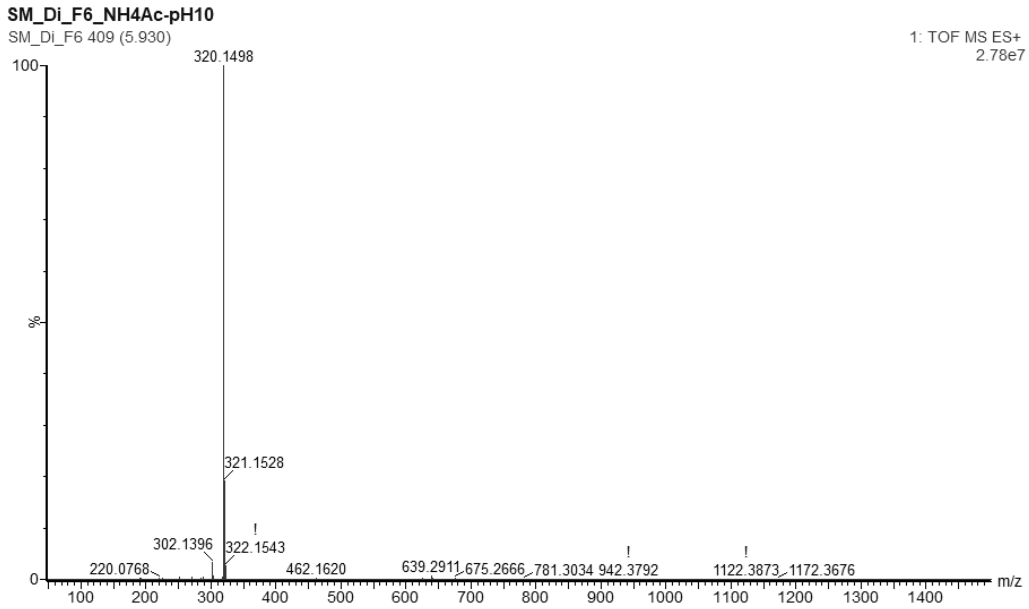


Fig. S1. HRMS spectrum of compound **1**

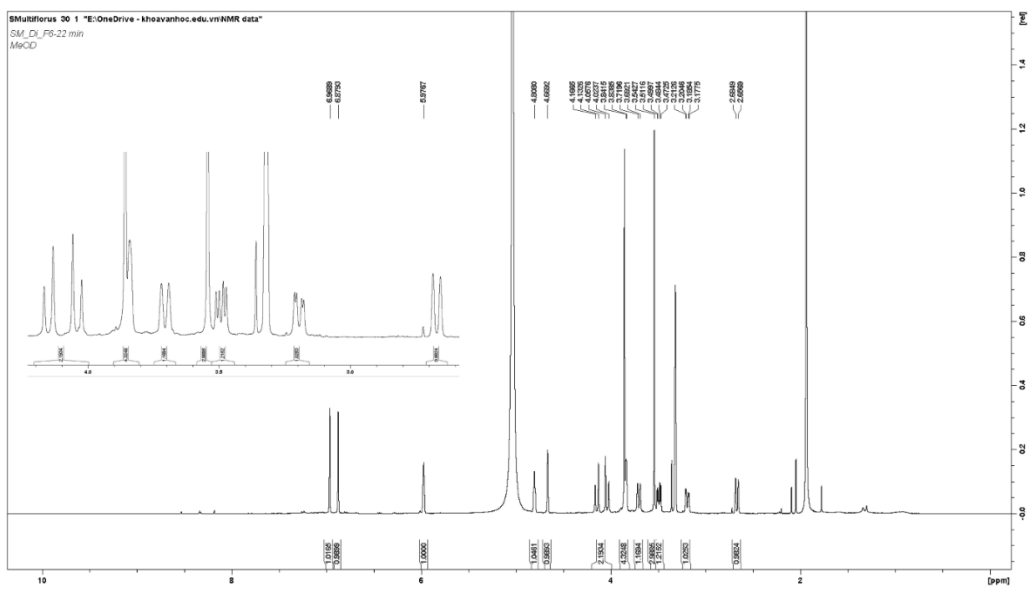


Fig. S2. ^{1}H spectrum of compound **1** in CD_3OD

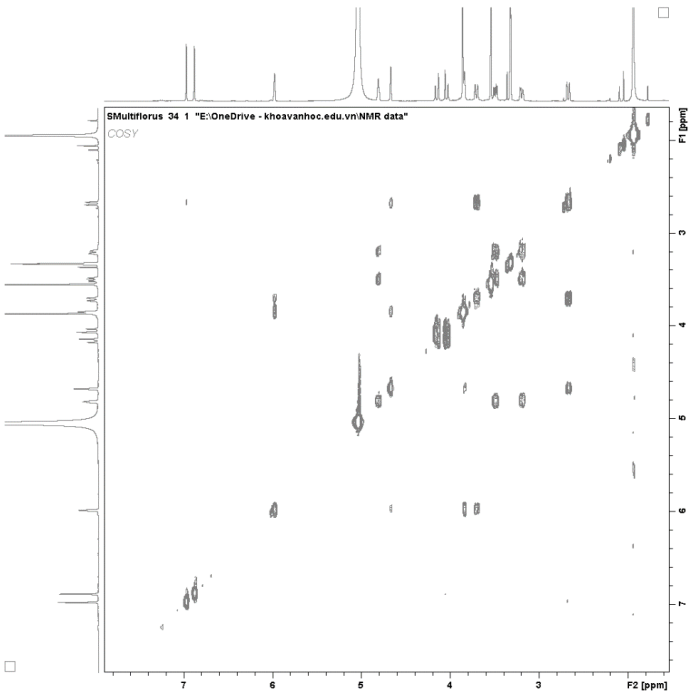


Fig. S2. COSY spectrum of compound **1** in CD_3OD

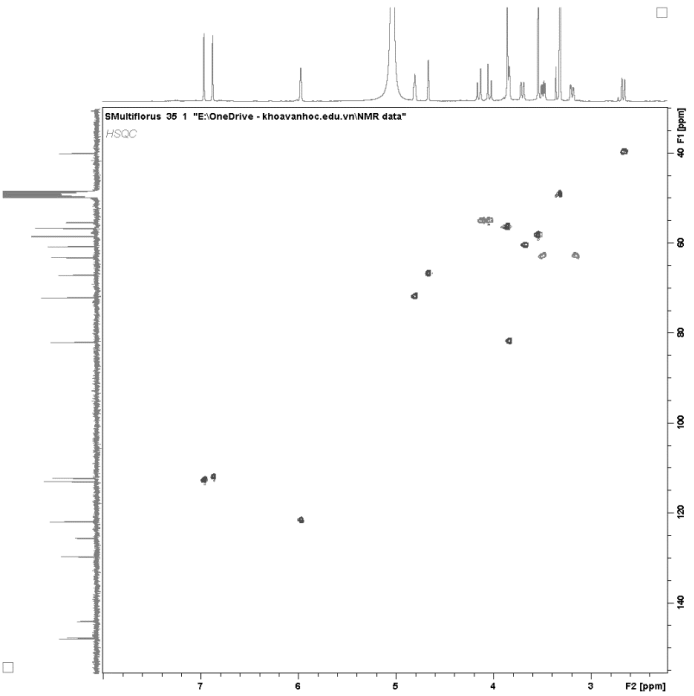


Fig. S4. HSQC spectrum of compound **1** in CD_3OD

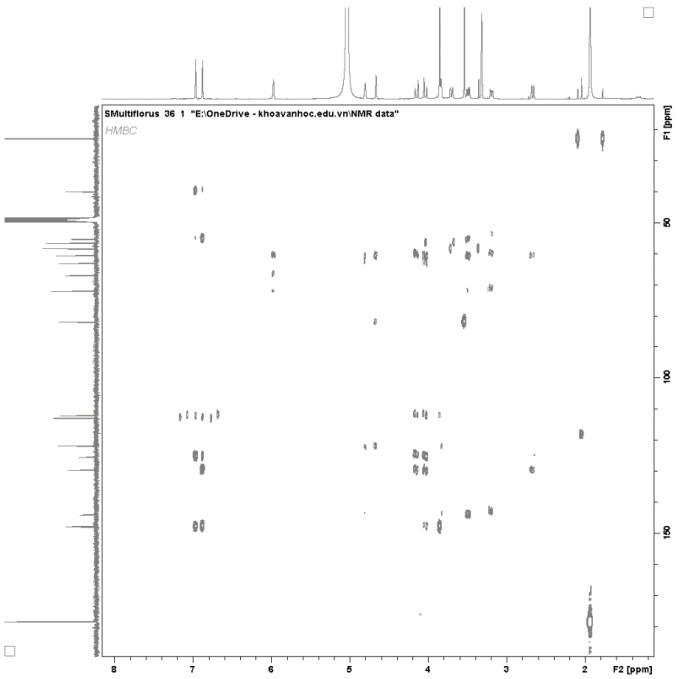


Fig. S5. HMBC spectrum of compound **1** in CD_3OD

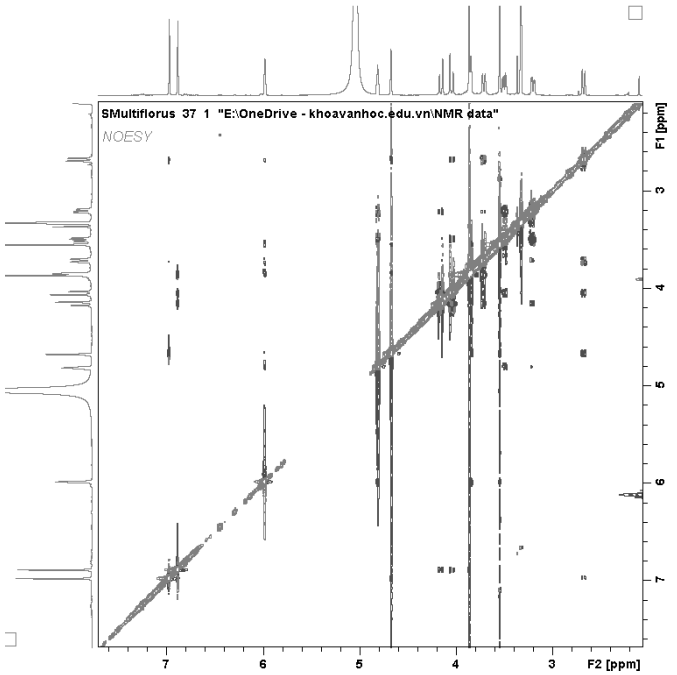


Fig. S6. NOESY spectrum of compound **1** in CD_3OD

Table S1. Anti-SARS-CoV-2 activity and cytotoxicity on Vero E6-EGFP cells

Compounds		**EC_{50} (µM)**		**CC_{50} (µM)**	**av. EC_{50} (µM)**	**av. CC_{50} (µM)**	**SI**
1	=	57	>	100	53	>100	>2
	=	48	>	100			
3	=	65	>	100	≥82	>100	>1
	>	100	>	100			
4	>	100	>	100	>100	>100	n.a.
	>	100	>	100			
5	=	92	>	100	≥96	>100	n.a.
	>	100	>	100			
6	>	100	>	100	>100	>100	n.a.
	>	100	>	100			
7		n.a.	<	0.05	n.a.	<0.05	n.a.
		n.a.	<	0.05			
8		n.a.	=	0.37	n.a.	0.33	n.a.
		n.a.	=	0.28			
9	=	63	>	100	70	>100	>1
	=	77	>	100			
10		n.a.	=	9.63	n.a.	8.80	n.a.
		n.a.	=	7.97			
11	=	58	>	100	49	>100	>2
	=	40	>	100			
12	=	86	>	100	92	>100	n.a.
	=	99	>	100			
13	=	61	>	100	62	>100	>1
	=	63	>	100			
14		n.a.	=	0.61	n.a.	0.55	n.a.
		n.a.	=	0.48			
15	=	84	>	100	≥92	>100	n.a.
	>	100	>	100			

n.a. = not applicable

6.6. References

Bally, T.; Rablen, P. R. Quantum-Chemical Simulation of 1H NMR Spectra. 2. Comparison of DFT-based procedures for computing proton-proton coupling constants in organic molecules. J. Org. Chem. 2011, 76 (12), 4818–4830.

Barfield, M., Chakrabarti, B., 1969. Long-range proton spin-spin coupling. Chem. Rev. 69, 757–778.

Berkov, S., Osorio, E., Viladomat, F., Bastida, J., 2020. Chapter Two - Chemodiversity, chemotaxonomy and chemoecology of Amaryllidaceae alkaloids, in: Knölker, H.-J.B.T.-T.A.C. and B. (Ed.). Academic Press, pp. 113–185.

Boudewijns, R., Thibaut, H.J., Kaptein, S.J.F., 2020. STAT2 signaling restricts viral dissemination but drives severe pneumonia in SARS-CoV-2 infected hamsters. Nat. Commun. 11, 1–10.

Cahlíková, L., Benešová, N., Macáková, K., Urbanová, K., Opletal, L., 2011. GC/MS analysis of three Amaryllidaceae species and their cholinesterase activity. Nat. Prod. Commun. 6, 1934578X1100600912.

Cedrón, J.C., Del Arco-Aguilar, M., Estévez-Braun, A., Ravelo, Á.G., 2010. Chemistry and biology of Pancratium alkaloids. Alkaloids Chem. Biol. 68, 1–37.

Chen, C.K., Lin, F.H., Tseng, L.H., Jiang, C.L., Lee, S.S., 2011. Comprehensive study of alkaloids from Crinum asiaticum var. sinicum assisted by HPLC-DAD-SPE-NMR. J. Nat. Prod. 74, 411–419.

Das, R., Goswami, S., Barman, A., Barman, M., Ray, S., 2021. Antimitotic and cytotoxic effects of leaf aqueous extract of scadoxus multiflorus (martyn) raf. In allium cepa root apical meristem cells. Cytologia (Tokyo). 86, 267–271.

Das, R., Barman, A., Ray, S., 2022. Traditional knowledge, phytochemistry, and pharmacological properties of the African genus Scadoxus (Amaryllidaceae). South African J. Bot. 151, 565–577.

Garbisch, E.W., 1964. Conformations. VI. Vinyl-allylic proton spin couplings. J. Am. Chem. Soc. 86, 5561–5564.

Grimblat, N.; Zanardi, M. M.; Sarotti, A. M., 2015. Beyond DP4: An improved probability for the stereochemical assignment of isomeric compounds using quantum chemical calculations of NMR shifts. J. Org. Chem. 80 (24), 12526–12534.

Govaerts, R., 2022. The World Checklist of Vascular Plants (WCVP). Royal Botanic Gardens, Kew. (accessed via GBIF.org on 2023-01-12).

Ivens, T., Van Den Eynde, C., Van Acker, K., Nijs, E., Dams, G., Bettens, E., Ohagen, A., Pauwels, R., Hertogs, K., 2005. Development of a homogeneous screening assay for automated detection of antiviral agents active against severe acute respiratory syndrome-associated coronavirus. J. Virol. Methods 129, 56–63.

Jin, Y.H., Min, J.S., Jeon, S., Lee, J., Kim, S., Park, T., Park, D., Jang, M.S., Park, C.M., Song, J.H., Kim, H.R., Kwon, S., 2021. Lycorine, a non-nucleoside RNA dependent RNA polymerase inhibitor, as potential treatment for emerging coronavirus infections. Phytomedicine 86, 153440.

Kihara, M., Koike, T., Imakura, Y., Kida, K., Kobayashi, S., Shingu, T., 1987. Alkaloidal constituents of Hymenocallis rotata herb. (Amaryllidaceae). Chem. Pharm. Bull. 35, 1070–1075.

Kitajima, M., Kinoshita, E., Kogure, N., 2009. Two new alkaloids from bulbs of Lycoris squamigera. Heterocycles 77, 1389–1396.

Kobayashi, S., Takeda, S., Ishikawa, H., Matsumoto, H., Kihara, M., Shingu, T., Numata, A., Uyeo, S., 1976. Alkaloids of the Amaryllidaceae. A new alkaloid, sanguinine, from Lycoris sanguinea Maxim. var. Kiushiana Makino, and pretazettine from Lycoris radiata Herb. J-STAGE. 24, 1537–1543.

Kobayashi, S., Ishikawa, H., Kihara, M., Shingu, T., Hashimoto, T., 1977. Isolation of carinatine and pretazettine from the bulbs of Zephyranthes carinate Herb. (Amaryllidaceae). Jstage. Chem. Pharm. Bull. 25(9), 2244–2248.

Kwembeya, E.G., 2021. Tracking biological footprints of climate change using flowering phenology of the geophytes: Pancratium tenuifolium and Scadoxus multiflorus. Int. J. Biometeorol. 65, 577–586.

Labrana, J., Machocho, A.K. ori, Kricsfalusy, V., Brun, R., Codina, C., Viladomat, F., Bastida, J., 2002. Alkaloids from Narcissus angustifolius subsp. transcarpathicus (Amaryllidaceae). Phytochemistry 60, 847–852.

Le, N. T. H., Van Roy, E., Dendooven, E., Peeters, L., Theunis, M., Foubert, K., Pieters, L., Tuenter, E., 2021. Alkaloids from Lepidium meyenii (Maca), structural revision of macaridine and UPLC-MS/MS feature-based molecular networking. Phytochemistry, 190, 112863.

Le, N.T.H., De Jonghe, S., Erven, K., Vermeyen, T., Baldé, A.M., Herrebout, W.A., Neyts, J., Pannecouque, C., Pieters, L., Tuenter, E., 2023a. Anti-SARS-CoV-2 activity and cytotoxicity of Amaryllidaceae alkaloids from Hymenocallis littoralis. Molecules, 28(7), 3222.

Le, N. T. H., Vermeyen, T., Aerts, R., Herrebout, W. A., Pieters, L., Tuenter, E., 2023b. Epimeric mixture analysis and absolute configuration determination using an integrated spectroscopic and computational approach — A case study of two epimers of 6-hydroxyhippeastidine. Molecules, 28(1), 214.

Lima, R. de C.L., Gramsbergen, S.M., Van Staden, J., Jäger, A.K., Kongstad, K.T., Staerk, D., 2017. Advancing HPLC-PDA-HRMS-SPE-NMR analysis of coumarins in Coleonema album by use of orthogonal reversed-phase C18 and pentafluorophenyl separations. J. Nat. Prod. 80, 1020–1027. https://doi.org/10.1021/acs.jnatprod.6b01020

Machocho, A.K., Bastida, J., Codina, C., Viladomat, F., Brun, R., Chhabra, S.C., 2004. Augustamine type alkaloids from Crinum kirkii. Phytochemistry 65, 3143–3149.

Nair, M.S., Huang, Y., Fidock, D.A., Polyak, S.J., Wagoner, J., Towler, M.J., Weathers, P.J., 2021. Artemisia annua L. extracts inhibit the in vitro replication of SARS-CoV-2 and two of its variants. J. Ethnopharmacol. 274, 114016.

Nair, J.J., van Staden, J., 2022. Chemical principles of Boophone, Nerine, Crossyne, Clivia, Cryptostephanus, Haemanthus and Scadoxus of the South African Amaryllidaceae and their biological properties. Planta Med. Nair, J.J., van Staden, J., 2023. Antiviral alkaloid principles of the plant family Amaryllidaceae. Phytomedicine 108, 154480.

Owen, L., Laird, K., Shivkumar, M., 2022. Antiviral plant-derived natural products to combat RNA viruses: Targets throughout the viral life cycle. Lett. Appl. Microbiol. 75, 476–499.

Pigni, N.B., Ríos-Ruiz, S., Martínez-Francés, V., Nair, J.J., Viladomat, F., Codina, C., Bastida, J., 2012. Alkaloids from Narcissus serotinus. J. Nat. Prod. 75, 1643–1647.

Raman, K., Rajagopal, K., Islam, F., Dhawan, M., Mitra, S., Aparna, B., Varakumar, P., Byran, G., Choudhary, O.P., Emran, T. Bin, 2022. Role of natural products towards the SARS-CoV-2: A critical review. Ann. Med. Surg. 80, 104062.

Richomme, P., Pabuccuoglu, V., Gozler, T., Freyer, A., Shamma, M., 1989. (-)-siculinine: a lycorine-type alkaloid from Sternbergia sicula. J. Nat. Prod. 52 (5), 1150–1152.

Southgate, E.H., Holycross, D.R., Sarlah, D., 2017. Total synthesis of lycoricidine and narciclasine by chemical dearomatization of bromobenzene. Angew. Chemie - Int. Ed. 56, 15049–15052.

Suau, R., Gómez, A.I., Rico, R., Tato, M.P.V., Castedo, L., Riguera, R., 1988. Alkaloid N-oxides of amaryllidaceae. Phytochemistry 27, 3285–3287.

Teklehaymanot, T., 2009. Ethnobotanical study of knowledge and medicinal plants use by the people in Dek Island in Ethiopia. J. Ethnopharmacol. 124, 69–78.

Tuenter, E., Foubert, K., Staerk, D., Apers, S., Pieters, L., 2017. Isolation and structure elucidation of cyclopeptide alkaloids from Ziziphus nummularia and Ziziphus spina-christi by HPLC-DAD-MS and HPLC-PDA-(HRMS)-SPE-NMR. Phytochemistry 138, 163–169.

Yadav, J.S., Satheesh, G., Murthy, C.V.S.R., 2010. Synthesis of (+)-lycoricidine by the application of oxidative and regioselective ring-opening of aziridines. Org. Lett. 12, 2544–2547.

Zhan, G., Zhou, J., Liu, R., Liu, T., Guo, G., Wang, J., Xiang, M., Xue, Y., Luo, Z., Zhang, Y., Yao, G., 2016. Galanthamine, plicamine, and secoplicamine alkaloids from Zephyranthes candida and their anti-acetylcholinesterase and anti-inflammatory activities. J. Nat. Prod. 79, 760–766.

Zhang, Y.N., Zhang, Q.Y., Li, X.D., Xiong, J., Xiao, S.Q., Wang, Z., Zhang, Z.R., Deng, C.L., Yang, X. Lou, Wei, H.P., Yuan, Z.M., Ye, H.Q., Zhang, B., 2020. Gemcitabine, lycorine and oxysophoridine inhibit novel coronavirus (SARS-CoV-2) in cell culture. Emerg. Microbes Infect. 9, 1170–1173.

CHAPTER 7

Alkaloids from *Lepidium meyenii* (Maca)

Publication:

Le Ngoc-Thao-Hien, Elias Van Roy, Ella Dendooven, Laura Peeters, Mart Theunis, Kenn Foubert, Luc Pieters, Emmy Tuenter. "Alkaloids from *Lepidium meyenii* (Maca), structural revision of macaridine and UPLC-MS/MS feature-based molecular networking". Phytochemistry, Volume 190, 2021, 112863, ISSN 0031-9422.

Abstract: *Lepidium meyenii* Walp., known as Peruvian ginseng, is widely used in ethnomedicine. To date, *L. meyenii* is cultivated worldwide at high-altitude and is commonly used as a food supplement. However, its medicinal value is still controversial and its mechanism of action remains unknown, due to limited knowledge about the phytochemical constituents of this plant species. In this study, a multidisciplinary approach comprising conventional NMR- and HRMS-based structure elucidation, quantum mechanical calculation of NMR chemical shifts and UPLC-MS/MS feature-based molecular networking was applied to analyse the phytochemical profile of *L. meyenii*. In the current work, three previously undescribed imidazole alkaloids were identified using extensive spectroscopic techniques (HRMS, NMR), for which the names lepidiline E, F and G were adopted. In addition, two amidine alkaloids were reported, representing an undescribed class of alkaloids in *L. meyenii*, and 1,2,3,4-tetrahydro-*β*-carboline-3-carboxylic acid, a well-known *β*-carboline alkaloid, was also isolated from *L. meyenii* for the first time. Molecular networks of imidazole, amidine and *β*-carboline alkaloids in *L. meyenii* were constructed by the Global Natural Products Social Molecular Networking (GNPS) web platform, resulting in the tentative identification of three undescribed analogues. In addition, the structure of a previously reported compound named 'macaridine' was revised as macapyrrolin C based on density functional theory (DFT) calculations and comprehensive comparison of NMR data.

Keywords: *Lepidium meyenii*; Brassicaceae; lepidilines; amidine alkaloids; *β*-carboline alkaloids; DFT calculation; feature-based molecular networking

7.1. Introduction

Lepidium meyenii Walp. (Brassicaceae), known as "Maca", originates from the Central Andes in Peru. For many centuries, it has been cultivated because of the nutritional value of the root and for various medicinal purposes. Nowadays, it is widely sold worldwide as a dietary supplement (sometimes referred to as Peruvian Ginseng) to increase vitality and longevity, and most notably, to enhance fertility (Beharry and Heinrich, 2018). Characteristic constituents, which have been proposed as chemical markers, include polyunsaturated fatty acids (macaenes) and their amides (macamides) (Ganzera et al., 2002). In recent years, increasing attention has been paid to various other classes of constituents, such as *β*-carboline, pyrrole and imidazole alkaloids, glucosinolates, thioamides, hydantoins and thiohydantoins (Huang et al., 2018) (Carvalho and Ribeiro, 2019). According to Carvalho and Ribeiro (2019), 26 macamides and macaenes, 9 glucosinolates and 2 isothiocyanates were reported in the timeframe 2000-2018. In 2017-2018, a number of papers was published in which the alkaloid fraction was investigated in more detail, including pyrrole alkaloids (Zhou et al., 2017), thiohydantoin alkaloids (Zhou et al., 2017) (Yu et al., 2017), urea alkaloids (Kitamura et al., 2017), hexahydroimidazo[1,5-c]thiazole alkaloids (Zhou et al., 2017), hydantoin and thioamide alkaloids (Geng et al., 2018) (Tian et al., 2018). Not less than 26 alkaloids belonging to these skeleton types were reported; they can be considered as minor constituents because mostly large amounts of dried plant material (10-25 kg) were needed to obtain around 1-20 mg pure compound.

Although imidazole alkaloids, biogenetically derived from histidine, commonly exist in nature as tertiary alkaloids, two quaternary imidazole alkaloids, *i.e*. lepidilines A and B, were isolated from *L. meyenii* for the first time (Cui et al., 2003). Their atypical structures were elucidated by NMR spectroscopy and confirmed by X-ray crystallography. Afterwards, lepidilines C and D, which are methoxylated derivatives of lepidilines A and B, were reported by Jin et al. (2016). To the best of our knowledge, this type of quaternary imidazole alkaloid exists exclusively in *L. meyenii*, indicating that lepidilines can be used as chemical markers of this species. Interestingly, one of the main imidazole alkaloids, *i.e.* lepidiline A, was recently proposed as a potential active ingredient of Maca to enhance fertility (Cheng et al.,

2020). This may indicate that the quaternary imidazole alkaloids could contribute to its sexual and fertility enhancing properties, together with the macamides and macaenes. Because of this unique quaternary imidazole moiety, it is important to further explore the chemical profile of the alkaloid fraction of *L. meyenii*. Therefore, in the present work, a detailed study of the alkaloidal profile of *L. meyenii* was carried out by means of NMR spectroscopy, UPLC-MS/MS and feature-based molecular networking, resulting in the isolation and identification of several previously undescribed imidazole alkaloids. In addition, the occurrence of amidine alkaloids was reported for the first time in Maca, and a structural revision of macaridine is proposed, based on density functional theory (DFT) calculations.

7.2. Results and Discussion

7.2.1. Structure Elucidation

Seven quaternary imidazole alkaloids were isolated from *L. meyenii*, three of which are reported here for the first time (**1** – **3**), whereas four were known (**4** – **7**) (Fig.1).

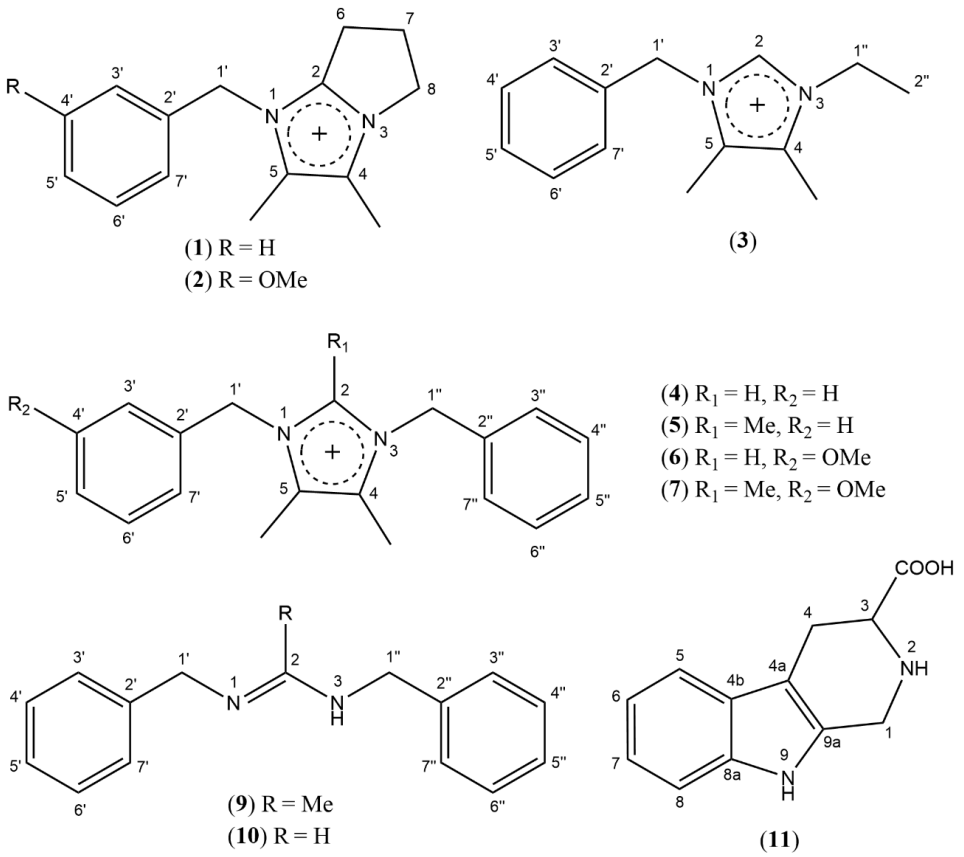


Fig. 1. Structure of isolated imidazole, amidine and *β*-carboline alkaloids

Comparison of the NMR spectra of compounds **4**, **5**, **6** and **7** with reported data confirmed their identity as lepidilines A-D, respectively (Cui et al., 2003; Jin et al., 2016). Compounds **4** and **5** showed NMR patterns typical for benzyl moieties, while compounds **6** and **7** contained both benzyl and methoxy-benzyl moieties. Compounds **4** and **5** produced simple ^{1}H NMR spectra with only a few (3-5) signals, but their symmetrical structures caused some difficulties in taking correct integrals. The only difference between compounds **4** and **5** is the methyl group at position 2 of compound **5**, which induces deshielding on the 4,5-Me groups. This is also observed for compounds **6** and **7**. Besides, the presence of a 4'-OMe group also produces significant changes in the UV spectra. The 4'-OMe functionality acts as an auxochrome attached to a benzene ring and causes a bathochromic shift from λ_{max} 230 to λ_{max} 274 nm. Therefore, compounds **4** and **5** only show an absorption maximum at λ_{max} 230 nm, but compounds **6** and **7** at both λ_{max} 230 and

λ_{max} 274 nm. This was crucial for monitoring the separation of compounds **5** and **6** (the most difficult couple to separate) and for purification based on the UV spectra. In the HMBC spectra, key cross-peaks of the four compounds occur between H-1' and C-2, C-5, C-2', C-3' and C-7'; and between H-1" and C-2, C-4, C-2", C-3" and C-7". It should be noted that for compounds **4** and **6** the H-2 signal disappeared in the ^{1}H and 2D NMR spectra when $CDCl_3$ and CD_3OD were used as solvent. However, it could be observed in DMSO-d_6; the assignment of H-2 was further confirmed by HSQC (cross-peak between H-2 and C-2) and HMBC correlations (cross-peaks between H-2 and C-4, C-5, C-1' and C-2').

The NMR spectra of compound **1** showed the typical NMR pattern of a benzyl moiety, and those of compound **2** of a meta-methoxy-benzyl moiety as previously described for compounds **4 – 7**. Apart from this, other signals in their ^{1}H NMR spectra were identical. Hence, compounds **1** and **2** possess the same backbone, and compound **2** (yielding a molecular ion in MS at *m/z* 257.1653) bears an additional methoxy group compared to compound **1** (*m/z* 227.1553). In ^{1}H NMR, the benzyl methylenes of compounds **1** and **2** were observed within the same chemical shift range (~5.35-5.45 ppm) as compounds **4** – **7**, which are imidazole alkaloids. Therefore, it was hypothesized that compounds **1** and **2** also possessed an imidazole ring. Nevertheless, in contrast to compounds **4** – **7**, compounds **1** and **2** contained a bicyclic structure, consisting of an imidazole ring fused to a cyclopentane ring, based on the following observations: (1) three -CH_2- functionalities were evident from the DEPT-135 spectrum; (2) the -CH_2-CH_2-CH_2- moiety was confirmed by COSY correlations and by the coupling pattern in ^{1}H NMR, i.e. a triplet (*J* ~ 7.5 Hz), a quintet (*J* ~ 7.5 Hz) and a triplet (*J* ~ 7.5 Hz), each with an integration of two; (3) all three -CH_2- moieties showed HMBC correlations with C-2 of the imidazole ring, indicating a 2- or 3-bond distance; (4) one head of the -CH_2-CH_2-CH_2- functionality is connected to a nitrogen atom (N-3), as one -CH_2- has a remarkably more downfield shift in ^{1}H and ^{13}C NMR (H-8, ~4.20 ppm; C-8, ~47.0 ppm) than the other two -CH_2- units, due to its connection to an electronegative element; this was further confirmed by a HMBC correlation between H-8 and C-4; (5) the other head of the -CH_2-CH_2-CH_2- functionality is linked to C-2, as the chemical shift of C-2 is 151.6 ppm, indicating it must be substituted. Finally, the correct assignment of 4-Me and 5-Me was deduced from NOE effects between H-

1’ and 5-Me and between H-8 and 4-Me. All NMR assignments are listed in Table 1, and key correlations are shown in Fig. 2. The structures proposed for compounds **1** and **2** were confirmed by HRMS. In UV spectroscopy, compound **1** showed a λ_{max} at 230 nm and compound **2** at both 230 and 274 nm. Compounds **1** and **2** are reported for the first time, and the names lepidiline E and F, respectively, were adopted.

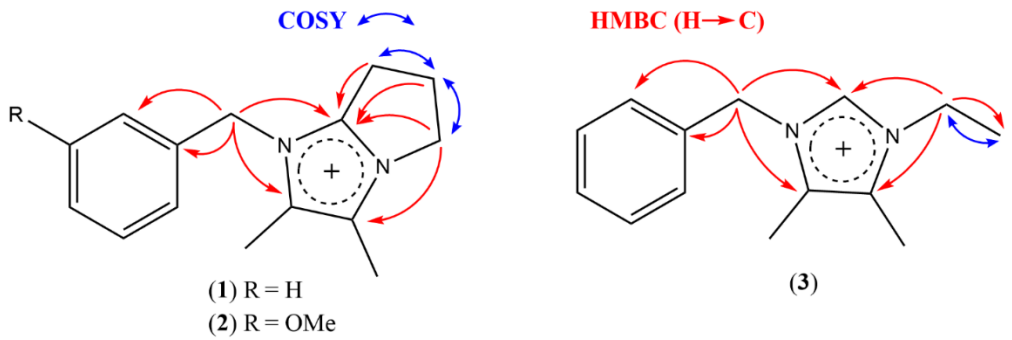


Fig. 2. Key COSY and HMBC correlations observed for lepidiline E (**1**) F (**2**) and G (**3**)

Compound **3** was isolated as a mixture with compound **1**. Thus, NMR analysis of the mixture was carried out, of which compound **1** was the major and compound **3** the minor compound (ratio 5:1). Similar to compound **1**, the NMR pattern of a benzyl moiety was observed for compound **3** and its benzyl methylene resonated at 5.37 ppm, suggesting the possible presence of an imidazole ring. Structure elucidation of compound **3** started with the HMBC correlations of H-1’ (benzyl methylene), since this peak did not overlap with proton signals of compound **1**. H-1’ correlated with carbons at 128.7 and 134.5 ppm, similar to H-1’ and H-1” of compound **4** and H-1” of compound **6**. Therefore, it was deduced that compound **3** contains an imidazole ring bearing a benzyl group on one of its two nitrogen atoms. The ^{13}C NMR signal at 128.7 ppm was then assigned to C-3’, C-7’ and C-4; the signal at 134.5 ppm was assigned to C-2, indicating that this position is not substituted. Two methyl groups attached to the imidazole ring were assigned as follows: (1) a methyl group was observed in the ^{1}H NMR spectrum at 2.31 ppm with an integral of three; (2) a second methyl group of compound **3** was observed at 2.19 ppm, but overlapped with the 5-Me signal of compound **1**, as indicated by its integral. The peak area of this signal equals the sum of the peak areas of two

methyl groups at 2.25 ppm (4-Me of compound **1**) and 2.31 ppm (a methyl-group of compound **3**). Finally, it was supported by a HMBC correlation with the ^{13}C NMR signal at 128.7 ppm (C-4). The correct assignment of these two methyl groups was confirmed by NOE effects. Likewise, the ethyl group attached to a nitrogen of the imidazole ring was elucidated as follows: (1) a triplet at 1.52 ppm with an integral of three, which showed a COSY correlation with a signal at 4.20 ppm, indicating an overlap with a signal of compound **1** with the same chemical shift; (2) the peak area of the signal at 4.20 ppm equals the sum of the peak areas of a methylene of the major compound **1** and a methylene of the minor compound **3**; (3) HMBC cross-peaks between the methylene of the minor compound **1** with ^{13}C NMR signals at 134.5 ppm (C-2), 128.7 ppm (C-5) and most importantly, at 15.0 ppm (CH_3 of the ethyl group). Taken altogether, compound **3** was identified as 1-benzyl-3-ethyl-4,5-dimethylimidazolium. NMR assignments are listed in Table 1, and key correlations are shown in Fig. 2. HRESIMS yielded a molecular ion at *m/z* 215.1550 $[M]^+$, confirming the chemical formula $C_{14}H_{19}N_2$ of the proposed structure. As a remark, H-2 was only observed in DMSO-*d*$_6$ at 9.71 ppm, but not observed in MeOH-*d*$_4$. Compound **3** is reported for the first time, and the name lepidiline G was adopted.

Table 1. ^1H and ^{13}C NMR assignments of compounds **1** – **3** (lepidilines E, F and G), recorded in CD_3OD

Position	**Compound 1**		**Compound 2**		**Compound 3**	
	δ_H (mult., *J* in Hz)	δ_C	δ_H (mult., *J* in Hz)	δ_C	δ_H (mult., *J* in Hz)	δ_C
1	-	-	-	-	-	-
2	-	151.6	-	151.6	n.o.	134.5
3	-	-	-	-	-	-
4	-	124.9	-	124.6	-	128.9
5	-	131.1	-	131.1	-	128.7
6	3.08 (t, 7.6)	24.5	3.08 (t, 7.6)	24.4	-	-
7	2.75 (quintet, 7.5)	26.1	2.75 (quintet, 7.4)	25.9	-	-
8	4.20 (t, 7.4)	47.2	4.19 (t, 7.4)	46.9	-	-
1’	5.29 (s)	50.6	5.26 (s)	50.2	5.37 (s)	51.5
2’	-	135.1	-	135.8	-	135.1
3’	7.28 *	128.6	6.83 *	114.4	7.30 *	128.7

4’	7.39-7.45 *	130.4	-	161.5	7.37-7.46 *	130.4
5’	7.39-7.45 *	129.9	6.97 (dd; 8.2, 2.2)	114.6	7.37-7.46 *	129.9
6’	7.39-7.45 *	130.4	7.34 (t, 8.0)	131.3	7.37-7.46 *	130.4
7’	7.28 *	128.6	6.81 *	120.4	7.30 *	128.7
1’’	-	-	-	-	4.20 *	43.4
2’’	-	-	-	-	1.52 (t, 7.35)	15.0
4-Me	2.25 (s)	8.1	2.25 (s)	7.8	2.31 (s)	8.0
5-Me	2.19 (s)	8.8	2.19 (s)	8.6	2.19 (s)	8.4
4’-OMe	-	-	3.81	55.5	-	-

*overlapping signals; n.o.: not observed

Table 2. ^{1}H and ^{13}C NMR assignments of compounds **9** and **10** recorded in $CDCl_3$

Position	**Compound 9**		**Compound 10**			
	Major isomer		**Major isomer**		**Minor isomer**	
	δ_H (mult., *J* in Hz)	δ_C	δ_H (mult., *J* in Hz)	δ_C	δ_H (mult., *J* in Hz)	δ_C
1	-	-	-	-	-	-
2	-	168.0	7.28-7.38 *	168.7	7.28-7.38 *	166.3
3	11.00 (brs)	-	10.89 (brs)	-	8.79 (brs)	-
1’	4.52 (s)	47.2	4.50 (s)	46.7	4.65 (s)	46.9
2’	-	135.0	-	135.1	-	133.9
3’	7.31 *	126.7	7.27 *	127.1	7.26 *	127.9
4’	7.35-7.40 *	129.2	7.28-7.38 *	129.0	7.28-7.38 *	129.0
5’	7.35-7.40 *	128.3	7.28-7.38 *	129.0	7.28-7.38 *	129.0
6’	7.35-7.40 *	129.2	7.28-7.38 *	129.0	7.28-7.38 *	129.0
7’	7.31 *	126.7	7.27	127.1	7.26 *	127.9
1”	4.53 (s)	47.2	3.78 (s)	36.2	4.07 (s)	38.7
2”	-	135.0	-	129.9	-	131.5
3”	7.31 *	126.7	7.17	129.1	7.34-7.39 *	129.5
4”	7.35-7.40 *	129.2	7.28-7.38 *	129.0	7.28-7.38 *	129.0
5”	7.35-7.40 *	128.3	7.28-7.38 *	129.0	7.28-7.38 *	129.0
6”	7.35-7.40 *	129.2	7.28-7.38 *	129.0	7.28-7.38 *	129.0
7”	7.31 *	126.7	7.17	129.1	7.34-7.39 *	129.5
2-Me	2.04 (s)	13.07	-	-	-	-

*overlapping signals

The NMR spectra of compounds **9** and **10** again indicated the presence of two benzyl moieties. Their elemental composition was established as $C_{16}H_{18}N_2$ (*m/z* 239.1549, $[M+H]^+$) and $C_{15}H_{16}N_2$ (*m/z* 225.1402, $[M+H]^+$), respectively, based on HRESIMS. Since the two benzyl moieties alone already account for 182 Da, along with 28 Da for two nitrogen atoms, the remaining part of the structure of compounds **9** and **10** must correspond to 27 and 13 Da, respectively, which indicated the presence of a -C-CH_3 and a -CH- moiety, respectively. In the HMBC spectra, cross-peaks of the benzyl methylenes (H-1' and H-1") with deshielded carbon signals around 163-169 ppm in ^{13}C NMR inferred that these carbons are positioned between the two nitrogen atoms, forming an -N-C=N- backbone. This is in agreement with several synthetic amidine analogues reported in literature (Maccallini et al., 2015). Therefore, the structure of compound **9** was elucidated as an acetamidine derivative, i.e. *N,N'*-dibenzylacetamidine, and compound **10** as a formamidine derivative, i.e. *N,N'*-dibenzylformamidine (Fig. 1). In organic chemistry, amidine analogues are quite common, but to the best of our knowledge, this is the first time amidine alkaloids are reported as plant-derived natural products.

In literature, considerable efforts have been made to clarify the chemistry of acyclic amidines, and in non-aqueous solutions, amidines exhibit tautomerism with respect to the position of the double bond in the -N-C=N- moiety and *Z-E* isomerism with respect to the C=N double bond (Minkin and Mikhailov, 2010). As for compounds **9** and **10**, it appears that *Z-E* isomerism was observed when changing the solvent from DMSO-d_6 to $CDCl_3$. In DMSO-d_6, only the *E*-isomer exists, as the *E*-isomer is energetically preferred over the *Z*-isomer (mainly due to the steric effect of the benzyl moieties). This phenomenon in $CDCl_3$ or $CHCl_3$ was previously described for amidine derivatives (Minkin and Mikhailov, 2010). However, in-depth research on tautomerism and isomerism is beyond the scope of this study, and we only report the phenomena observed according to the two solvents used: $CDCl_3$ and DMSO-d_6. As for compound **10**, the ratio between the two isomers is approximately 1.5:1 in $CDCl_3$. The ^1H NMR spectrum of compound **9** in $CDCl_3$ showed the presence of at least three isomers, and we assume that the most abundant isomer is the one observed in DMSO-d_6. In COSY, cross-peaks between H-1" and 3-NH were clearly observed. It can be noticed that the chemical shift of

3-NH varies for the different isomers from 8.85 to 11.05 ppm, of which the most downfield 3-NH belongs to the most abundant isomer. Table 2 shows the NMR assignments for the major isomer of compound **9** and two isomers of compound **10** in $CDCl_3$.

β-Carboline alkaloids, isolated from several plant genera and animals, were reported to exhibit multiple pharmacological effects, including neuro-pharmacological and antitumoural activities. However, some *β*-carbolines showed mutagenic and carcinogenic properties (Cao et al., 2007). Therefore, it is crucial to check their presence in a common food supplement like *L. meyenii*. Until now, only (1*R*,3*S*)-1-methyl-1,2,3,4-tetrahydro-*β*-carboline-3-carboxylic acid (MTCA), which is a well-known *β*-carboline alkaloid (Gutsche and Herderich, 1997) was isolated from a butanolic extract of *L. meyenii*. Compound **11** was identified as 1,2,3,4-tetrahydro-*β*-carboline-3-carboxylic acid (Fig. 1), showing resemblance to MTCA. Briefly, the ^{1}H NMR spectrum of compound **11** showed NMR signals typical for an indole moiety: (1) four adjacent aromatic hydrogens (H-5, H-6, H-7 and H-8) appeared as a doublet (*J*=7.8 Hz), a doublet of triplets (*J*=7.8, 1.5 Hz), a doublet of triplets (*J*=7.8, 1.5 Hz) and a doublet (*J*=7.8 Hz), respectively; (2) in DMSO-d_6, the NH signal of the indole nucleus was highly deshielded and resonated at 10.87 ppm (9-NH); (3) HMBC correlations of the 9-NH with aromatic carbons (C-4a, C-4b, C-8a and C-9a) were observed. Next, two AB systems with a large coupling constant of 15-16 Hz indicated two methylene functionalities attached to the indole nucleus (1-CH_2- and 4-CH_2-), constituting a 4a, 9a-disubstituted pyrrole moiety. The 4-CH_2-group correlated with H-3 (3.54 ppm) in COSY, and H-3 correlated with the carboxyl-group at 170.5 ppm in the HMBC spectrum. Notably, hydrogen signals of the carboxyl-group and 2-NH were not observed in DMSO-d_6. Although this compound is known, this was the first time it was isolated from *L. meyenii*.

7.2.2. Structure revision of “macaridine”

Compound **8** surprisingly showed ^{1}H and ^{13}C NMR chemical shift values similar to two compounds previously reported by Muhammad et al. (2002) and Zhou et al. (2018) for macaridine and macapyrrolin C (see Table 3 and Fig.3), respectively, which both have a molecular formula $C_{13}H_{13}NO_2$ (Muhammad et al., 2002; Zhou et al., 2018). The only small difference in reported spectroscopic data is the UV

absorption: 'macaridine' showed a λ_{max} in methanol at 208, 255 and 294 nm, while for macapyrrolin C only a λ_{max} in methanol at 292 nm was reported. In our study, we followed the isolation procedure described by Muhammad et al. (2002) with the aim of isolating macaridine. However, on the basis of extensive 2D NMR analysis, the structure of compound **8** was elucidated as macapyrrolin C. In the 'macaridine' structure, HMBC correlations are expected between H-8/C-8 (δ_H 5.73/ δ_C 48.9) and H-2/C-2 (δ_H 4.54/ δ_C 57.0). However, surprisingly, Muhammad et al. (2002) did not observe such HMBC correlations, in spite of the favorable three-bond distance between positions 2 and 8 in the proposed structure. Nevertheless, absence of this HMBC correlation would be more likely in case of a four-bond distance between these specific signals, as would be the case for H-1'/C-1' and -CH_2- of 5-CH_2OH (δ_H 5.76/ δ_C 49.0 and δ_H 4.57/ δ_C 57.1, respectively) in macapyrrolin C. Apart from that, Zhou et al. (2018) reported three pyrrole alkaloids from *L. meyenii*, i.e. macapyrrolins A-C, while no other macaridine-like analogues were reported so far.

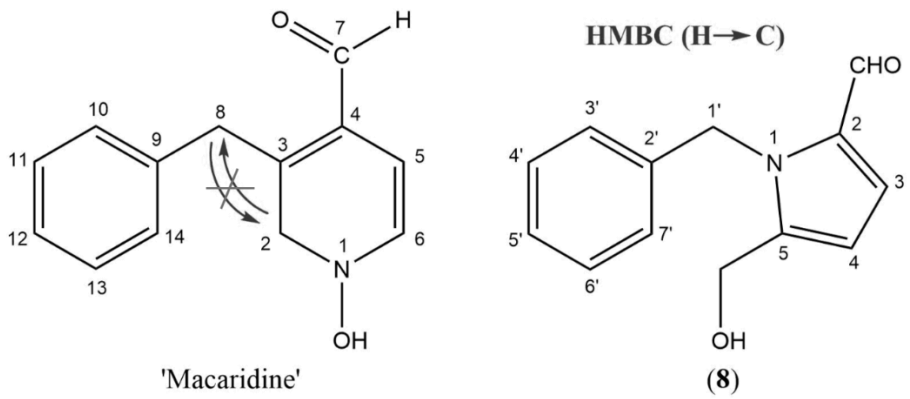


Fig. 3. Structures of the proposed 'macaridine' (left) and macapyrrolin C (right)

Table 3. Comparison between experimental (exp) and calculated (calc) ^{1}H and ^{13}C NMR chemical shifts of 'macaridine' and macapyrrolin C

Position	Macaridine in $CDCl_3$ (Muhammad et al. 2001)				Position	Macapyrrolin C in $CDCl_3$ (current study)			
	δ_H (exp)	δ_H (calc)	δ_C (exp)	δ_C (calc)		δ_H (exp)	δ_H (calc)	δ_C (exp)	δ_C (calc)

1	-	-	-		1	-	-	-	-
8	5.73	3.82	48.9	35.9	1’	5.76	5.89	49.0	49.0
2	4.54	3.58	57.0	55.1	5-CH_2OH	4.57	4.12	57.1	57.3
6	6.29	6.19	111.2	94.5	4	6.32	6.08	111.2	110.7
5	6.94	5.33	124.8	126.1	3	6.96	6.69	124.7	124.2
10	6.98	7.15	126.5	127.6	3’	6.99	6.96	126.7	125.6
14	6.98	7.15	126.5	127.6	7’	6.99	6.96	126.7	125.6
12	7.22	7.18	128.0	128.1	5’	7.25	7.12	127.9	125.9
11	7.26	7.18	129.1	128.1	4’	7.29	7.17	129.3	127.5
13	7.26	7.18	129.1	130.8	6’	7.29	7.17	129.3	127.5
4	-	-	133.2	137.4	2	-	-	133.3	132.6
9	-	-	138.2	140.5	2’	-	-	138.3	137.9
3	-	-	142.5	151.6	5	-	-	142.3	143.0
7	9.52	9.69	180.2	184.6	2-CHO	9.57	9.26	180.2	177.9
CMAE		0.529		4.477	CMAE		0.183		1.008

Over the last decades, the calculation of nuclear magnetic shielding constants by the Density Functional Theory (DFT) has been developed and widely applied to support structure elucidation and to predict relative configurations (Lodewyk et al., 2012). In the current study, DFT calculations were performed for the two structures, ‘macaridine’ and macapyrrolin C, in order to obtain additional proof for the structural revision of ‘macaridine’. Since the two structures are highly different, the DFT calculation can be expected to match with only one of the two. The mPW1PW91 functional, which has been reported as one of the best overall functionals for computing ^{13}C NMR chemical shifts, was used with a triple-ζ Pople’s basis set 6-311+G(2d,p) to calculate quantitative computed data (Jensen, 2008; Ermanis et al., 2017). Consequently, as can be seen in Table 3, calculated chemical shifts of macapyrrolin C are in complete agreement with the experimental data with a corrected-mean absolute error (CMAE) of 0.183 ppm for hydrogen shifts and 1.008 ppm for carbon shifts. The deviations for ‘macaridine” are much more pronounced. The calculated data also explain the deshielding effect observed for H-1’ (δ_H 5.76) and C-1’ (δ_C 49.0) of macapyrrolin C, which is caused by attachment to the nitrogen of the pyrrole ring, but which is not the case for ‘macaridine’. Therefore, the structure of ‘macaridine’ is to be revised as macapyrrolin C. Detailed

information about all contributing conformers subjected to DFT calculations can be found as Supporting Information.

7.2.3. Molecular networking and metabolomics analysis

In nature, the quaternary imidazole alkaloids represent a unique class of compounds reported exclusively in *L. meyenii*. In order to gain more insight in the compositional profile of this class of compounds, a molecular network-guided approach was followed. Fig. 4 shows the molecular network containing the imidazole alkaloids obtained in the present work, i.e. compounds **1** – **7** (nodes (6), (16), (3), (24), (29), (36) and (40), respectively).

Fig. 4. Cluster containing identified imidazole and amidine alkaloids. Nodes were numbered from lowest to highest *m/z* values. Detailed information on unknown nodes can be found as Supporting Information.

Known compounds are shown as orange nodes, unknown compounds as grey nodes and tentatively identified compounds as grey nodes with black borders. The node size correlates with the relative abundance of each feature. Since lepidilines A – F are quaternary alkaloids that can be detected in positive ion mode without the need of ionisation in the ion source, and since their structures are rather similar, their peak intensity shows a more or less linear correlation to their concentration. Therefore, the peak intensities of the lepidiline signals were used to estimate the ratios of lepidilines in our samples. Hence, the node sizes of the lepidilines A-F were adjusted according to their relative intensities determined in the CH_2Cl_2 (I) fraction of procedure 1 (see Fig. S62), which were 7.14+E5, 8.88+E5, 3.83+E5, 3.06+E5, 1.44+E5, 2.29+E4 and 2.60+E4, respectively. It is clear that lepidilines A and B are the major imidazole alkaloids, followed by lepidilines C, D and E. Lepidilines F and G are found in levels approximately ten-fold less than lepidilines C, D and E, and twenty-fold less than lepidilines A and B.

Interestingly, two amidine alkaloids (nodes (5) and (10)) are also present in the cluster, but are found further away from the region of the known imidazole alkaloids. Most probably, this cluster contains compounds with a -N=C-N- moiety, including imidazole, amidine and possibly other unknown scaffolds.

Another noteworthy point is the fact that there are three *m/z* values (227.15, 215.15 and 239.15) which were reported before in several HRMS-based metabolomics studies of *L. meyenii*. However, the structures of our purified compounds with corresponding *m/z* values (compounds **1**, **3** and **9**, respectively) are not in agreement with the previously proposed structures. Zhou et al. (2017) suggested that the product with *m/z* 227.15 corresponded to 1-dibenzyl-2-propane-4,5-dimethylimidalium, but this chemical name is incorrect, and instead should be 1-benzyl-2-propane-4,5-dimethylimidazolium, if the compound would exist. Furthermore, this chemical name was assigned to two different *m/z* values: 227.15 and 229.16, which were different peaks eluting at different retention times (Zhou et al., 2017). Unfortunately, this misinterpretation was restated by Tafuri et al. (2019). Concerning the product with *m/z* 215.15, Zhou et al. (2017) and Geng et al. (2020) proposed the structure of 1-benzyl-2-ethyl-4,5-dimethylimidazolium, which was also erroneously named as 1-dibenzyl-2-ethyl-4,5-dimethylimidazolium.

Apart from these mistakes in chemical names, it can be deduced that these studies did not consider the possibility of substitution on both N-1 and N-3, and therefore, a substituent was erroneously located in position C-2 instead of N-3. Since the present study describes the presence of amidine alkaloids in *L. meyenii* for the first time, misinterpretations in previously published metabolomics research for this class of compounds were inevitable. For instance, the ion with *m/z* 239.15 was proposed to correspond to an imidazole alkaloid by Zhou et al. (2017).

As stated by the GNPS platform, molecular networks are built based on similarities between MS fragmentation data of all compounds. Therefore, compounds within close proximity are expected to show the highest structural similarities. Three grey nodes with black borders (13), (46) and (49) correspond to *m/z* values of 245.1654, 337.1913 and 351.2079, respectively. Node (13) is one of the first neighbors of lepidiline F, while nodes (46) and (49) share lepidilines C and D as first neighbors. Based on the mass difference of 30 Da observed between three couples (245.1654 and 215.1550 (lepidiline F), 337.1913 and 307.1806 (lepidiline C), 351.2079 and 321.1968 (lepidiline D)), the signal at *m/z* 245.1654 likely represents a methoxylated derivative of lepidiline F, the signal at *m/z* 337.1913 a methoxylated derivative of lepidiline C, and the signal at *m/z* 351.2079 a methoxylated derivative of lepidiline D. The position of the additional -OMe groups can be deduced based on knowledge about the -OMe position of the imidazole alkaloids identified in this study: lepidilines C, D and F possess only one -OMe group, which is meta-oriented. Furthermore, the effect of –OMe substitution on the retention properties on a RP-C18 column was assessed. It was observed that compounds with a -OMe group attached to a benzyl moiety are more strongly retained on the RP-C18 stationary phase (Fig. S62-S65): lepidiline C elutes slower than lepidiline A; lepidiline D is more retained than lepidiline B; and lepidiline F elutes slower than lepidiline E; a similar behavior was observed for products with *m/z* values 245.1654, 337.1913 and 351.2079. Thus, we tentatively identified the structures corresponding to these three *m/z* values as 3-ethyl-1-(3'-methoxybenzyl)-4,5-dimethyl-imidazolium for *m/z* 245.1654; 1,3-bis(3'-methoxybenzyl)-4,5-dimethyl-imidazolium for *m/z* 337.1913; and 1,3-bis(3'-methoxybenzyl)-2,4,5-trimethyl-imidazolium for *m/z* 351.2079 (Fig. 5). Finally, due to co-elution, fragmentation patterns of these compounds were mixed with those of other compounds in the MSe spectra.

Therefore, we performed MRM scans to confirm specific fragments in positive ion mode, and three fragmentation reactions were set for each parent ion (Fig. 5).

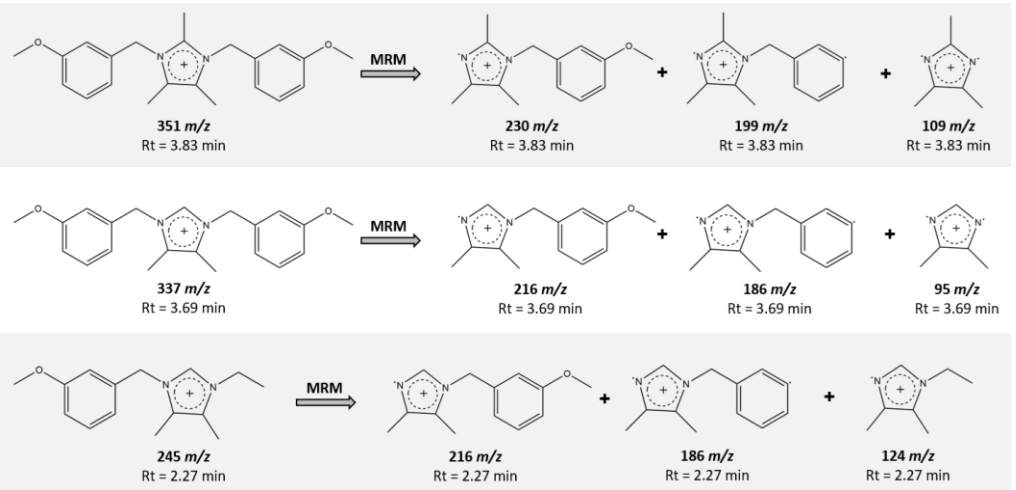


Fig. 5. Results of three channel MRM scans using UPLC conditions for the TQD system.

7.3. Conclusion

In this study, the alkaloid fraction of *L. meyenii* was explored. Three undescribed imidazole alkaloids were identified, for which the names lepidiline E, F and G were adopted, along with three tentatively identified imidazole derivatives. Amidine alkaloids were also described for the first time in this species. UPLC-HRMS-based molecular networking indicated the presence of imidazole, amidine and *β*-carboline alkaloids. Thorough comparison of our findings and previous profiling studies pointed out several misinterpretations with regard to the imidazole alkaloids. From our point of view, it is crucial to restate that isolation and structure elucidation by NMR is the key to undoubtedly identify phytochemical constituents. To avoid misinterpretations, care should be taken in metabolomics research based on HRMS data, and tentative identification should only be carried out in combination with thorough proof.

Furthermore, macapyrrolin C was purified and was used as a reference to revise the structure of previously reported ‘macaridine’. Based on the comprehensive

comparison of NMR data and DFT calculations, it was confirmed that the structure reported for 'macaridine' should be revised as macapyrrolin C.

7.4. Materials and Methods

7.4.1. Isolation and Purification

Two batches of plant material were purchased from two suppliers: (1) 2 kg hydroglycerin organic flour of the roots of *Lepidium meyenii* Walp. (Brassicaceae), drug/extract ratio 4:1, batch number 0001-29708/18 supplied by Soria Bel N.V. (Ichtegem, Belgium) and (2) 1 kg of "Bio-Maca Pulver" (root powder), batch number 87619, supplied by Herbis Natura GmbH (Berlin, Germany). Two different isolation procedures were used for the two batches of plant material: the first one followed a general fractionation scheme, the second one was more dedicated.

7.4.1.1. Procedure 1 (applied to the first batch of plant material):

Root flour (2 kg) was macerated and percolated with approximately 25 L of 80% ethanol, filtered and dried under reduced pressure. Next, liquid-liquid partition was applied in order to obtain an alkaloid-rich fraction. Briefly, crude extract was suspended in water and was acidified with 10% HCl to a pH < 3 before performing liquid-liquid partition with CH_2Cl_2 (I). Next, the pH of the acidified phase was increased to a value ≥ 9 by adding NH_4OH (25%), followed by a second liquid-liquid partition with CH_2Cl_2 (II). In this way, three fractions (CH_2Cl_2 (I), CH_2Cl_2 (II) and a residual aqueous fraction) were obtained and were examined by TLC with a mobile phase of EtOAc/MeOH/NH_4OH/H_2O (80:20:1:1). Observation of the TLC plates after spraying with the Dragendorff reagent indicated that the majority of alkaloids was present in the CH_2Cl_2 (I) fraction (16.2 g).

To confirm the presence of alkaloids in the CH_2Cl_2 (I) fraction, UPLC-TQD-MS analysis in positive ionization mode was performed using either a general MS scan detection method or a Selected Ion Recording (SIR). The latter aimed at the specific identification of the four reported major alkaloids from *L. meyenii*: lepidilines A-D and *m/z* values used in our method were derived from earlier studies performed by Cui et al. (2003) and Jin et al. (2016): *m/z* values of 277, 291, 307 and 321 (all $[M^+]$) were set for the identification of 1,3-dibenzyl-4,5-dimethylimidazolium

chloride (lepidiline A), 1,3-dibenzyl-2,4,5-trimethylimidazolium chloride (lepidiline B), 3-benzyl-1-(3-methoxybenzyl)-4,5-dimethylimidazolium chloride (lepidiline C) and 3-benzyl-1-(3-methoxybenzyl)-2,4,5-trimethylimidazolium chloride (lepidiline D), respectively. As a result, the presence of lepidilines A-D in the CH_2Cl_2 (I) fraction was confirmed by SIR, while the MS scan analysis revealed the presence of several other unknown constituents. Next, the CH_2Cl_2 (I) fraction was further fractionated using a Diaion® HP-20 column. The column was sequentially eluted with MeOH/H_2O (10:90), MeOH/H_2O/AA (40:60:0.5) and MeOH/H_2O/AA (50:50:0.5). Fractions of 250 mL each were collected and concentrated under reduced pressure followed by freeze-drying. Fractions showing identical spots on TLC were combined. This resulted in four subfractions, namely OC1, OC2, OC3 and OC4. UPLC-TQD-MS profiling was carried out on the four subfractions using the aforementioned method and lepidilines A-D were all found to be present in fractions OC3 and OC4, while fractions OC1 and OC2 mainly contained unknown *m/z* values.

Based on these results, further fractionation by flash chromatography was performed. All four fractions OC1 (41 mg), OC2 (45 mg), OC3 (659 mg) and OC4 (342 mg) were fractionated using Claricep™ C-series cartridges containing 12 g deactivated silica gel (40-60 µm). Solvents were CH_2Cl_2 (A) and MeOH + 0.1% AA (B) and the flow rate was 27 mL/min. The gradient used was: 0-5 min (0% B), 45-47 min (40% B), 50-60 min (100% B), while the gradient was halted during elution of detected compounds. The UV detector was set at 230 and 274 nm. After flash chromatography, fractions were combined based on the resulting chromatograms and on TLC analysis and were profiled by UPLC-TQD-MS.

From fraction OC1, a mixture of two compounds was obtained with *m/z* values of 227 (major compound) and 215 (minor compound) in positive ionization mode. However, the amount of the mixture was only 2 mg, and therefore, further analysis by NMR was carried out on the mixture as such with the aim of identifying both compounds. From fraction OC2, subfraction H3 was derived, containing two main constituents with *m/z* values of 227 and 257 in positive ionization mode. From subfractions OC3 and OC4, five new subfractions (FC1-FC5) were obtained, of which subfractions FC1 and FC2 were selected for further purification.

Semi-preparative HPLC-DAD with manual collection was performed for isolation of pure compounds from the fractions FC1, FC2 and H3. More specifically, the preparative method for fractions FC1 and FC2 was: (1) sample concentration: 1 mg/mL in water; (2) injection volume: 600 µL; (3) solvent A: $H_2O/NH_4Ac/FA$ (100:0.5:0.3, v/v) and solvent B: $MeOH/NH_4Ac/FA$ (100:0.5:0.3, v/v) (NH_4Ac was prepared as a 20 µM solution); (4) gradient: 0-5 min (50% B), 14 min (56% B), 16 min (90% B), 18-23 min (50% B); (5) UV detection: 230 and 274 nm; (6) flow rate: 3 mL/min. To be noticed, lepidilines B and C co-eluted during earlier analysis, but the addition of NH_4Ac resulted in baseline separation of these two alkaloids. The HPLC conditions for fraction H3 were: (1) sample concentration: 5 mg/mL in MeOH; (2) injection volume: 100 µL; (3) solvent A: $H_2O/FA/DEA$ (100:0.5:0.3, v/v) and solvent B: ACN/MeOH/FA/DEA (80:20:0.5:0.3, v/v); (4) gradient: 0-10 min (25% B), 15 min (38% B), 20-23 min (90% B), 25-30 min (25% B); (5) UV detection: 230 and 274 nm; (6) flow rate: 3 mL/min. Finally, to remove the diethylamine formate, solid-phase extraction with a Chromabond® C18 ec (500 mg, 6 mL) column (Macherey-Nagel, Düren, Germany) was carried out. Each sample was dissolved in water, loaded on an activated column and subsequently, the column was rinsed with 4 column volumes of water to remove the salt. In a next step, the column was rinsed with 3 column volumes of MeOH to elute the purified alkaloids.

Lepidilines E and F (compounds **1** and **2**) were purified by semi-preparative HPLC-DAD of fraction H3. Lepidilines E and G (compounds **1** and **3**) were isolated from fraction OC1 by flash chromatography as a mixture, as previously mentioned. Lepidilines A-D (compounds **4**-**7**) were isolated by semi-preparative HPLC-DAD from fractions FC1 and FC2.

7.4.1.2. Procedure 2 (applied to the second batch of plant material):

Similar to the first procedure, root powder (1 kg) was macerated and percolated with approximately 15 L of 80% ethanol, filtered and dried under reduced pressure to obtain a crude extract. However, taking into consideration the publication of Muhammad et al. (2002) and the chemical nature of the targeted compounds, the partitioning procedure was modified in order to optimize the isolation of the lepidilines, macaridine and (1R,3S)-1-methyl-1,2,3,4-tetrahydro-*β*-carboline-3-carboxylic acid (MTCA) as follows: (1) First, the crude extract was suspended in

water and partitioned with CH_2Cl_2 (I), (2) Second, the water phase was acidified to pH < 3 and was partitioned with CH_2Cl_2 (II), (3) Finally, the acidified water phase was partitioned with *n*-butanol (Muhammad et al., 2002).

The CH_2Cl_2 (I) extract (12 g) was fractionated by flash chromatography using a Claricep™ C-series cartridge containing 40 g deactivated silica gel (40-60 µm). Briefly, settings of the flash system were as follows: (1) solvents: CH_2Cl_2 (A), MeOH + 0.5% AA (B); (2) flow rate: 30 mL/min; (3) gradient: 0-5 min (0% B), 50-55 min (60% B), 60-70 min (100% B) and the gradient was halted during elution of detected compounds; (4) UV detection: 230 and 274 nm. The *n*-butanol extract (15 g) was fractionated using an MCI column (200 g) and a gradient of increasing MeOH:H_2O ratios. In both cases, collected fractions were combined based on TLC analysis, resulting in 15 subfractions originating from the CH_2Cl_2 (I) extract and 40 subfractions derived from the n-BuOH extract. All subfractions were analyzed by UPLC-TQD-MS for the identification of lepidilines, macaridine and MTCA. The UPLC-TQD-MS conditions were identical to those described in procedure 1, but for SIR, *m/z* values of 216.1 $[M+H]^+$ and 231.2 $[M+H]^+$, were added for the identification of macaridine and MTCA in positive ionization mode, respectively (Muhammad et al., 2002). This analysis revealed that the lepidilines, 'macaridine' and MTCA were all present in different fractions.

The isolation of lepidilines A, B, C, D and E was repeated according to the methods described in procedure 1, in order to obtain a higher amount of each.

Isolation of 'macaridine' was carried out by preparative HPLC-DAD of fraction CH_2Cl_2 (I), using the following conditions: (1) sample concentration: 10 mg/mL in MeOH; (2) injection volume: 60 µL; (3) solvent A: H_2O + 0.1% FA and solvent B: ACN + 0.1% FA (v/v); (4) gradient: 0-5 min (25% B), 20 min (40% B), 28 min (72% B), 30-32 min (90% B) and 34-38 min (25% B); (5) UV detection: 210, 274 and 294 nm; (6) flow rate: 3 mL/min.

Semi-preparative HPLC-DAD was performed on fraction *n*-BuOH-11, the fraction presumably containing MTCA. The conditions were as follows: (1) sample concentration: 5 mg/mL in MeOH; (2) injection volume: 120 µL; (3) solvent A: H_2O and solvent B: ACN/MeOH – 4/6 (v/v); (4) gradient: 0-5 min (10% B), 20 min (25%

B), 23 min (100% B) and 25-30 min (10% B); (5) UV detection: 210 and 254 nm; (6) flow rate: 3 mL/min.

Accurate mass measurements were performed for all isolated compounds **1-11**. 1D and 2D NMR spectra were recorded for structure elucidation of all compounds.

7.4.1.3. Spectroscopic data

Lepidiline E (**1**)

Yellow amorphous powder (3 mg); UV λ_{max} 230 nm; ^{1}H and ^{13}C NMR (CD_3OD, 400 and 100 MHz): see Table 1; Positive HRESIMS *m/z* 227.1552 $[M]^+$ (calcd for $C_{15}H_{19}N_2$, 227.1548)

Lepidiline F (**2**)

Yellow amorphous powder (< 1 mg); UV λ_{max} 230, 274 nm; ^{1}H and ^{13}C NMR (CD_3OD, 400 and 100 MHz): see Table 1; Positive HRESIMS *m/z* 257.1656 $[M]^+$ (calcd for $C_{16}H_{21}N_2O$, 257.1654)

Lepidiline G (**3**)

Yellow amorphous powder (< 1 mg); UV λ_{max} 230 nm; ^{1}H and ^{13}C NMR (CD_3OD, 400 and 100 MHz): see Table 1; ^{1}H NMR (DMSO-d_6, 400 and 100 MHz): δ_H 9.71 (s, H-2), 7.29-7.41 (m, overlapped, H-3', H-4', H-5', H-6', H-7', H-3", H-4", H-5", H-6", H-7"), 5.44 (s, H-1'), 4.16 (H-1"), 2.23 (s, 4-Me), 2.11 (overlapped, 5-Me), 1.41 (overlapped, H-2"); Positive HRESIMS *m/z* 215.1551 $[M]^+$ (calcd for $C_{14}H_{19}N_2$, 215.1548)

N,N'-dibenzylacetamidine (**9**)

Yellow amorphous powder (1.5 mg); UV λ_{max} 230, 256 nm; ^{1}H and ^{13}C NMR ($CDCl_3$, 400 and 100 MHz): see Table 2; ^{1}H and ^{13}C NMR (DMSO-*d6*, 400 and 100 MHz): δ_H 7.35-7.40 (overlapped, H-4', H-5', H-6', H-4", H-5", H-6"), 7.28-7.30 (overlapped, H-3', H-7', H-3", H-7"); 4.56 (s, H-1'), 4.53 (s, H-1"), 2.25 (s, 2-Me); δ_C 164.3 (C-2), 128.9 (C-4', C-5', C-6', C-4", C-5", C-6"), 127.4 (C-3', C-7', C-3", C-7"), 47.0 (C-1'),

45.4 (C-1”), 17.3 (2-Me); Positive HRESIMS *m/z* 239.1549 $[M+H]^+$ (calcd for $C_{16}H_{19}N_2$, 239.1548)

N,N’-dibenzylformamidine (**10**)

Yellow amorphous powder (1 mg); UV λ_{max} 230, 256 nm; ^1H and ^{13}C NMR ($CDCl_3$, 400 and 100 MHz): see Table 2; ^1H and ^{13}C NMR (DMSO-d_6, 400 and 100 MHz): δ_H 7.37-7.41 (overlapped, H-4’, H-5’, H-6’), 7.32-7.39 (overlapped, H-2, H-3”, H-4”, H-5”, H-6”, H-7”); 4.46 (s, H-1’), 3.77 (s, H-1”); δ_C 165.9 (C-2), 136.0 (C-2’), 134.9 (C-2”), 129.2 (C-3”, C-7”), 129.0 (C-4’, C-5’, C-6’, C-4”, C-5”, C-6”), 128.2 (C-3’, C-7’), 45.6 (C-1’), 38.8 (C-1”); Positive HRESIMS *m/z* 225.1402 $[M+H]^+$ (calcd for $C_{15}H_{17}N_2$, 225.1392)

Lepidilines A-D (**4**, **5**, **6**, **7**), macapyrrolin C (**8**), 1,2,3,4-tetrahydro-*β*-carboline-3-carboxylic acid (**11**): see as Supporting information

7.4.2. Computational Details

Monte Carlo conformational search was performed by PCMODEL (version 10.0) using Merck Molecular Force Field (MMFF94) applying 8 and 7 kcal.mol^{-1} energy windows for two consecutive conformational search cycles. Afterwards, for chemical shift calculations, all resulting conformers were subjected to geometry optimization using B3LYP/6-311+G(2d,p) level of theory in gas-phase, and shielding tensors were then computed at mPW1PW91/6-311+G(2d,p) level of theory with polarizable continuum model (PCM). After dereplication, resulting isotropic shielding values were scaled based on slopes and intercepts provided by the CHESHIRE website (http://cheshirenmr.info) (for ^1H nuclei, slope = -1.0933 and intercept = 31.9088; for ^{13}C nuclei, slope = -1.0449 and intercept = 187.1018). Geometry optimization, frequencies and shielding tensor calculations were carried out with the Gaussian16 package (Frisch et al., 2016). Boltzmann distributions were estimated from the B3LYP/6-311+G(2d,p) level of theory using the sum of electronic and thermal free energies at 298.15 K. Only conformers with energies within 2.5 kcal mol^{-1} from the global minimum were submitted to the GIAO (Gauge-Independent Atomic Orbital) calculation. Avogadro 1.2.0 software was used for visualization of computed outputs.

7.4.3. Molecular networking and metabolomics analysis

All crude extracts and fractions mentioned in section 4.3 were profiled by UPLC-QTOF-MS and -MS/MS in positive ionization mode (MS scan range: 50-1500 *m/z*). The UPLC conditions were as follows: sample concentrations: 10 µg/mL and 1 µg/mL in 80% MeOH, injection volume: 5 µL, mobile phase solvent A: H_2O/FA (100:0.1, v/v) and solvent B: ACN/FA (100:0.1, v/v), and gradient: 0-2 min (2% B), 22-24 min (100% B), 25-27 min (2% B). Raw data obtained from the UPLC-QTOF-MS/MS analysis were converted to *abf format using the Analysis Base File converter. Then, data processing was carried out with MS-Dial (Tsugawa et al., 2015), followed by the development of a feature-based molecular network, using the GNPS platform. The MS data processing workflow of MS-Dial (version 4.24) comprises *abf file import, data collection, peak detection, MS2 deconvolution, ion and adduct definition, peak alignment and isotope tracking. All MS data of the crude extracts and fractions were combined, in order to create a single molecular network. Setting parameters were as follows: positive ionization mode, centroid data, time analysis: 0.5-27.0 min, MS1 and MS2 mass range: 50-1500 Da, MS1 tolerance: 0.01 Da, MS2 tolerance: 0.025 Da, peak detection limit: 30,000, mass slice width: 0.1, MS/MS abundance cut off: 100 amplitude, ion and adducts: $[M]^+$, $[M+H]^+$, $[M+Na]^+$. Afterwards, processed files including an *mgf file and a text table were uploaded to the GNPS website and a molecular network was created with the feature-based molecular networking workflow (Nothias et al., 2019; Wang et al., 2016). To visualize the network, the output was imported into Cytoscape version 3.8.0 (Shannon et al., 2003). Peak annotation was performed manually for known compounds isolated by the authors and/or previously reported in literature. Due to co-elution, typical fragments of tentatively identified compounds were confirmed by MRM (multiple-reaction monitoring), conducted with the UPLC-TQD-MS system and using the LC conditions described in the section 4.3 and the MS setting parameters as demonstrated in Fig. 5.

7.5. Supporting information – Chapter 7

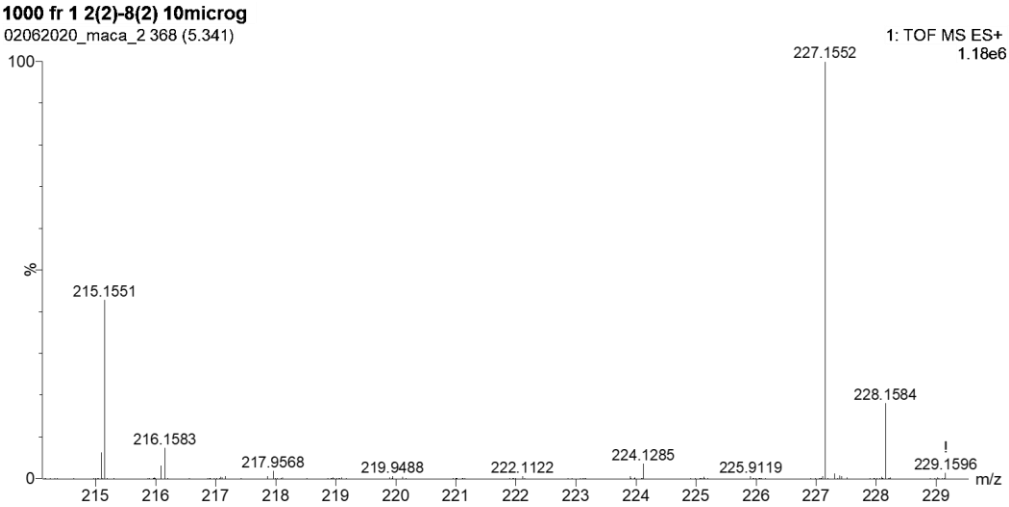


Fig. S1. HRMS data of the mixture of compounds **1** and **3**

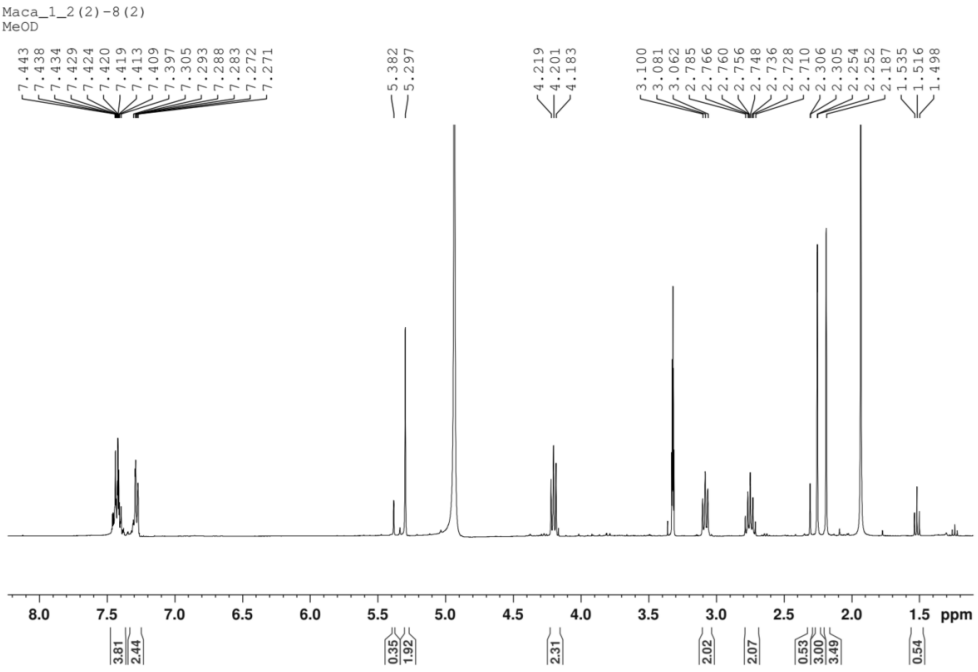


Fig. S2. ^{1}H spectrum of compounds **1** and **3** in MeOD

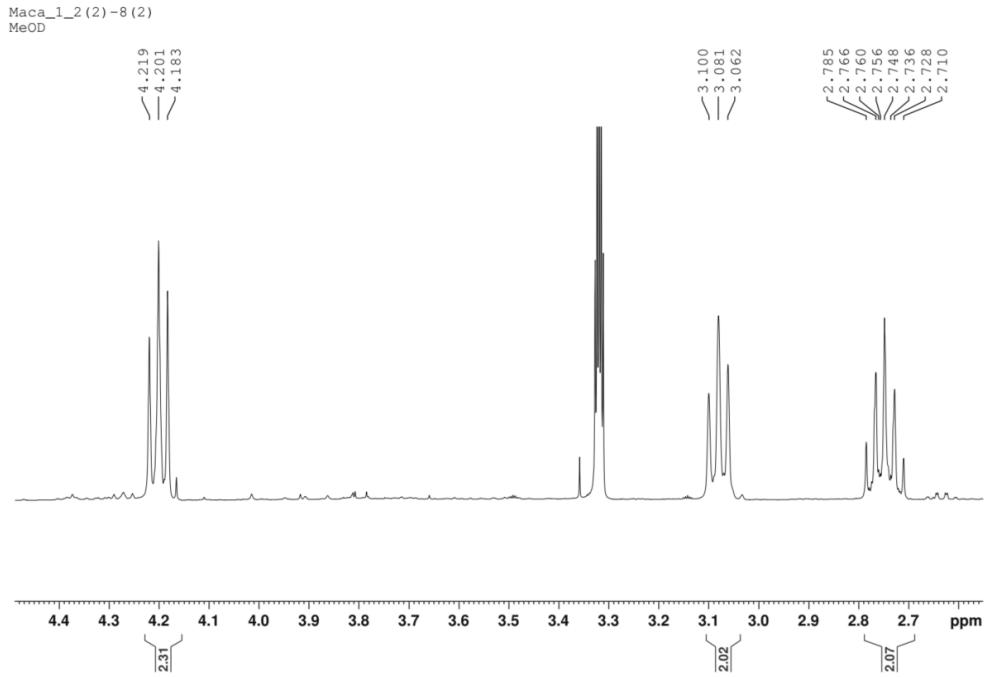


Fig. S3. Typical ^{1}H patterns of the additional fused ring of compound **1** in MeOD

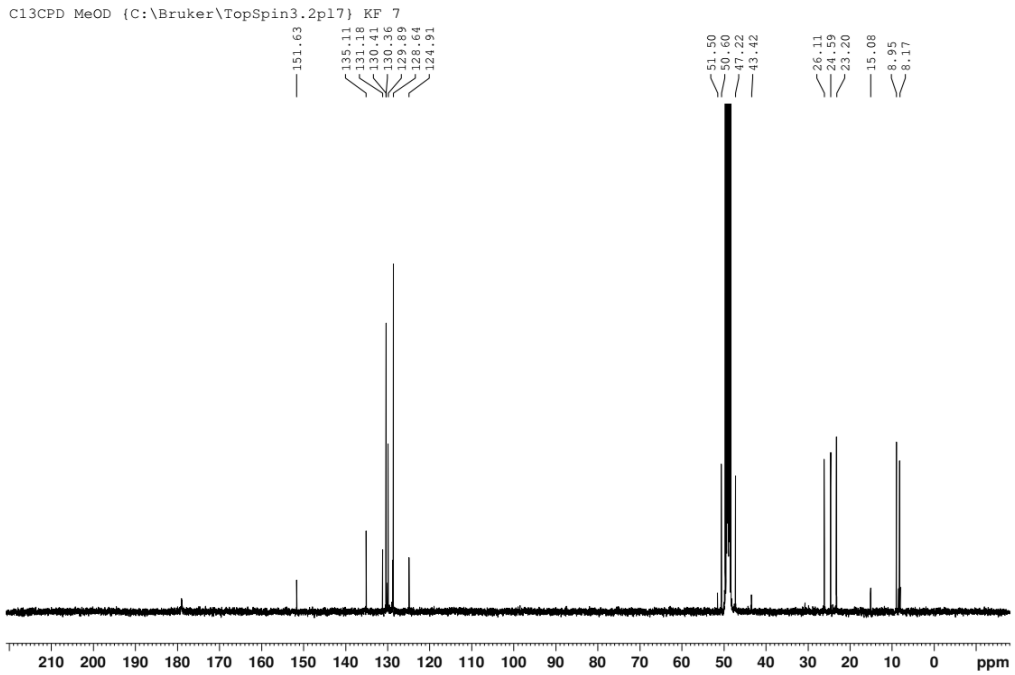


Fig. S4. ^{13}C spectrum of compounds **1** and **3** in MeOD

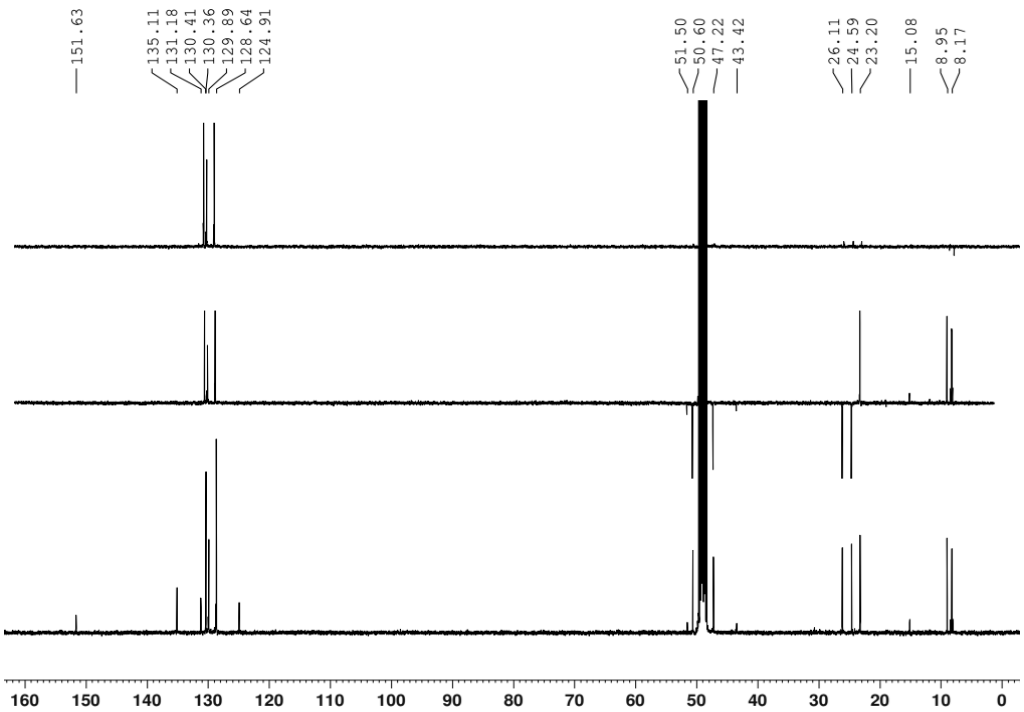


Fig. S5. ^{13}C and DEPT spectra of compounds **1** and **3** in MeOD

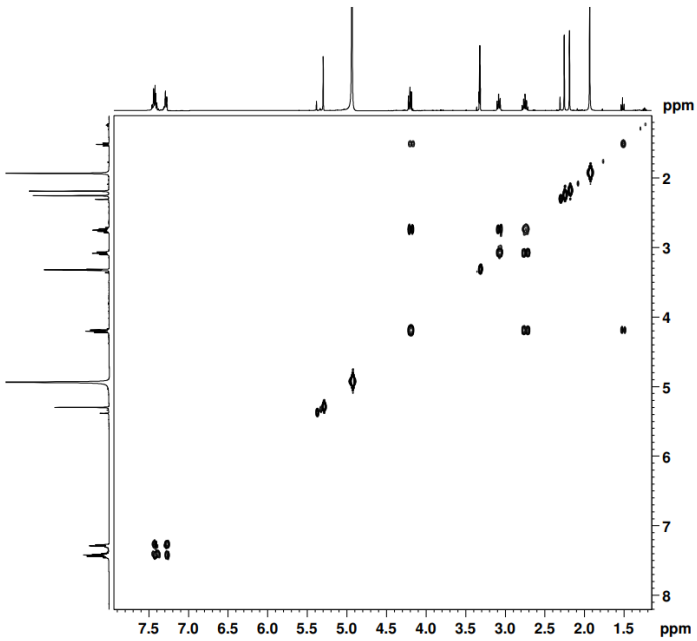


Fig. S6. COSY spectrum of compounds **1** and **3** in MeOD

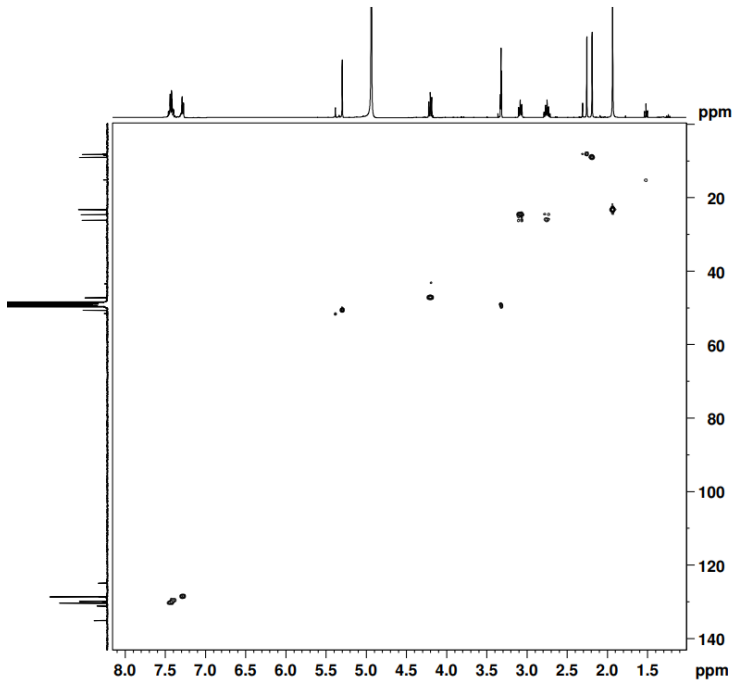


Fig. S7. HSQC spectrum of compounds **1** and **3** in MeOD

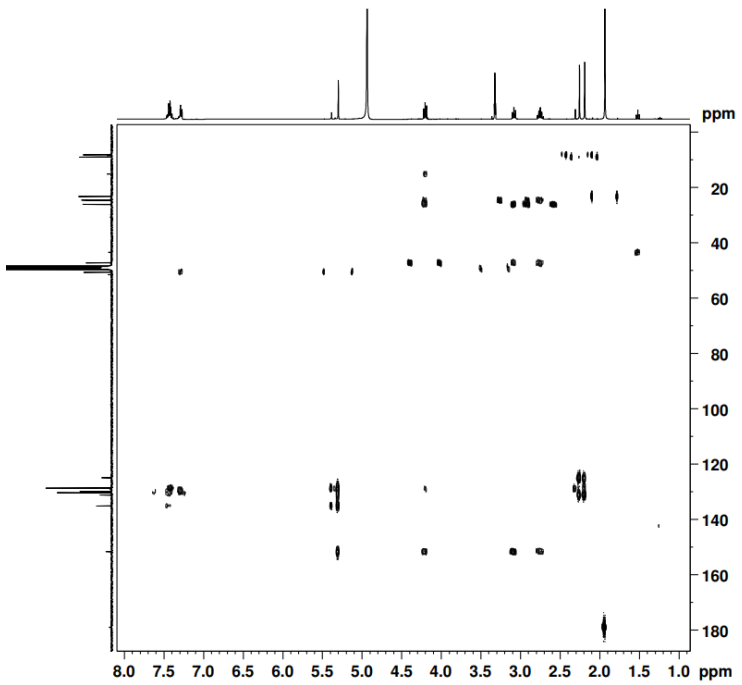


Fig. S8. HMBC spectrum of compounds **1** and **3** in MeOD

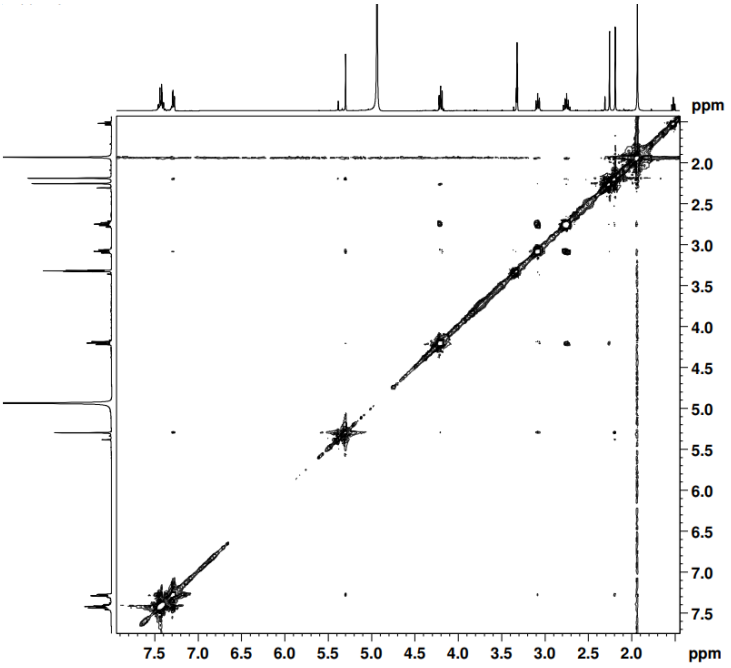


Fig. S9. NOESY spectrum of compounds **1** and **3** in MeOD

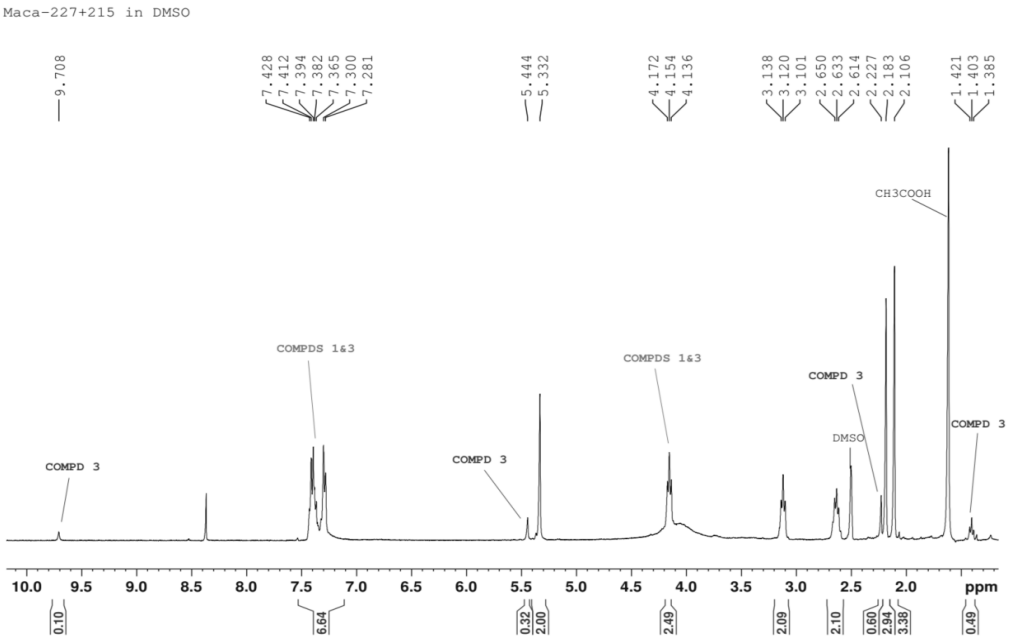


Fig. S10. ^{1}H spectrum of compounds **1** and **3** in DMSO-d_6 with H-2 of compound **3**.

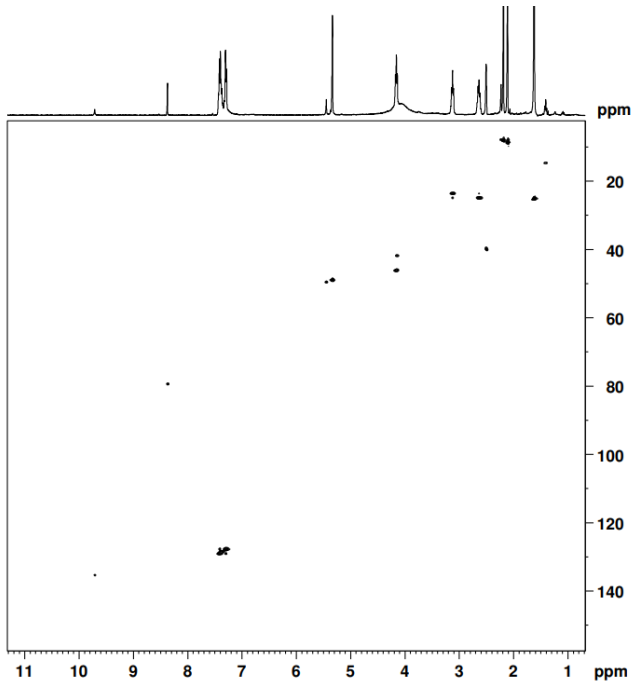


Fig. S11. HSQC spectrum of compounds **1** and **3** in DMSO-d_6 showing the correlation of H-2 with C-2 of compound **3**

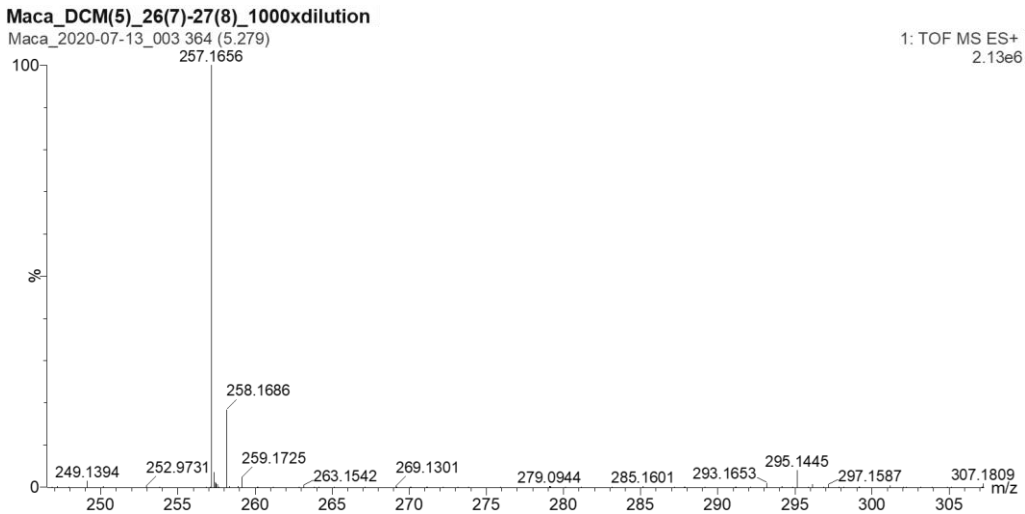


Fig. S11. HRMS data of the mixture of compound **2**

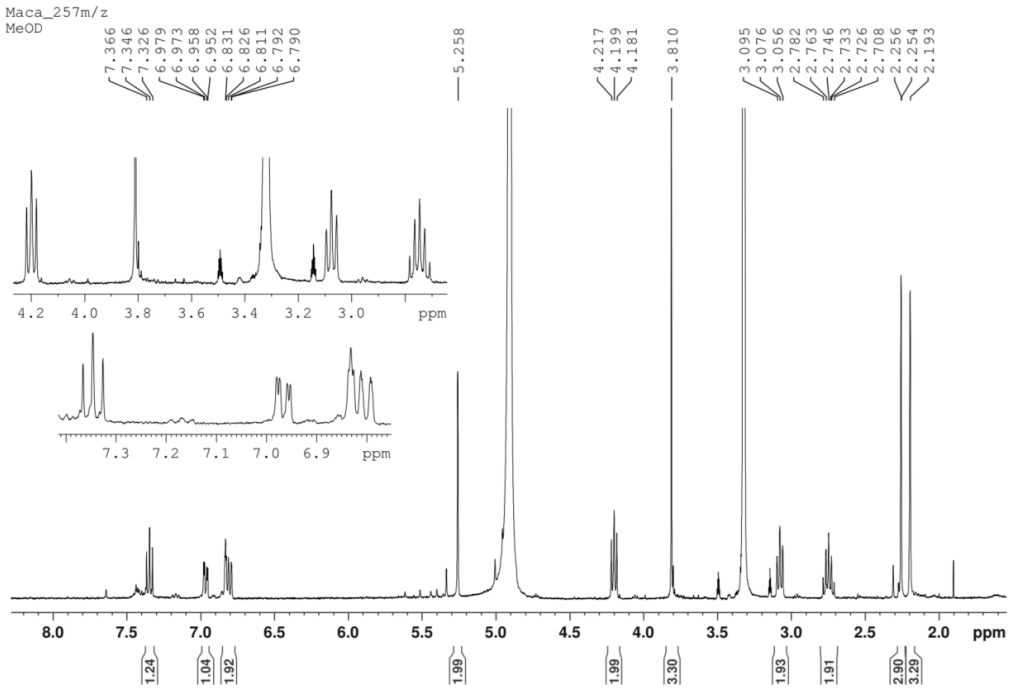


Fig. S12. ^{1}H spectrum of compound **2** in MeOD

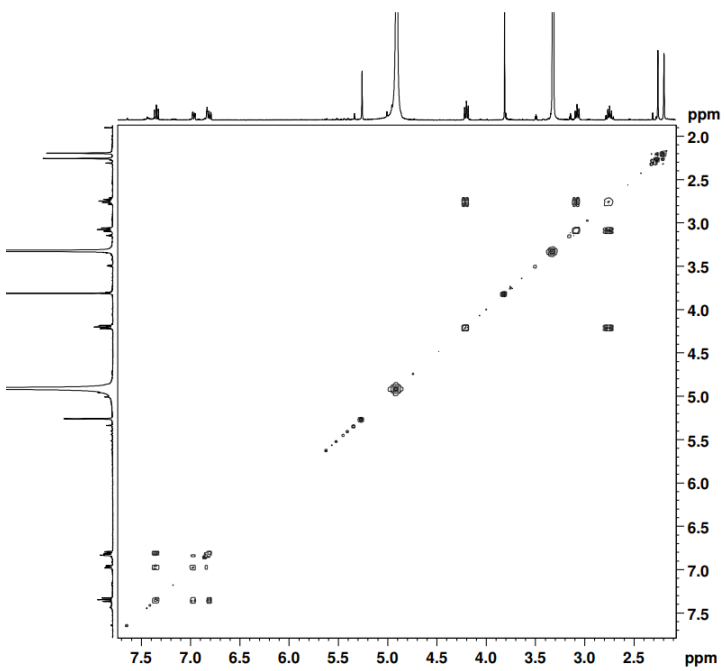


Fig. S13. COSY spectrum of compound **2** in MeOD

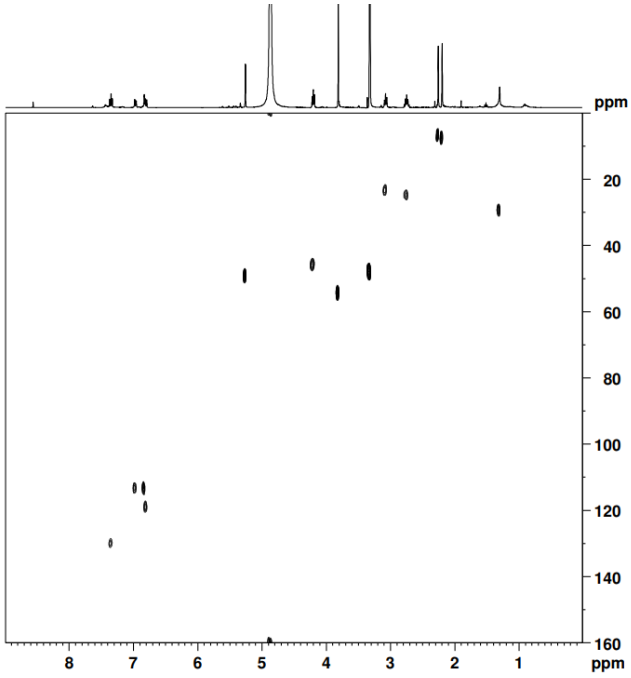


Fig. S14. HSQC spectrum of compound **2** in MeOD

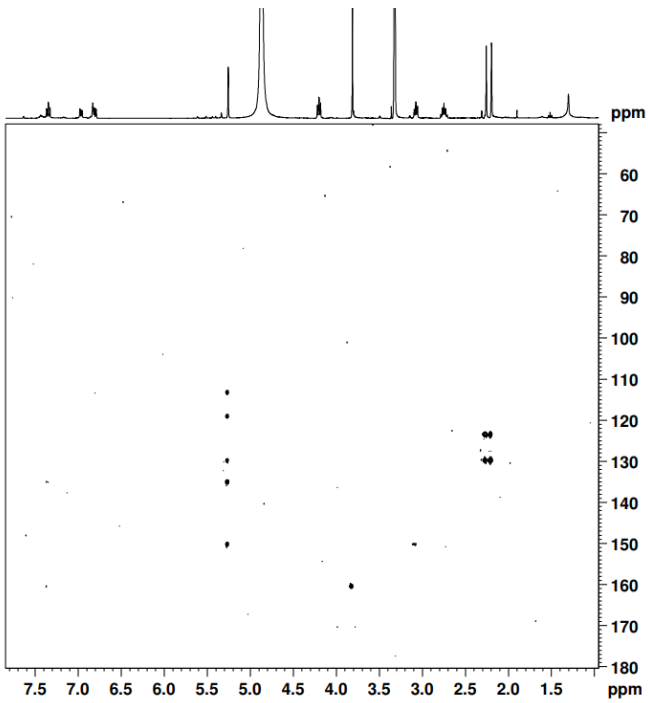


Fig. S15. HMBC spectrum of compound **2** in MeOD

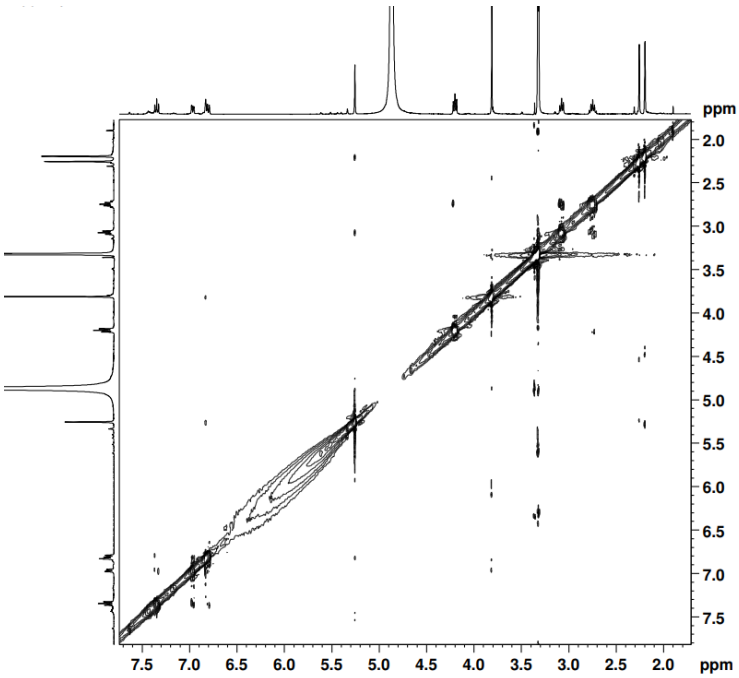


Fig. S16. NOESY spectrum of compound **2** in MeOD

7.6. References

Beharry, S., Heinrich, M., 2018. Is the hype around the reproductive health claims of maca (*Lepidium meyenii* Walp.) J. Ethnopharmacol. 211, 126–170.

Cao, R., Peng, W., Wang, Z., Xu, A., 2007. Carboline Alkaloids: Biochemical and Pharmacological Functions. Curr. Med. Chem. 14, 479–500.

Carvalho, F. V., Ribeiro, P.R., 2019. Structural diversity, biosynthetic aspects, and LC-HRMS data compilation for the identification of bioactive compounds of *Lepidium meyenii*. Food Res. Int. 125, 108615.

Cheng, C., Shen, F., Ding, G., Liu, A., Chu, S., Ma, Y., Hou, X., Hao, E., Wang, X., Yuanyuan Hou Y., Bai, G., 2020. Lepidiline A improves the balance of endogenous sex hormones and increases fecundity by Targeting HSD17B1. Mol. Nutr. Food Res. 64, 1900706.

Cui, B., Zheng, B., He, K.L., Zheng, Q., 2003. Imidazole alkaloids from *Lepidium meyenii*. J. Nat. Prod. 66, 1101–1103.

Ermanis, K., Parkes, K., Agback, T., Goodman, J., 2017. Doubling the power of DP4 for computational structure elucidation. Org. Biomol. Chem. 15.

Ganzera, M., Zhao, J., Muhammad, I., Khan, I.A., 2002. Chemical profiling and standardization of *Lepidium meyenii* (Maca) by reversed phase high performance liquid chromatography. Chem. Pharm. Bull. 50, 988–991.

Gaussian 16, Revision A.03, M. J. Frisch, G. W. Trucks, , Gaussian, Inc., Wallingford CT, 2016.

Geng, P., Sun, J., Chen, P., Brand, E., Frame, J., Meissner, H., Stewart, J., Gafner, S., Clark, S., Miller, J., Harnly, J., 2020. Characterization of Maca (*Lepidium meyenii/Lepidium peruvianum*) using a mass spectral fingerprinting, metabolomic analysis, and genetic sequencing approach. Planta Med. 86.

Gutsche, B., Herderich, M., 1997. HPLC-MS/MS profiling of tryptophan-derived alkaloids in food: identification of tetrahydro-β-carbolinedicarboxylic acids. J. Agric. Food Chem. 45, 2458–2462.

Geng, H.C., Yang, D.S., Chen, X.L., Wang, L.X., Zhou, M., Mei, W.Q., 2018. Meyeniihydantoins A–C, three novel hydantoin derivatives from the roots of Lepidium meyenii Walp. Phytochem. Lett. 26, 208–211.

Huang, Y.-J., Peng, X.-R., Qiu, M.-H., 2018. Progress on the Chemical Constituents Derived from Glucosinolates in Maca (Lepidium meyenii). Nat. Products Bioprospect. 8, 405–412.

Jensen, F., 2008. Basis set convergence of nuclear magnetic shielding constants calculated by density functional methods. J. Chem. Theory Comput. J. Chem. Theory Comput. 2008, 4, 5, 719–727

Jin, W., Chen, X., Dai, P., Yu, L., 2016. Lepidiline C and D: two new imidazole alkaloids from *Lepidium meyenii* Walpers (Brassicaceae) roots. Phytochem. Lett. 17, 158–161.

Kitamura, S., Morisseau, C., Harris, T.R., Inceoglu, B., Hammock, B.D., 2017. Occurrence of urea-based soluble epoxide hydrolase inhibitors from the plants in the order Brassicales. PLoS One 12, 4–9.

Lodewyk, M.W., Siebert, M.R., Tantillo, D.J., 2012. Computational prediction of ^{1}H and ^{13}C chemical shifts: a useful tool for natural product, mechanistic, and synthetic organic chemistry. Chem. Rev. 112, 1839–1862.

Maccallini, C., Montagnani, M., Paciotti, R., Ammazzalorso, A., De Filippis, B., Di Matteo, M., Di Silvestre, S., Fantacuzzi, M., Giampietro, L., Potenza, M.A., Re, N., Pandolfi, A., Amoroso, R., 2015. Selective acetamidine-based nitric oxide synthase inhibitors: synthesis, docking, and biological studies. ACS Med. Chem. Lett. 6, 635–640.

Minkin, V., Mikhailov, I., 2010. Amidines and imidates: Volume 2 (1991), in: Cheminform. pp. 527–621.

Muhammad, I., Zhao, J., Dunbar, D.C., Khan, I.A., 2002. Constituents of *Lepidium meyenii* 'maca.' Phytochemistry 59, 105–110.

Nothias, L.F., Petras, D., Schmid, R., 2019. Feature-based molecular networking in the GNPS analysis environment. bioRxiv 812404.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T., 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 13, 2498–2504.

Tian, X.X., Peng, X.R., Yu, M.Y., Huang, Y.J., Wang, X., Zhou, L., Qiu, M., 2018. Hydantoin and thioamide analogues from Lepidium meyenii. Phytochem. Lett. 25, 70–73.

Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., VanderGheynst, J., Fiehn, O., Arita, M., 2015. MS-DIAL: data-independent MS/MS deconvolution for comprehensive metabolome analysis. Nat. Methods 12, 523–526.

Wang, M., Carver, J.J., et al., 2016. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. Nat. Biotechnol. 34, 828–837.

Yu, M.Y., Qin, X.J., Peng, X.R., Wang, X., Tian, X.X., Li, Z.R., Qiu, M.H., 2017. Macathiohydantoins B–K, novel thiohydantoin derivatives from Lepidium meyenii. Tetrahedron 73, 4392–4397.

Zhou, M., Zhang, R.-Q., Chen, Y.-J., Liao, L.-M., Sun, Y.-Q., Ma, Z.-H., Yang, Q.-F., Li, P., Ye, Y.-Q., Hu, Q.-F., 2018. Three new pyrrole alkaloids from the roots of *Lepidium meyenii*. Phytochem. Lett. 23, 137–140.

Zhou, M., Ma, H.Y., Xing, H.H., Li, P., Li, G.P., Geng, H.C., Hu, Q.F., Yang, G.Y., 2017. Biomimetic synthesis of macahydantoins A and B from *Lepidium meyenii*,

and structure revision of macahydantoin B as a class of thiohydantoin with a 4-methyl-hexahydropyrrolo[1,2-c]imidazole skeleton. Org. Lett. 19, 4952–4955.

Zhou, Y., Li, P., Brantner, A., Wang, H., Shu, X., Yang, J., Si, N., Han, L., Zhao, H., Bian, B., 2017. Chemical profiling analysis of Maca using UHPLC-ESI-Orbitrap MS coupled with UHPLC-ESI-QqQ MS and the neuroprotective study on its active ingredients. Sci. Rep. 7, 44660.

CHAPTER 8

Quality control of *Lepidium meyenii* (Maca)

Manuscript under review:

Le Ngoc-Thao-Hien, Kenn Foubert, Mart Theunis, Tania Naessens, Murat Bozdag, Pieter Van Der Veken, Luc Pieters, Emmy Tuenter. "Method validation and quality control of alkaloid content in *Lepidium meyenii* (Maca) containing food and food supplements".

Abstract:

Lepidium meyenii Walp. (Brassicaceae), also known as Maca or Peruvian ginseng, is a common ingredient in food supplements with many claimed health benefits, such as improved endurance, increased energy level and enhanced sexual properties. Due to potential toxicity of its alkaloids, some regulatory authorities, e.g. in Belgium, Germany, United States, expressed concerns about the safe consumption of Maca root. However, due to the lack of commercial standards, no established analytical method currently exists for this purpose. The current study presents the first analytical method for quality control of Maca-containing food supplements, assessing the presence of eleven major alkaloids belonging to three different classes, i.e. imidazole, β-carboline and pyrrole alkaloids. A rapid and sensitive UPLC-ESI-MS/MS method is reported, which was fully validated according to the ICH guidelines. To ensure the method's applicability and practicability in the absence of primary standards, validation of secondary standards (SS) alongside primary standards (PS) was also conducted. As a result, all quantified imidazole alkaloids were ascertained to be the major alkaloids, followed by the β-carboline and pyrrole alkaloids. Ultimately, from a regulatory point of view, it seems advisable not to request the absence of any detectable amount of alkaloids, but to impose a maximum level based on safety considerations.

Keywords: Alkaloids; *Lepidium meyenii*; Maca; UPLC-ESI-MS/MS; validated method; secondary standard

8.1. Introduction

Lepidium meyenii Walp. (Brassicaceae) or 'Maca' is an herbaceous plant native to the high plateaus of the Andes of Peru (Carvalho & Ribeiro, 2019). The climatic and environmental conditions in this region are exceptionally distinctive, characterized by high altitudes, low temperatures, humidity, strong winds, abundant rainfall and sunlight, low atmospheric pressure, as well as high levels of UV and cosmic radiation (Valerio Jr. et al., 2005; Chen et al., 2021). Growing and thriving in this harsh living conditions is considered a significant evolutionary factor of Maca, forcing the production of protective and biologically active compounds. Nowadays, Maca is grown not only in Peru, but also in some other South American countries, as well as in China (Chen et al., 2021). Maca is primarily used as food supplement due to its claimed medicinal properties and nutritional value, such as improved endurance and energy level, increased memory, enhanced fertility and sexual performance, and protection against UV radiation, to name a few (Tafuri et al., 2021; Beharry & Heinrich, 2018). Many studies have indicated that the distinctive presence of macamides and macaenes, which are fatty acid derivatives, in Maca likely plays a crucial role in the purported medicinal properties (Beharry & Heinrich, 2018). Glucosinolates and isothiocyanates, widely distributed in the Brassicaceae family, are two other major classes of compounds in Maca, which recently have been considered as new plant-derived antimicrobial agents (Huang et al., 2018).

In the early 2000s, the imidazole alkaloids lepidilines A and B were reported, as well as the β-carboline alkaloid 1-methyl-1,2,3,4-tetrahydro-β-carboline-3-carboxylic acid (MTCA) and macaridine (Muhammad et al., 2002; Cui et al., 2003). However, in 2021, the structure of macaridine was revised as macapyrrolin C, a pyrrole alkaloid (Zhou et al., 2018; Le et al., 2021). Since 2016, discovery of minor yet distinctive agents containing nitrogen and sulfur from Maca has emerged as an increasingly captivating research topic, disclosing novel classes of compounds including lepipyrrolins, meyeniins, hydantoins, thiohydantoins and macathioureas (Carvalho & Ribeiro, 2019; Liu et al., 2021; Geng et al., 2023).

Although in general the use of Maca is considered as safe, cytotoxicity has been reported for some of the imidazole alkaloids, including lepidilines A and B (Beharry & Heinrich, 2018). In addition, some local publications from Peru document the

toxicological effects of Maca on cell cultures and on mice (Valerio Jr. et al., 2005). Among all alkaloids, toxicity and pharmacological activities of β-carbolines have been studied extensively. On the one hand, some β-carbolines exhibited potent neuropharmacological, antitumor, antioxidant, anti-inflammatory and antimicrobial effects (Bower-Cargill et al., 2022; Xie et al., 2021). On the other hand, some β-carbolines were reported as co-mutagens or precursors of mutagens, neurotoxin and carcinogens, such as norharman, harman, 1-trichloromethyl-1,2,3,4-tetrahydro-β-carboline and aminophenyl-β-carboline derivatives (Jiménez et al., 2008; Xie et al., 2021). Although not directly related to Maca, a case of β-carboline alkaloids intoxication was reported, following the digestion of *Peganum harmala* seed extract (Frison et al., 2008). Therefore, concerns have been raised about the safety of Maca as food supplements, e.g. in Belgium, Germany and United States. More specifically, this has led to the legal requirement that no detectable amounts of alkaloids may be present in food supplements containing Maca (Belgian Royal Decree of 21/08/2021). According to the German Federal Institute for Risk Assessment (published 03/04/2007), insufficient information and concrete evidence for safety and quality control of maca-containing food supplements were available, and therefore no safe intake of Maca could be derived. Although Maca (*L. meyenii* and *L. peruvianum*) has been listed as EU Novel Food, no report for Maca from the European Food Safety and Authority (EFSA) was found at the present time. On the other hand, in the United States Pharmacopoeia (USP) safety review of Maca (published in 2012), 50–100 g of dried hypocotyls per day as food, and a dose range of 1.5 – 3.0 g of dried hypocotyls (corresponding to about 10 g of fresh maca) as dietary supplement were suggested. However, USP also documents numerous nonserious adverse effects of Maca consumption in humans, including rash, syncope, burning sensation, nausea, abdominal pain, vomiting, diarrhea, dehydration, impaired concentration, headache, epistaxis, anxiety, insomnia, tachycardia, and thirst.

Several advancements in the development of analytical methods for macamides, macaenes, glucosinolates and isothiocyanates in *L. meyenii* and some plant species were already made (Xia et al., 2019; S. X. Chen et al., 2017; Hwang et al., 2019; Theunis et al., 2022). However, quantitative analysis of the alkaloids has largely been overlooked. From an analytical perspective, the main challenge lies in the

scarcity of commercially available standards of Maca alkaloids that can be used as primary standards, which hampers the progress of analytical method development. Therefore, our study systematically addresses the issue in the following manner: (1) - conducting a phytochemical and metabolomic investigation to assess the alkaloidal composition of Maca and to isolate appropriate primary standards - this research was previously published by our group (Le et al., 2021); (2) - devising a straightforward synthetic scheme for generating pyrrole-containing standards; (3) - developing and validating an analytical method employing secondary standards, thus enabling the quality control of Maca even in the absence of primary standards, and (4) - as a last step, analyzing some commercial products available on the Belgian market using the method developed and validated hereinafter.

8.2. Results and Discussion

8.2.1. Method development

8.2.1.1. Optimization of mass spectrometric parameters

Initially, all compounds were analyzed in both positive and negative electrospray ionization (ESI) modes. The imidazole compounds (**1** – **7**) exclusively exhibited MS responses in positive mode, whereas the β-carbolines (**8** – **11**) and pyrroles (**12** – **14**) displayed signals in both modes (see **Fig. 1**). The absence of MS signals in negative mode for the imidazoles can be attributed to the positively charged nature of the quaternary imidazole moiety. Interestingly, this phenomenon also explains the exceptional sensitivity of imidazoles to mass spectrometry, as these compounds are inherently charged and do not require additional ionization for detection. Molecular ions $[M]^+$ can be detected for the imidazoles, while the pyrroles and β-carbolines showed the highest intensity as protonated adducts $[M+H]^+$.

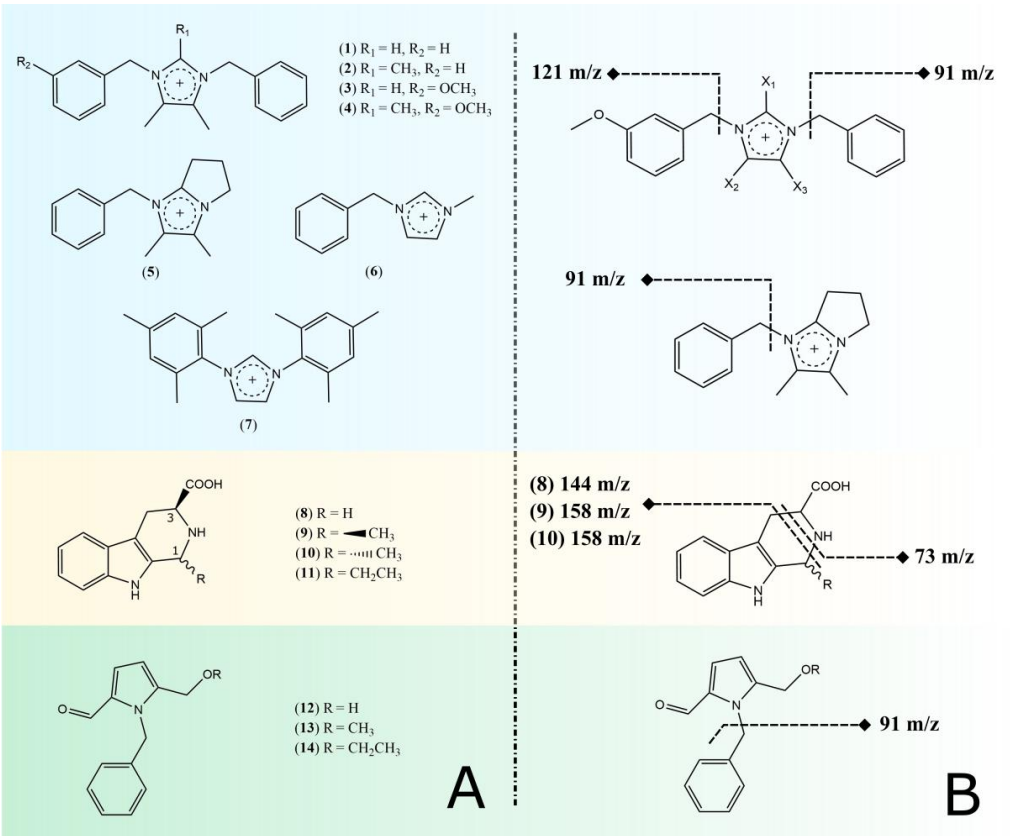


Fig. 1. A - Structures of compounds **1**-**14**: (1) lepidiline A, (2) lepidiline B, (3) lepidiline C, (4) lepidiline D, (5) lepidiline E, (6) 1-benzyl-3-methylimidazolium chloride (BMI) (SS), (7) 1,3-dimesitylimidazolidinium chloride (DMI) (IS), (8) 1,2,3,4-tetrahydro-β-carboline-3-carboxylic acid (TCCA), (9) 1-(R)-methyl-1,2,3,4-tetrahydro-β-carboline-3-(S)-carboxylic acid (RS-MTCA), (10) 1-(S)-methyl-1,2,3,4-tetrahydro-β-carboline-3-(S)-carboxylic acid (SS-MTCA), (11) 1-ethyl-1,2,3,4-tetrahydro-β-carboline-3-carboxylic acid (ETCA) (IS), (12) macapyrrolin C, (13) macapyrrolin A, (14) macapyrrolin G. **B** – Characteristic fragmentation patterns of all compounds

For the optimization of cone and collision voltages to be used in a multiple reaction monitoring (MRM) method, each standard was directly infused into the mass spectrometer in positive ESI mode. A quantifier (transition with the highest intensity) and a qualifier (transition used for identification) were identified for each compound. **Table S1** presents the optimal cone and collision energies for obtaining

the transitions of interest. Characteristic fragmentation patterns were identified for all three classes of compound, and **Fig. 1** demonstrates the stable and most intense fragments formed. The *m/z* 91 fragment was observed for lepidilines **1** – **6** and macapyrrolins **12** – **14**, corresponding to the benzylic moiety, while the *m/z* 121 fragment was observed only for lepidilines C and D (**3** – **4**), corresponding to the methoxylated benzylic moiety. The neutral loss of 73 Da is characteristic for the β-carbolines, resulting in the major fragment of *m/z* 144 for TCCA (**8**) and the *m/z* 158 for RS-MTCA (**9**) and SS-MTCA (**10**).

In the next step, the final experimental conditions for instrument operation, including capillary voltage, RF lens, source temperature, desolvation temperature and desolvation flow, were optimized using a design-of-experiment approach. Initially, a screening study was performed with the aim to select the most influential parameters and to reduce the complexity of the study. To achieve this, a standard mixture consisting of target compounds, internal and secondary standards was screened applying the established MRM method and the optimized chromatographic conditions described in section 2.1.2, which later on revealed the significant influence of temperature on the peak area. For this reason, central composite design was employed to optimize the two critical parameters: source temperature and desolvation temperature. The remaining parameters were fixed as follows: capillary voltage (3.5 kV), RF lens (0.1 V) and desolvation gas flow (900 L/h). As can be seen in the contour plots in **Fig. 2**, the variation of the peak area obtained for target compounds is presented based on the modification of the two influential factors, i.e., the source temperature (130 – 150 °C) and desolvation temperature (400 – 500 °C). The desolvation temperature displayed a much greater impact on the area than the source temperature at the tested conditions. The higher the desolvation temperature, the greater the area, whereas the influence of source temperature was solely observed for TCCA (**8**), MTCA (**9**-**10**) and macapyrrolin A (**13**). The optimal point (source temperature of 140 °C and desolvation temperature of 500 °C) provided the largest peak area for most of the target compounds (red zones), and reasonably high instrumental responses for quantifying TCCA (**8**), MTCA (**9**-**10**) and macapyrrolin A (**13**) (red-orange zones).

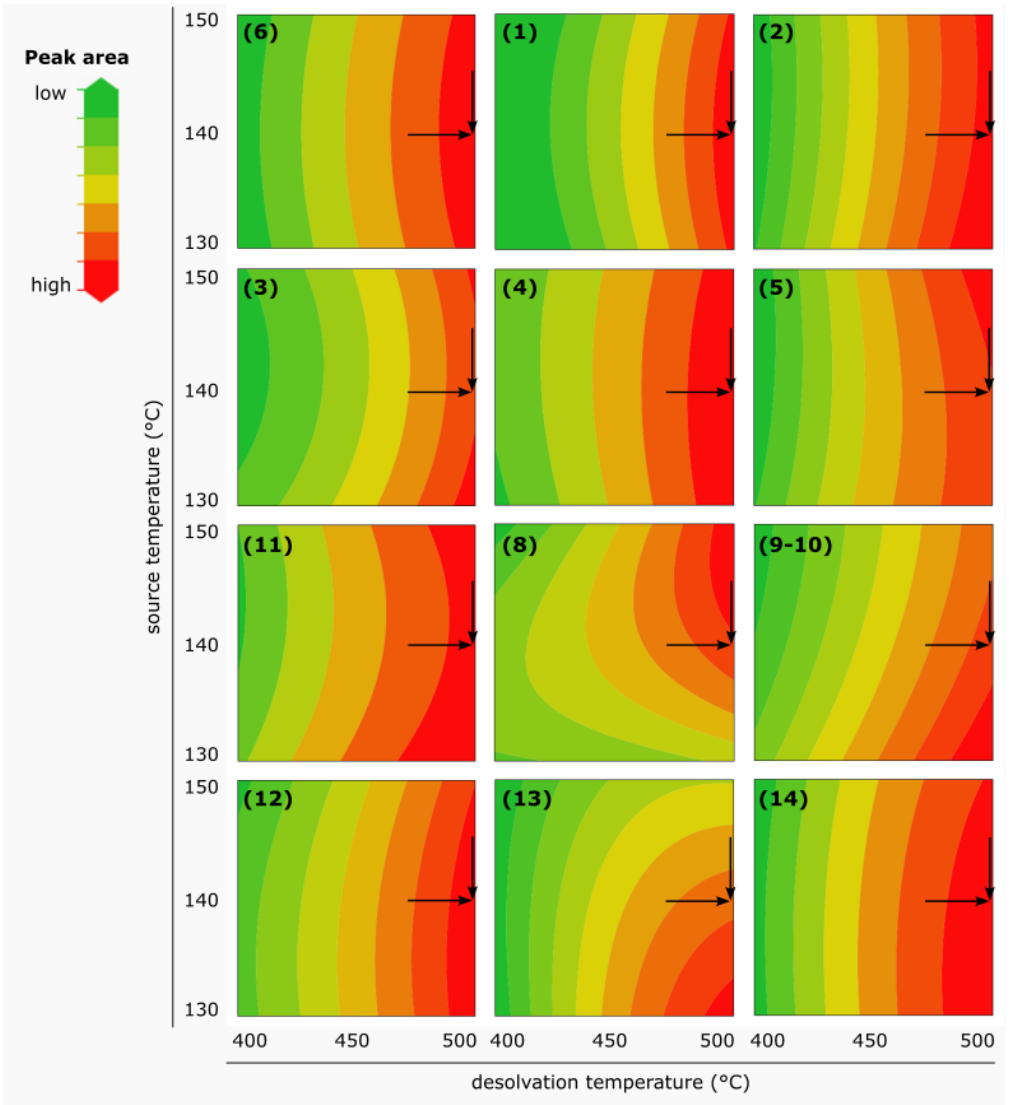


Fig. 2. Contour plots of all target compounds obtained from the central composite design for the optimization of source temperature (130 – 150 °C) and desolvation temperature (400 – 500 °C). Color gradient from green to red represents the increase of peak area as the optimized instrumental response.

8.2.1.2. Optimization of chromatographic conditions

At first, preliminary screening of chromatographic conditions was conducted for acidic (FA 0.1%) and basic (NH_4OH 0.1%) mobile phases with three different columns (BEH C18, BEH C18 Shield and HSS C18). The choice of FA 0.1% was straightforward, since peak splitting was observed for the β-carbolines when NH_4OH 0.1% was used as additive. This can be explained by the presence of a carboxylic group in the structures of the β-carbolines. On the other hand, FA could provide sharp peaks as none of the target compounds possesses strong basic properties. In all cases, the nitrogen is involved in an aromatic ring system where its free electron pair participates in the conjugated system instead of playing an electron-donating role. For the selection of the column, no significant difference was found between the tested ones. The BEH C18 was chosen due to its ubiquity.

The chromatographic separation was investigated with respect to resolution of the peaks and time-effectiveness. The target compounds exhibit a wide range of polarity: the β-carbolines are the most hydrophilic, while the lepidilines and the macapyrrolins are more hydrophobic. Therefore, the gradient slope is gradually increased in the first half of the chromatogram, allowing an adequate retention of the β-carbolines, and steadily increased in the second half, allowing the elution of the lepidilines and the macapyrrolins. Results indicated a chromatographic separation of all analytes, except for lepidilines B (**2**) and C (**3**) (**Fig. 3**). Separation of these two compounds cannot be obtained with FA as the additive, regardless of the gradient slope and the column used. While our previous study reported the use of a mixture of ammonium acetate and formic acid as additive for isolation of the two compounds, we opted for FA for reasons of simplicity and routine analytical application, considering the fact that chromatographic separation is not mandatory for an MRM method.

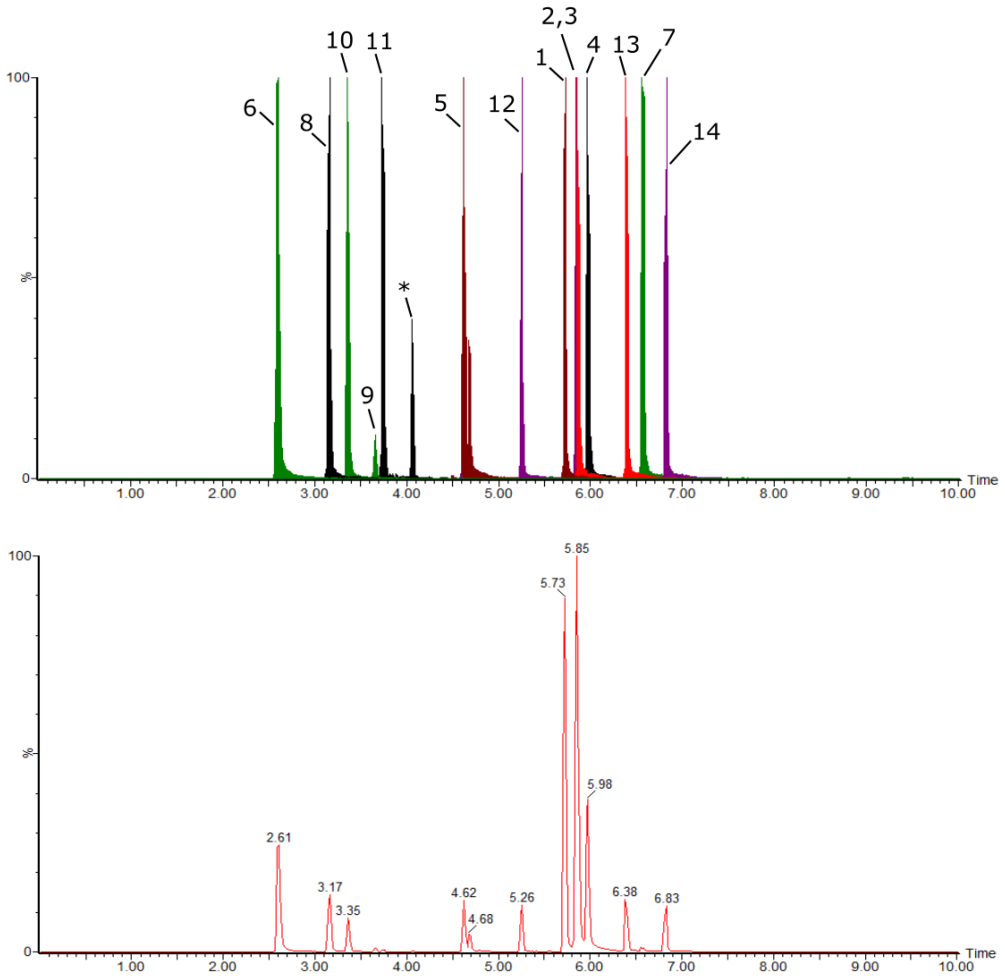


Fig. 3. Normalized EICs (upper) and MRM chromatogram (lower) of the optimal separation achieved for all compounds under the chromatographic conditions described in the experimental section. Co-elution occurs for lepidiline B and C (**2** and **3**). Commercial ETCA used in this study is a mixture of RS-ETCA and SS-ETCA. SS-ETCA, denoted as * on the chromatogram is not included in the method.

8.2.1.3. Optimization of sample preparation procedure

The sample preparation included extraction and purification. Sonication-assisted extraction was initially tested due to its high extraction efficiency and cost-effectiveness. Since the stability of the compounds of interest remained unaffected by sonication, this extraction method was adopted. As no marked changes of

ultrasonic extraction efficiency were observed when the extraction time exceeded 30 min, the optimal extraction conditions were determined as 30-min ultrasonic extraction with 2-3 time manual shaking (repeated three times). Knowledge obtained from our reported phytochemical study indicated the difference in polarity of the target compounds: the imidazoles and the pyrroles were isolated from the chloroform extract (lipophilic) and the β-carbolines from the *n*-BuOH extract (relatively hydrophilic). Therefore, optimization of the extraction solvent was required. To this end, six ratios of the aqueous / methanolic extraction solvents were evaluated with an increasing percentage of MeOH i.e. 0 – 20 – 40 – 60 – 80 – 100% MeOH. As a result, it was observed that the amount of β-carbolines was low when pure MeOH was used, indicating the necessity of adding a certain amount of water in the extraction solvent, while an obvious decline in the amount of the lepidilines was observed when H_2O, 20% MeOH, and 40% MeOH were used, implying the need of a high proportion of organic solvent. Considering the practicality and the fact that lepidilines are the major components among the alkaloids, 75% MeOH was selected as the optimal extraction solvent.

The impact of the plant sample matrix on the intensity of signal was assessed through SPE purification and dilution experiments. For solid-phase purification, several types of SPE cartridges were evaluated, including Oasis HLB, Oasis MCW, Chromabond C18, Chromabond HR-XC, and Chromabond HR-XCW. Among these, Oasis HLB demonstrated a straightforward procedure and satisfactory efficacy in removing impurities, particularly polar ones (see **Table S6**, **Table S7**, **Fig. S22 – S24**). Furthermore, the matrix effect was examined on finished product matrices (commercial products analyzed at the end of the study), with the Oasis HLB cartridge exhibiting superior cleaning capabilities. Consequently, the use of Oasis HLB was adopted for samples encompassing a wide variety of matrices.

The optimal amount of Maca powder, designated as 100% level, was established at 100 mg following an evaluation of different powder weights ranging from 50 mg to 500 mg. This quantity represents the minimum amount of plant material at which the presence of minor compounds (**9**, **10**, **13**) can be accurately quantified without any interference from the plant sample matrix. Compound **14**, while previously identified in Maca, was not even detected during the analysis of the 500

mg powder. Thus, this compound was spiked to Maca sample when validating the method.

At this point, screening results showed that the concentration of imidazoles (**1**-**5**) and macapyrrolin C (**12**) in pure Maca powder is approximately 10-50 times higher than that of the minor compounds. Additionally, the imidazoles exhibit a much higher sensitivity, roughly 3-5 times more than the other compounds, which poses challenges in accurately and plausibly quantifying all compounds within a single sample. Hence, we adopted the sample preparation procedure as depicted in **Fig. 4**, wherein two dilutions were analyzed for each sample: one for quantifying the minor compounds (first dilution - in a 25 mL volumetric flask; compounds **8** – **11**, **13**, **14**) and another for quantifying the major compounds (second dilution = 50-fold diluted solution; compounds **1** – **7**, **12**). After drying, a buffer (5 mM $HCOONH_4$ in 10% MeOH (pH 3)) was used for reconstituting and diluting samples for the following reasons: (1) – in case of MeOH proportions higher than 20%, the peak shape of the early eluting peaks was affected; (2) – an acidic medium allows the dissolution of the alkaloids in 10% MeOH; (3) – the use of the buffer allows a relatively matrix-matched quantification when samples are adequately diluted.

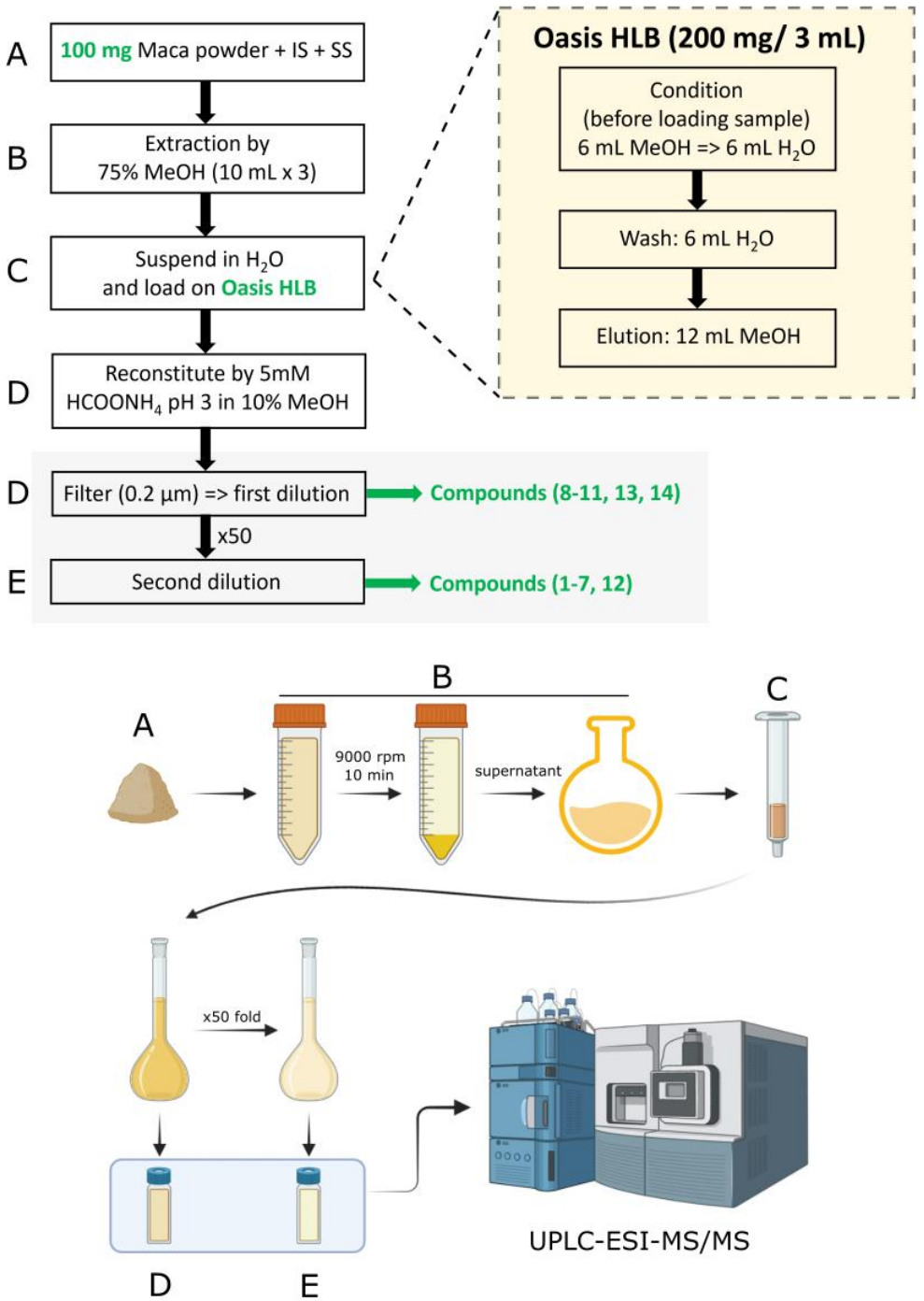


Fig. 4. Sample preparation procedure applied in this study with the double-dilution strategy.

8.2.1.3. Selection of internal and secondary standards

Internal and secondary standards were chosen with respect to the structural similarity with the target compounds and commercial availability. Since the quaternary imidazole nucleus determines the behavior of the lepidilines in the TQD (triple quadrupole detector), it is required that this moiety is maintained. BMI, DMI and DDI possess this moiety and are readily commercially available. After preliminary screening with the optimal operational conditions, BMI was selected as secondary standard and DMI as internal standard (IS) for the lepidilines, based on fragmentation pathways and chromatographic behavior. Only the lepidilines require a secondary standard, since they are not commercially available. Primary standards are commercially available for TCCA (**8**) and MTCA (**9** – **10**), and the macapyrrolins (**12** – **14**) were obtained in gram scale following a two-step synthetic scheme (see section 2.3). A synthetic β-carboline, ETCA, was also used as IS after checking for its absence in Maca raw powder. To compromise with the two-dilution strategy as described in the sample preparation section, ETCA was considered the IS of the first dilution and DMI of the second dilution. Macapyrrolins A and G were identified in the first dilution and their contents were corrected based on ETCA. Similarly, macapyrrolin C (**12**) was quantified in the second dilution and, therefore, its recovery was calculated according to DMI.

8.2.2. Method validation

The method was validated in terms of selectivity, linearity, LOD, LOQ, repeatability, recovery and robustness. Identification of target compounds was confirmed: the same retention time, quantifier and qualifier were observed for each compound in both processed samples and standard mixture, except for macapyrrolin G (**14**) which was not detected in commercial Maca samples.

For validation, concentrations of the target compounds were not equal in the standard mixture and were adjusted according to their sensitivity to avoid overloading the MS detector and contaminating the instrument. Specifically, in all dilutions of the standard mixture, concentrations of the β-carbolines doubled those of the macapyrrolins and were four times higher than those of the lepidilines. Logarithmic transformation of concentration and peak area prior to performing regression analysis provided extended linear ranges and fulfilled homoscedasticity.

The 11 analytes demonstrated good linearity ($R^2 \geq 0.9991$) in a wide concentration range (see **Table S2**). In parallel, response factors were established for the lepidilines from linearity data of two different days by comparing instrumental responses of each compound with BMI (**6**) at the same concentrations (**Table 1**). The response factors were stable in the concentration range of 109 – 56 250 pg/mL. It is noteworthy that the response factors slightly varied on a day-to-day basis, but still remained comparable.

Table 5. Response factors established for lepidilines

	SS-day1	**SS-day2**	**Mean-SS**
Lepidiline A (**1**)	2.0146 ± 0.0844	2.2400 ± 0.0092	2.1273 ± 0.0468
Lepidiline B (**2**)	2.3137 ± 0.1244	2.4444 ± 0.0118	2.3791 ± 0.0681
Lepidiline C (**3**)	0.5740 ± 0.0466	0.6047 ± 0.0553	0.5894 ± 0.0510
Lepidiline D (**4**)	0.5456 ± 0.0437	0.6607 ± 0.0072	0.6032 ± 0.0255
Lepidiline E (**5**)	0.5770 ± 0.0453	0.6149 ± 0.0154	0.5960 ± 0.0304

The LOD and LOQ of all compounds were determined, revealing varying sensitivity. The lepidilines were the most sensitive (LODs < 0.05 ppm = picogram/mL scale), while the β-carbolines were the least sensitive (LODs ≈ 1 ppm). Lepidilines A –D, due to their high content in samples and outstanding sensitivity, always appeared as the most intense peaks in any sample chromatogram in this study. Compared to major compounds reported from Maca, the total content of lepidilines A – E (around 400 ppm in this study) is relatively close to those previously reported for macamides and macaenes (80 - 3000 ppm) (Ganzera et al., 2002; Xia et al., 2021), thus providing a new insight into the chemical composition of Maca. Notably, it was reported that macamides and macaenes are by-products of air-drying process according to Xia and Esparza, while freeze-dried Maca contains solely around 80 ppm of macamides and macaenes (Esparza et al., 2015; Xia et al., 2021), which is significantly lower than the lepidilines content. Based on this result, we suggest considering the lepidilines as analytical chemical markers for identification of Maca. Our previous phytochemical study already confirmed the exclusive existence of lepidilines in Maca, together with macamides and some macaenes (Le et al., 2021; Carvalho & Ribeiro, 2019). Literature research also supports this

finding. More specifically, in *Lepidium sativum*, a known species in the *Lepidium* genus, no quaternary imidazole alkaloid is reported so far (Maier et al., 1998). Furthermore, regarding the sensitivity to MS detection, the lepidilines are much more sensitive to MS detection than macaenes and macamides (fatty acid derivatives), allowing the detection of very small amounts.

The intraday precision (run-to-run precision) showed RSDs < 8% for all compounds. For the interday precision (day-to-day precision), RSDs of all compounds were lower than 15%, except for macapyrrolin A (**Table S3**). Nonetheless, since macapyrrolin A appears to be a minor compound in comparison with the others, it is not deemed necessary to increase the amount of plant material for analysis, risking the impact of matrix effects and the possible contamination of the MS instrument. Accuracy was evaluated by the recovery study. The mean recovery of all analytes at three spiked concentration levels ranged from 82.4 to 122.4 % (**Table S3**), meaning the method was fully validated.

Minor shifts of retention time (± 0.3 min) were observed with slight changes of column temperature or when another BEH C18 column (of a different production batch) was used. Quantitative results, however, remained within RSD of 5%. Thus, the applicability of the method when the column is changed or with minor temperature variations in the column chamber is ensured.

8.2.3. Application of the validated method on commercial food supplements

A range of food supplements, available on the Belgian market, was analyzed for its alkaloidal constituents with the newly developed and validated method. The total alkaloid content was expressed as the overall content of the target compounds. As can be seen in **Table S4** and **Fig. 5**, the lepidilines (**1** – **5**) and macapyrrolin C (**12**) were the major alkaloids, and macapyrrolin A (**13**) was the minor one. Macapyrrolin G (**14**) was not detected in any sample.

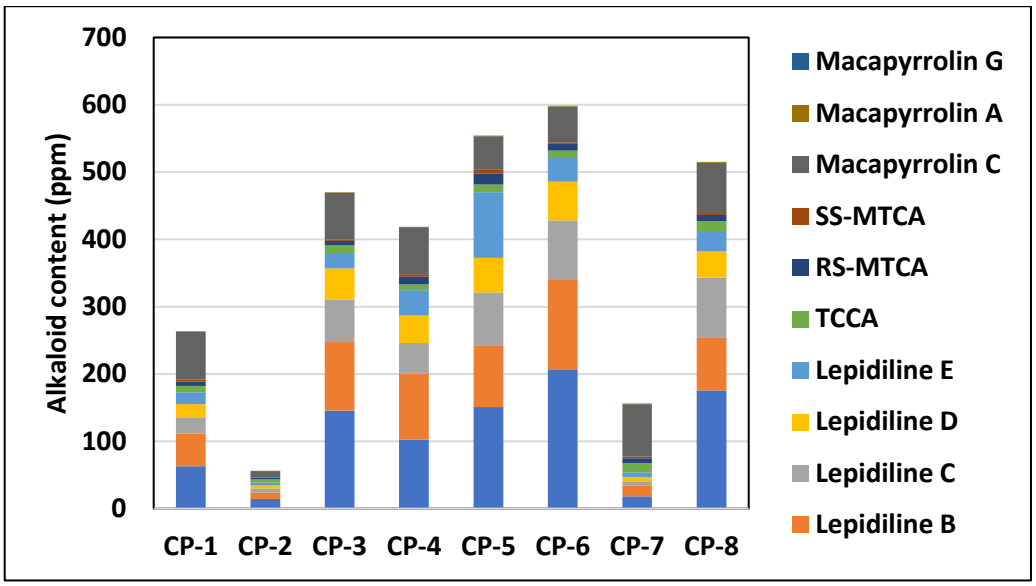


Fig. 5: Alkaloid content in commercial products (ppm) estimated by primary standards (n = 3)

In Maca raw powder (CP-4, CP-5 and CP-8), the total alkaloid content was approximately 500 ppm. As for the other commercial preparations (finished products) (CP-1, CP-2, CP-3, CP-6 and CP-7), total alkaloid content varied from 50 to 600 ppm. CP-2 contained the lowest alkaloid content, which could be expected since CP-2 was a mixture of several plant species. CP-7 was low in lepidilines, indicating a possible influence of different processing and preparation steps of finished products. Influence of different harvest times and locations might also be responsible for the observed differences in quantitative alkaloid content of the products.

The toxicity and pharmacological properties of β-carbolines were the most well-studied among all alkaloid classes. The current investigation of Maca-containing products revealed the total β-carboline content ranged from 20 to 30 ppm in all samples. TCCA (**8**) exists as a single isomer, while the diastereoisomers of MTCA (RS-MTCA (**9**) and SS-MTCA (**10**)) were found in all samples. It appears that the RS-MTCA content is roughly three times higher than the SS-MTCA content.

Table S5 shows the alkaloid content obtained from the established response factors using BMI (**6**) as secondary standard. Most quantitative errors were within 15% in comparison with those calculated by primary standards. Hence, the use of BMI as a secondary standard for the lepidilines can be considered as a valid alternative to the primary standard. With the obtained results in our analysis, the use of the averaged response factors as shown in **Table 1** in all cases led to quantitative errors less than 15%.

8.3. Conclusion

In recent years, Maca-containing food supplements have become ubiquitous on the European market. This raises a need for quality control of these products, since some alkaloids isolated from Maca were reported as potentially cytotoxic agents. However, the lack of commercial standards for the lepidilines and macapyrrolins hindered the development of an analytical method. The present work is the first report of an accurate, rapid, sensitive and comprehensive method for quality control of the most abundant alkaloids in raw material and Maca-containing food supplements. Eleven alkaloids, belonging to three different classes (imidazole, pyrrole and β-carboline alkaloids) were simultaneously determined. Nine alkaloids were identified as the main alkaloids (the lepidilines, the β-carbolines and macapyrrolin C) and two others as minor components (macapyrrolins A and G). This method was fully validated and offered good linearity, accuracy, repeatability, precision and robustness. More importantly, we also validated in parallel the use of a secondary standard (BMI) to quantify lepidilines. In addition a simple, easy-to-scale-up synthetic scheme of pyrrole alkaloids was also proposed. Thus, this validated UPLC-TQD-MS/MS method represents the first and a valuable method for both qualitative and quantitative assays of major alkaloids in Maca. The authenticity and adulteration of Maca can also be investigated using the current validated method. The exclusive existence of lepidilines in Maca and the permanently positively charged nature of the quaternary imidazole ring leads to an excellent sensitivity of the lepidilines towards MS detection in positive ion mode. Therefore, the lepidilines represent as excellent chemical markers for characterization of Maca in all types of commercial preparations (raw material, mixtures, mono-preparations, poly-preparations). From a regulatory point of view,

and based on safety concerns about the alkaloidal constituents of *L. meyenii*, it may be advisable to establish a maximum level of particular alkaloids or alkaloid classes in food supplements containing Maca, rather than to request their total absence.

8.4. Materials and Methods

8.4.1. Isolation and purification of imidazole alkaloids

Isolation of imidazole standards was reported before (Le et al., 2021). Lepidilines A – E (**1** – **5**) were obtained in sufficient amount (10-20 mg). Structure elucidation was performed by means of extensive spectroscopic and computational techniques using the published protocol of the authors (Le et al., 2022). Standard purity was determined by means of the total intensity method applied on the ^{1}H NMR spectrum of the purified compounds.

8.4.2. Synthesis of pyrrole alkaloids

Macapyrrolin C (**12**) was synthesized by a modified procedure reported by (Adhikary et al., 2015). The condensation of β-D-glucose and benzyl amine in acidic medium created by oxalic acid allows the formation of macapyrrolin C (yield ~40%). The use of readily available starting materials allowed a multi-gram scale synthesis of macapyrrolin C, which was later on used as the starting material for synthesizing macapyrrolins A (**13**) and G (**14**) by alkylation. The synthetic scheme is demonstrated in **Fig. 6**.

Gram-scale synthesis of macapyrrolin C (modified Maillard reaction of carbohydrate and amine) was carried out as follows: : in a 500 mL round-bottom flask containing 90 mL of DMSO, 33 g of β-D-glucose, 20 mL of benzyl amine (1 eq.), and 16.5 g oxalic acid (1 eq.) were mixed and then the mixture was heated to 80 °C for 5 h. The reaction was monitored by UPLC-MS and quenched by water. Macapyrrolin C (**12**): Brownish oil (18 g); UV λ_{max} 210, 274, 294 nm; ^{1}H and ^{13}C NMR ($CDCl_3$, 400 and 100 MHz): δ_H 9.53 (s, -CHO, H-6), 7.28 (overlapped, H-4', H-6'), 7.25 (m, H-5'), 6.99 (overlapped, H-3', H-7'), 6.96 (d, 3.9, H-3), 6.31 (d, 3.9, H-4), 5.75 (s, H-1'), 4.56 (H-7); δ_C 180.1 (-CHO, C-6), 142.4 (C-5), 138.1 (C-2'), 133.0 (C-2), 129.0 (C-4',C-6'), 127.7 (C-5'), 126.4 (C-3', C-7'), 124.7 (C-3), 111.1 (C-4), 56.9 (C-7), 48.8 (C-1'); Positive HRESIMS *m/z* 216.1016 $[M+H]^+$ (calcd for $C_{13}H_{14}NO_2$, 216.1025).

Synthesis of macapyrrolin A: 200 mg of macapyrrolin C was first added into a 50 mL flask and then 5 mL of THF was added. The solution was stirred in an ice bath until it cooled. 0.11g of NaH (5 eq.) was added and the solution was stirred in the ice bath for 20 min. Then 0.069 mL of iodomethane (2 M solution in tert-buthyl methyl ether) was added, followed by stirring for 5 h. The reaction was monitored by UPLC-MS and quenched by water. Macapyrrolin A (**13**): Brownish oil (109 mg); UV λ_{max} 210, 274, 294 nm; ^{1}H and ^{13}C NMR ($CDCl_3$, 400 and 100 MHz): δ_H 9.57 (s, -CHO, H-6), 7.28 (overlapped, H-4', H-6'), 7.25 (m, H-5'), 6.99 (overlapped, H-3', H-7'), 6.96 (d, 3.9, H-3), 6.31 (d, 3.9, H-4), 5.75 (s, H-1'), 4.35 (H-7), 3.33 (6-OCH_3); δ_C 180.1 (-CHO, C-6), 139.5 (C-5), 138.1 (C-2'), 133.2 (C-2), 128.9 (C-4',C-6'), 127.5 (C-5'), 126.5 (C-3', C-7'), 124.3 (C-3), 112.2 (C-4), 66.1 (C-7), 58.3 (7-OCH_3), 48.9 (C-1'); Positive ESIMS *m/z* 230.1179 $[M+H]^+$ (calcd for $C_{14}H_{16}NO_2$, 230.1181).

Synthesis of macapyrrolin G: 200 mg of macapyrrolin C was first added into a 50 mL flask and then 5 mL of THF was added. The solution was stirred in an ice bath until it cooled. 0.11 g of NaH (5 eq.) was added and the solution was stirred in the ice bath for 20 min. Then 0.069 mL of iodoethane was added, followed by stirring for 5. The reaction was monitored by UPLC-MS and quenched by water. Macapyrrolin G (**14**): Brownish oil (88 mg); UV λ_{max} 210, 274, 294 nm; ^{1}H and ^{13}C NMR ($CDCl_3$, 400 and 100 MHz): δ_H 9.57 (s, -CHO, H-6), 7.28 (overlapped, H-4', H-6'), 7.25 (m, H-5'), 6.99 (overlapped, H-3', H-7'), 6.96 (d, 3.9, H-3), 6.31 (d, 3.9, H-4), 5.75 (s, H-1'), 4.40 (H-7), 3.48 (q, 7.0, CH_2 of 7-OCH_2CH_3), 3.48 (t, 7.0, CH_3 of 7-OCH_2CH_3); δ_C 180.0 (-CHO, C-6), 140.0 (C-5), 138.2 (C-2'), 133.2 (C-2), 128.9 (C-4',C-6'), 127.5 (C-5'), 126.4 (C-3', C-7'), 124.4 (C-3), 112.1 (C-4), 66.2 (C-7), 64.2 (CH_2 of 7-OCH_2CH_3), 48.8 (C-1'), 15.3 (CH_3 of 7-OCH_2CH_3); Positive ESIMS *m/z* 244.1330 $[M+H]^+$ (calcd for $C_{13}H_{14}NO_2$, 244.1338).

Isolation and purification of macapyrrolins C, A, G: The reaction mixture was extracted three times with EtOAc. The EtOAc extracts were combined and subsequently washed with water (2x) and brine solution before drying over Na_2SO_4. Then, the extract was dried under reduced pressure to obtain a dark brown residue. Purification of macapyrrolin A was performed by flash chromatography on normal phase silica gel. Solvents comprised *n*-heptane and EtOAc, and the separation was obtained by gradually increasing the proportion of EtOAc in the

gradient. The following instruments were used for fractionation, isolation and structure elucidation: for flash chromatography, a Reveleris iES system from Grace (Columbia, MD, USA) using the Reveleris NavigatorTM software and commercially packed Claricep flash columns containing 40 g irregular deactivated silica gel (Agela Technologies, Wilmington, DE, USA); for analytical purposes, an Agilent 1200 series HPLC-DAD system (Agilent Technologies, Santa Clara, CA, USA) equipped with OpenLAB v.A.01.05 software, together with a Phenomenex Kinetex EVO C18 (250 x 4.6 mm, 5 µm) column. NMR spectra were recorded on an Avance Nanobay III (Bruker BioSpin, Rheinstetten, Germany). NMR data processing was performed with TopSpin v.4.0.6 from Bruker. Finally, accurate mass measurements were conducted for all isolated compounds using a Xevo G2-XS QToF mass spectrometer (Waters, Milford, MA, USA) coupled to an Acquity UPLC system; an Acquity BEH UPLC column (100 x 2.1 mm, 1.7 µm) was used.

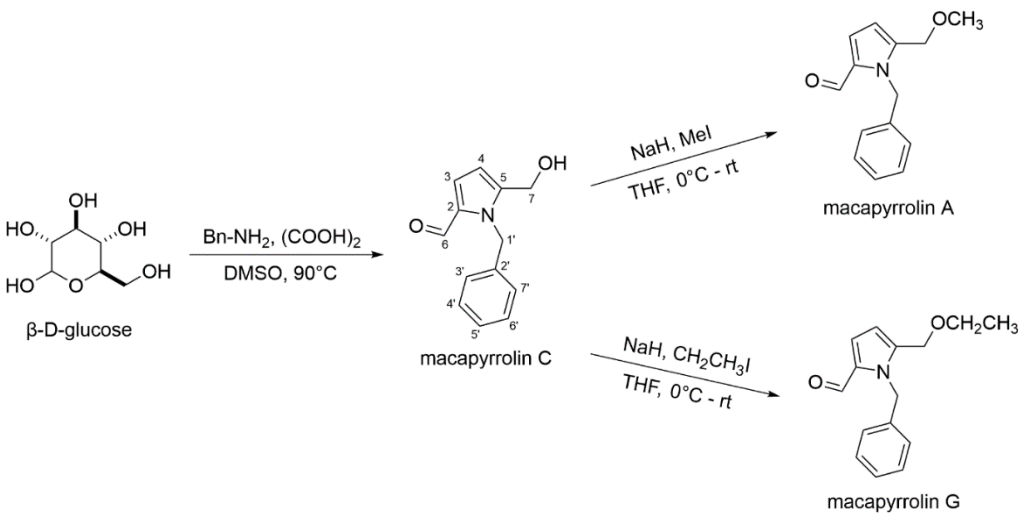


Fig. 6. Synthetic scheme of macapyrrolins C, A and G (**12** – **14**)

8.4.3. Sample collection

Method development was conducted using Maca root powder ("Bio-Maca Pulver", 1 kg), batch number 87619, supplied by Herbis Natura GmbH (Berlin, Germany). Seven representative commercial products were purchased on the Belgian market and were coded as CP-1 to CP-7. Of these, CP-1, CP-6, and CP-7 were capsule formulations containing Maca powder / extract, conservatives and additives; CP-2,

and CP-3 were tablets; CP-4 a pure powder; and CP-5 a hydroglycerin flour. Among these, CP-2 was the sole finished product in which Maca was mixed with other plant species.

8.4.4. Instrumentation

8.4.4.1. Chromatographic conditions

An Acquity ultra-performance liquid chromatography (UPLC) system, consisting of an autosampler and a binary pump (Waters, Milford, MA) equipped with a 11 µL loop, was used. Compounds were separated on an Acquity BEH C18 column (2.1 mm × 100 mm, 1.7 mm; Waters, Milford, MA). The column and autosampler were maintained at a temperature of 40 and 7 °C, respectively. 5 µL was injected using full loop injection. A flow rate of 0.5 mL/min was chosen and the following gradient was used: solvent A = water + 0.1% formic acid; solvent B = acetonitrile + 0.1% formic acid; gradient: 0 – 1 min (5% B), 1 - 3 min (from 5 to 15% B), 3 – 9 min (from 15 to 95% B), 9 – 10 min (95% B), 10 – 10.5 min (from 95 to 5% B), 10.5 – 12 min (5% B).

8.4.4.2. Mass spectrometric conditions

The UPLC system was coupled to a triple quadrupole (TQD) mass spectrometer (Waters, Milford, MA) equipped with an electrospray ionization (ESI) source operated in the positive ionization mode. The quantification of the molecular ion (imidazole alkaloids) and the protonated adducts (for all other compounds) was performed using the multiple reaction monitoring (MRM) mode to increase sensitivity and selectivity. A quantifier and a qualifier were determined for every compound using direct infusion at mixed mode. The optimal conditions were as follows: capillary voltage 3.5 kV, extractor voltage 4 V, source temperature 140 °C, desolvation temperature 500 °C, RF lens 0.1 V, desolvation gas flow 900 L/h, cone gas flow 20 L/h. The quadrupole was set for maximum resolution. All data were recorded and processed using the Quanlynx package in Masslynx software, version 4.1 (Waters, Milford, MA).

8.4.5. Sample preparation

The quantification of alkaloids in Maca powder was carried out using the following optimized procedure. Initially, approximately 100 mg of dried powder with internal and secondary standards added was sonicated in 10 mL of 75% methanol (v/v) for 30 min. Concentrations of added standards at the final solution: ETCA (50 ng/mL), BMI (5 ng/mL) and DMI (5 ng/mL). After centrifugation at 9000 rpm (RCF or G-force = 9050) for 5 min, the supernatant was transferred into a round-bottom flask. This extraction process was repeated two more times, and the three extracts were combined and subsequently dried under reduced pressure. The resulting dried sample was then reconstituted in 3 mL of water and loaded onto a SPE HLB cartridge (200 mg – 3 mL, Waters, Milford, MA) that had been pre-conditioned with methanol and water. The column was rinsed with 6 mL of water. Finally, the alkaloids were eluted from the column using 12 mL of methanol. The eluent was subsequently dried under vacuum before being transferred into a 25 mL volumetric flask. For the first solution, the sample was transferred and brought to the mark using a 5 mM $HCOONH_4$ pH 3 in 10% MeOH buffer. The sample was then further diluted 50-fold to obtain the second dilution. The quantification of macapyrrolin A, macapyrrolin G and β-carbolines took place in the first dilution, whilst all lepidilines and macapyrrolin C were quantified in the second dilution.

8.4.6. Method validation

Validation was performed according to the ICH guidelines for physico-chemical method validation and SANTE/11312/2021 guideline for LC-MS related specifications issued by European Union. The following validation criteria were investigated: selectivity, linearity, LOD, LOQ, accuracy, precision, matrix effect, and robustness.

8.4.6.1. Selectivity

Selectivity was determined using the optimal UPLC-TQD-MS conditions, and the obtained signals were compared with those of standard mixture in terms of retention time, precursor and product ions. The evaluation of selectivity involved identifying both quantitative and qualitative transitions for each compound. The selectivity of a compound is determined by the ratio between the quantifier and

the qualifier. This ratio was examined for every compound in primary standard solutions and processed samples.

8.4.6.2. Linearity, response factor and calibration model

Linearity was first examined in a wide concentration range for all compounds (0.1 – 1000 ng/mL) to identify linear ranges. Regression analysis of concentration and peak area was performed with and without log-log transformation. Analysis of variance, significance test of regression coefficient and homoscedasticity were examined and compared between the two models.

After obtaining linear ranges, response factors between the lepidilines (**1**-**5**) and BMI (**6**) were established as the ratio between a compound of interest with BMI at the same concentration, calculated form two independent linearity data of two days.

8.4.6.3. Limit of detection (LOD) and limit of quantification (LOD)

LODs and LOQs were identified at the concentrations that provided a signal-to-noise ratio higher than 3 and 10, respectively, from three consecutive injections.

8.4.6.4. Matrix effect

Matrix effects were first investigated by dilution tests. A series of 5 three-fold dilutions were run and calibration curves of lepidilines were compared. The procedure was carried out for a sample prepared from 500 mg of Maca powder.

8.4.6.5. Precision

To determine the intermediate precision, six independently prepared samples weighing 100% (100 mg) were analyzed using the method described above. This procedure was repeated on three different days. To assess repeatability across different concentration levels, six samples weighing 50% (50 mg) and six samples weighing 200% (200 mg) of the nominal mass were also analyzed. Standard mixture solutions were prepared daily: imidazole alkaloids (0.015 – 250 ng/mL), β-carbolines (0.060 – 1000 ng/mL), and pyrrole alkaloids (0.030 – 500 ng/mL). The standard mixture solutions were injected twice: once before the samples for

calibration, and once at the end of the sequence to assess the stability of the MS signal.

As for macapyrrolin G, the same set of samples above was used, and samples were spiked at three concentration levels (50%, 100%, and 200%). Among these levels, the 50% concentration was specifically chosen to align with the LOQ level.

8.4.6.6. Accuracy

To investigate the accuracy of the method, a recovery experiment was performed. To 50% of the plant material (50 mg) a known amount of standards was added until a total concentration of around 75%, 100% and 125%. For each level, samples were prepared in triplicate according to the described procedure.

8.4.6.7. Robustness

Three independent samples weighing 100% (100 mg) were prepared for this investigation. The ruggedness of the validated method was examined for a minor variation of column temperature (38 – 42 °C) and a different column batch. Quantitative results between the optimal and the changed methods were compared. The method was considered robust if deviation between them was within ± 10%.

8.4.7. Statistical analysis

Statistical analyses were performed by Excel 2016 with the add-in Analysis ToolPak. Design of experiment (central composite design) was modeled by JMP Pro 16.0.0 (JMP Statistical Discovery LLC).

Table 1. MS/MS transitions, cone and collision energies selected for the target compounds, internal and secondary standards

Compound	Chemical formula	Retention time (min)	Molecular weight	Ion / Adduct	Precursor ion (*m/z*)	Cone energy (V)	Quantifier (*m/z*), collision energy (V)	Qualifier (*m/z*), collision energy (V)
Lepidiline A (**1**)	$C_{19}H_{21}N_2$	5.69	277.1705	$[M]^+$	277.18	42	90.95 (26)	65.02 (56)
Lepidiline B (**2**)	$C_{20}H_{23}N_2$	5.83	291.1861	$[M]^+$	291.12	44	90.95 (32)	65.02 (68)
Lepidiline C (**3**)	$C_{20}H_{23}N_2O$	5.81	307.1810	$[M]^+$	307.18	38	91.01 (44)	121.04 (28)
Lepidiline D (**4**)	$C_{21}H_{25}N_2O$	5.94	321.1967	$[M]^+$	321.25	42	121.03 (30)	91.01 (40)
Lepidiline E (**5**)	$C_{15}H_{19}N_2$	4.60	227.1548	$[M]^+$	227.05	44	90.95 (32)	65.02 (50)
BMI (**6**) - SS	$C_{11}H_{13}N_2$	2.60	173.1078	$[M]^+$	172.99	26	90.95 (18)	96.93 (16)
DMI (**7**) - IS	$C_{21}H_{25}N_2$	6.49	305.2018	$[M]^+$	305.21	60	91.01 (42)	145.02 (24)
TCCA (**8**)	$C_{12}H_{12}N_2O_2$	3.14	216.0899	$[M+H]^+$	217.00	26	144.05 (14)	73.92 (28)
MTCA (**9**)	$C_{13}H_{14}N_2O_2$	3.34	230.1055	$[M+H]^+$	231.00	20	157.98 (14)	76.98 (30)
MTCA (**10**)	$C_{13}H_{14}N_2O_2$	3.65	230.1055	$[M+H]^+$	231.00	20	157.98 (14)	76.98 (30)
ETCA (**11**) - IS	$C_{14}H_{16}N_2O_2$	3.73	244.1212	$[M+H]^+$	245.07	20	182.06 (22)	167.23 (30)
Macapyrrolin C (**12**)	$C_{13}H_{13}NO_2$	5.24	215.0946	$[M+H]^+$	216.00	24	90.95 (20)	64.95 (44)
Macapyrrolin A (**13**)	$C_{14}H_{15}NO_2$	6.39	229.1103	$[M+H]^+$	230.07	20	90.95 (20)	198.06 (10)
Macapyrrolin G (**14**)	$C_{15}H_{17}NO_2$	6.82	243.1259	$[M+H]^+$	244.07	18	90.96 (26)	170.00 (14)

Table 2. Method validation parameters (linearity) for quantitative determination of target compounds

Compound	LOD (pg/mL)[a]	LOD (ppm)[b]	LOQ (pg/mL)[a]	LOQ (ppm)[b]	Concentration range (pg/mL)	Equation	R^2
Lepidiline A (**1**)	3.65	0.013	12.40	0.043	54 – 112500	Y = 0.8156X + 0.2418	0.9998
Lepidiline B (**2**)	3.08	0.012	10.51	0.040	54 – 56250	Y = 0.8216X + 0.2452	0.9998
Lepidiline C (**3**)	5.52	0.028	19.85	0.091	109 – 112500	Y = 0.8278X + 0.1068	0.9995
Lepidiline D (**4**)	4.91	0.025	16.67	0.081	109 – 112500	Y = 0.8395X + 0.0922	0.9998
Lepidiline E (**5**)	3.58	0.041	11.96	0.135	109 – 112500	Y = 0.8362X + 0.1012	0.9997
BMI (**6**) - SS	8.29	n.i.	28.27	n.i.	109 – 112500	Y = 0.7849X + 0.2119	0.9991
TCCA (**8**)	102.36	1.229	432.14	4.055	439 – 225000	Y = 1.1548X – 0.0364	0.9997
RS-MTCA (**9**)	189.46	1.455	662.66	4.800	817 – 418500	Y = 1.1439X – 0.1097	0.9994
SS-MTCA (**10**)	118.70	0.815	415.85	2.690	492 – 63000	Y = 1.1315X – 0.0819	0.9996
Macapyrrolin C (**12**)	33.45	0.354	114.03	1.168	219 - 225000	Y = 0.8098X – 0.0164	0.9991
Macapyrrolin A (**13**)	50.76	0.205	169.49	0.677	219 - 225000	Y = 1.1724X – 0.0421	0.9998
Macapyrrolin G (**14**)	43.37	n.i.	140.51	n.i.	219 - 450000	Y = 1.0242X – 0.1229	0.9996

[a]LOD, LOQ identified from standard mixtures

[b]LOD, LOQ estimated on Maca raw powder

n.i. = not identified

Table 3. Method validation parameters (precision, accuracy) for quantitative determination of target compounds

Compound	**Intermediate precision (RSD %)**				**Accuracy (recovery %)**		
	Intraday (n=18)	Interday (n=18)	Within levels (n=30)*	Between levels (n=30)*	75%	100%	125%
Lepidiline A (**1**)	5.38	9.88	4.81	8.01	114.1	113.9	99.9
Lepidiline B (**2**)	4.36	10.13	3.94	8.20	100.6	106.1	98.9
Lepidiline C (**3**)	4.84	11.42	4.24	8.91	111.0	100.9	111.6
Lepidiline D (**4**)	3.54	10.55	3.05	7.78	100.3	102.7	102.3
Lepidiline E (**5**)	2.91	12.05	3.30	8.90	100.8	102.7	98.9
TCCA (**8**)	5.05	9.97	4.56	9.02	82.8	116.6	102.0
RS-MTCA (**9**)	2.67	6.03	2.98	5.22	97.7	98.2	100.0
SS-MTCA (**10**)	1.97	2.36	2.69	5.59	101.5	90.7	102.5
Macapyrrolin C (**12**)	5.67	6.90	5.24	6.46	117.5	113.9	102.2
Macapyrrolin A (**13**)	5.28	29.85	7.22	61.62	122.4	109.3	107.1
Macapyrrolin G (**14**)*	1.43	1.71	2.00	4.16	114.7	102.7	102.8

*repeatability and intermediate precision on three different concentrations (50 – 100 – 200 %)

**spiked standard

Table 4. Alkaloid content in commercial products (ppm) estimated by primary standards (n = 3)

	CP-1	CP-2	CP-3	CP-4	CP-5	CP-6	CP-7	CP-8
Lepidiline A (1)	62.94 ± 1.96	14.56 ± 0.15	145.90 ± 4.20	102.59 ± 4.62	151.01 ± 0.18	206.02 ± 4.60	18.18 ± 0.73	175.80 ± 1.13
Lepidiline B (2)	48.40 ± 1.98	9.37 ± 0.06	101.72 ± 0.70	98.33 ± 6.95	91.78 ± 3.49	133.72 ± 1.88	15.92 ± 0.26	77.65 ± 0.50
Lepidiline C (3)	24.40 ± 0.59	6.06 ± 0.14	63.11 ± 0.29	45.20 ± 1.52	78.55 ± 0.75	88.35 ± 0.97	6.59 ± 0.37	89.65 ± 2.03
Lepidiline D (4)	19.71 ± 0.66	4.39 ± 0.07	46.21 ± 1.85	40.74 ± 1.12	51.14 ± 1.40	57.59 ± 0.74	6.04 ± 0.40	38.89 ± 1.18
Lepidiline E (5)	17.38 ± 0.42	4.43 ± 0.07	22.86 ± 0.49	37.89 ± 2.86	97.60 ± 0.70	35.37 ± 0.38	7.06 ± 0.23	30.71 ± 0.53
TCCA (8)	9.33 ± 0.24	4.65 ± 0.16	11.14 ± 0.35	8.47 ± 1.64	11.17 ± 0.41	11.06 ± 0.67	13.90 ± 0.49	14.66 ± 0.39
RS-MTCA (9)	7.02 ± 0.18	2.57 ± 0.08	6.61 ± 0.31	11.90 ± 0.28	17.01 ± 0.17	9.65 ± 0.24	6.93 ± 0.21	9.15 ± 0.07
SS-MTCA (10)	2.42 ± 0.09	0.90 ± 0.04	2.08 ± 0.06	3.27 ± 0.05	5.07 ± 0.12	2.61 ± 0.03	2.42 ± 0.07	2.63 ± 0.02
Macapyrrolin C (12)	71.47 ± 3.28	9.06 ± 0.28	69.41 ± 3.54	69.75 ± 8.41	50.17 ± 1.14	53.21 ± 3.62	78.08 ± 4.61	74.67 ± 2.59
Macapyrrolin A (13)	0.30 ± 0.01	0.03 ± 0.00	1.12 ± 0.02	0.46 ± 0.01	0.56 ± 0.01	0.80 ± 0.04	1.24 ± 0.06	1.45 ± 0.05
Macapyrrolin G (14)	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.
Total (ppm)	263.38 ± 9.41	56.02 ± 1.04	470.16 ± 11.80	418.60 ± 27.47	554.07 ± 8.36	598.37 ± 13.17	156.36 ± 7.44	515.25 ± 8.48

n.d. = not detected

Table 6. Lepidiline content in commercial products (CP) in ppm estimated by BMI (**6**) as secondary standard (n = 3)

		CP-1	CP-2	CP-3	CP-4	CP-5	CP-6	CP-7	CP-8
Content (ppm)	Lepidiline A (**1**)	70.34	16.13	169.22	112.86	167.81	231.62	19.49	197.96
SS-day1	Lepidiline B (**2**)	53.63	10.20	117.92	108.18	100.66	149.72	16.63	85.49
	Lepidiline C (**3**)	27.69	6.81	74.26	50.38	88.63	99.73	7.01	100.73
	Lepidiline D (**4**)	21.67	4.72	54.09	44.44	56.43	63.82	6.00	42.02
	Lepidiline E (**5**)	19.09	4.82	25.51	40.63	109.78	37.69	6.98	32.32
Content (ppm)	Lepidiline A (**1**)	62.74	14.39	151.14	100.81	149.90	206.99	17.44	177.07
SS-day2	Lepidiline B (**2**)	50.54	9.61	111.18	102.00	94.92	141.21	15.69	80.67
	Lepidiline C (**3**)	26.17	6.44	70.24	47.66	83.83	94.35	6.64	95.34
	Lepidiline D (**4**)	19.34	4.21	48.32	39.71	50.42	57.05	5.37	37.60
	Lepidiline E (**5**)	17.82	4.50	23.83	37.95	102.55	35.22	6.53	30.22
Error (%)*	Lepidiline A (**1**)	11.75	10.80	15.98	10.01	11.13	12.42	7.23	12.60
SS-day1	Lepidiline B (**2**)	10.80	8.81	15.92	10.01	9.68	11.96	4.46	10.11
	Lepidiline C (**3**)	13.46	12.50	17.67	11.47	12.84	12.88	6.40	12.36
	Lepidiline D (**4**)	9.95	7.51	17.06	9.08	10.35	10.83	0.68	8.06
	Lepidiline E (**5**)	9.87	8.82	11.57	7.21	12.47	6.55	1.16	5.25
Error (%)*	Lepidiline A (**1**)	0.32	1.18	3.59	1.73	0.74	0.47	4.05	0.72
SS-day2	Lepidiline B (**2**)	4.40	2.52	9.30	3.73	3.42	5.60	1.41	3.90
	Lepidiline C (**3**)	7.25	6.33	11.29	5.43	6.73	6.79	0.73	6.35
	Lepidiline D (**4**)	1.90	4.08	4.58	2.54	1.41	0.93	11.10	3.32
	Lepidiline E (**5**)	2.55	1.57	4.22	0.16	5.07	0.43	7.56	1.59

*Errors compared to alkaloid content obtained from primary standards in **Table 4**

8.5. Supporting information – Chapter 8

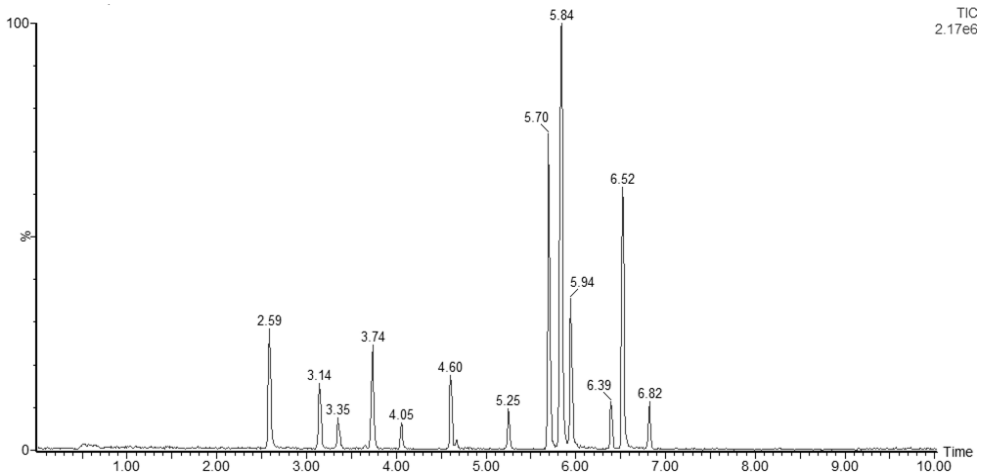


Fig. S1. Representative chromatogram of the standard mixture. Concentrations of the targeted compounds were not equal in the standard mixture and were adjusted according to their sensitivities to avoid overloading the MS detector and contaminating the instrument. In this example, the concentration of lepidilines (**1**-**5**), BMI (**6**) and DMI (**7**) was 3.5 ng/mL. The concentration of TCCA (**8**), MTCA (**9**-**10**), and ETCA (**11**) was 14.0 ng/mL. The concentration of macapyrrolins (**12**-**14**) was 7.0 ng/mL.

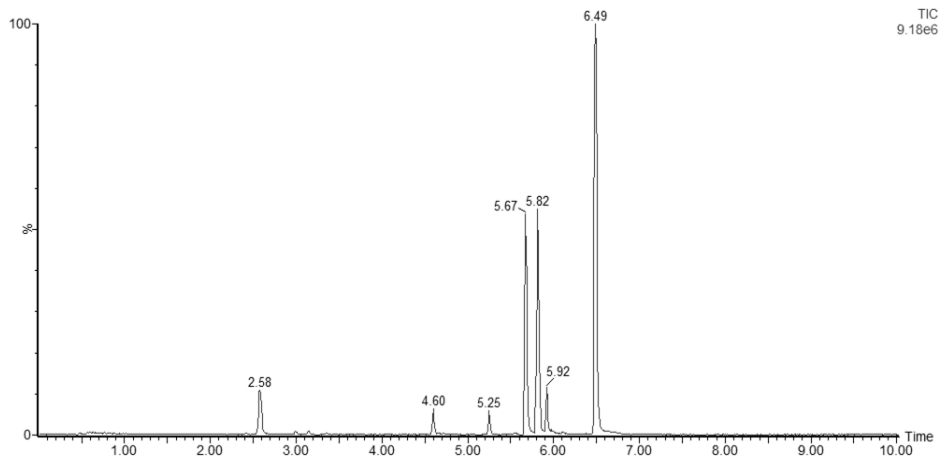


Fig. S2. Representative chromatogram of the CP-1

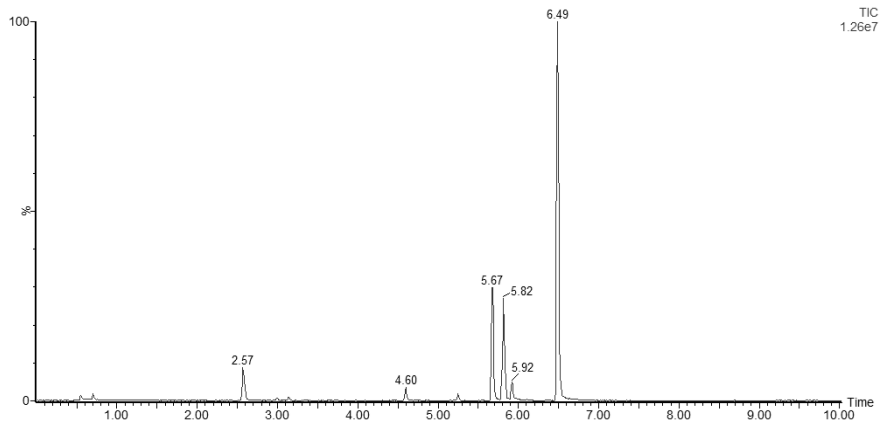


Fig. S3. Representative chromatogram of the CP-2

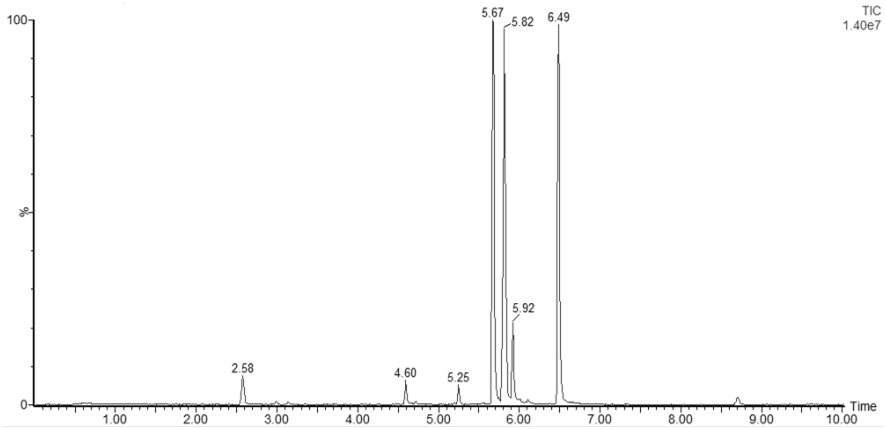


Fig. S4. Representative chromatogram of the CP-3

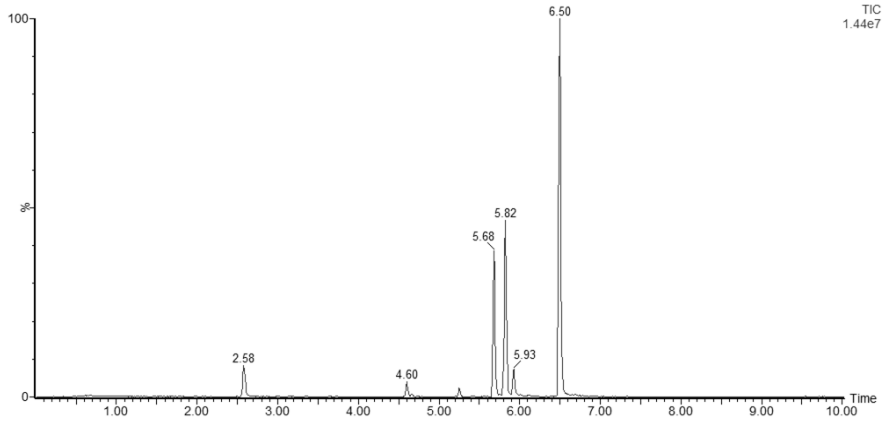


Fig. S5. Representative chromatogram of the CP-4

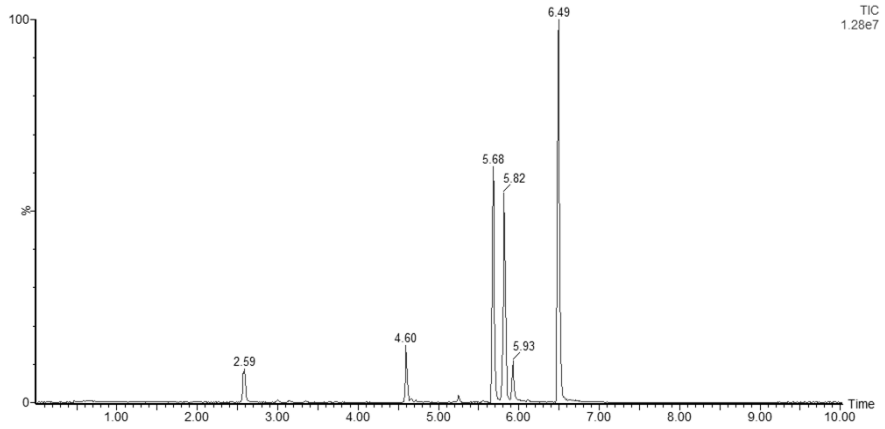


Fig. S6. Representative chromatogram of the CP-5

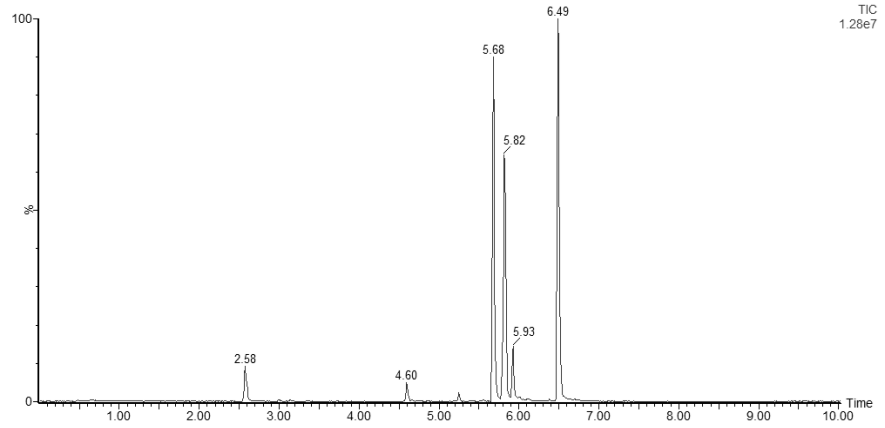


Fig. S7. Representative chromatogram of the CP-6

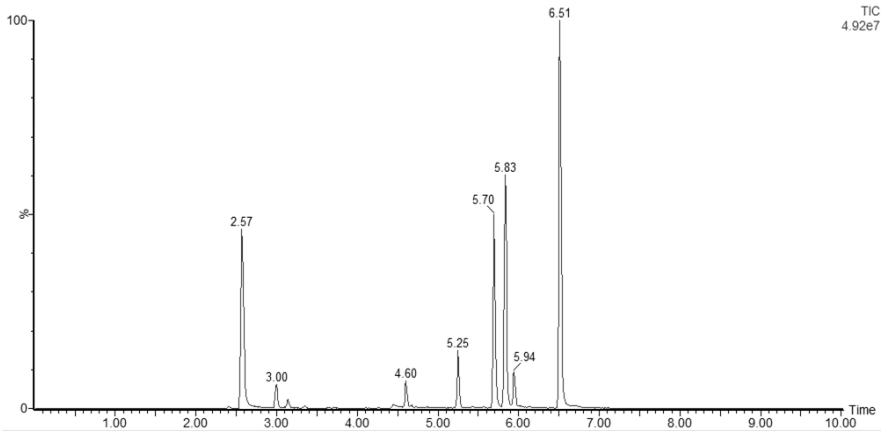


Fig. S8. Representative chromatogram of the CP-7

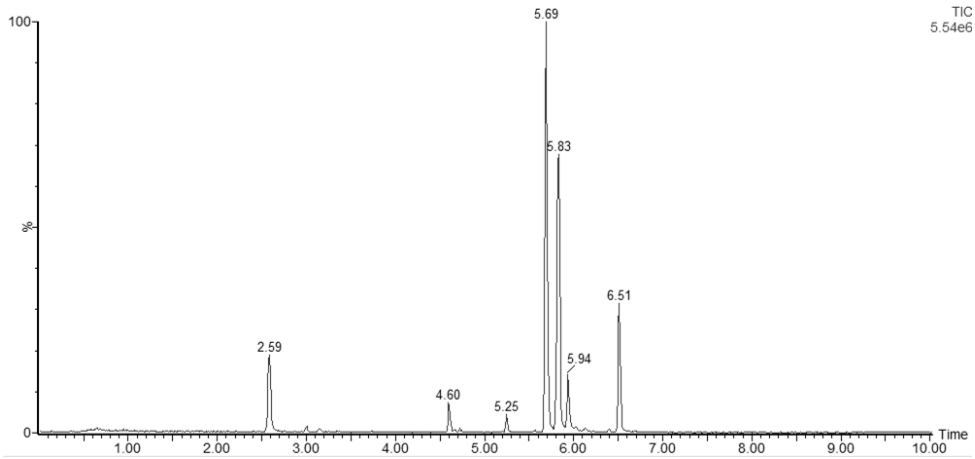


Fig. S9. Representative chromatogram of the CP-8

8.6. References

Adhikary, N. Das, Kwon, S., Chung, W. J., & Koo, S. (2015). One-Pot Conversion of Carbohydrates into Pyrrole-2-carbaldehydes as Sustainable Platform Chemicals. *Journal of Organic Chemistry*, *80*(15), 7693–7701.

Beharry, S., & Heinrich, M. (2018). Is the hype around the reproductive health claims of maca (*Lepidium meyenii* Walp.) justified? *Journal of Ethnopharmacology*, *211*(July 2017), 126–170.

Bower-Cargill, C., Yarandi, N., & Petróczi, A. (2022). A systematic review of the versatile effects of the Peruvian Maca Root (*Lepidium meyenii*) on sexual dysfunction, menopausal symptoms and related conditions. *Phytomedicine Plus*, *2*(4).

Carvalho, F. V., & Ribeiro, P. R. (2019). Structural diversity, biosynthetic aspects, and LC-HRMS data compilation for the identification of bioactive compounds of *Lepidium meyenii*. *Food Research International*, *125*(August).

Chen, R., Wei, J., & Gao, Y. (2021). A review of the study of active components and their pharmacology value in *Lepidium meyenii* (Maca). *Phytotherapy Research*, *35*(12), 6706–6719.

Chen, S. X., Li, K. K., Pubu, D., Jiang, S. P., Chen, B., Chen, L. R., Yang, Z., Ma, C., & Gong, X. J. (2017). Optimization of Ultrasound-assisted extraction, HPLC and

UHPLC-ESI-Q-TOF-MS/MS analysis of main macamides and macaenes from maca (cultivars of *Lepidium meyenii* Walp). *Molecules*, *22*(12).

Cui, B., Zheng, B. L., He, K., & Zheng, Q. Y. (2003). Imidazole Alkaloids from *Lepidium meyenii*. *Journal of Natural Products*, *66*(8), 1101–1103.

Frison, G., Favretto, D., Zancanaro, F., Fazzin, G., & Ferrara, S. D. (2008). A case of β-carboline alkaloid intoxication following ingestion of *Peganum harmala* seed extract. *Forensic Science International*, *179*(2–3), 37–43.

Geng, H. C., Wang, X. S., Chen, N., Lv, J. J., & Zhou, M. (2023). Macathioureas A–D, four new thiourea analogues from the roots of *Lepidium meyenii*. *Fitoterapia*, *168*(April), 105521.

Huang, Y.-J., Peng, X.-R., & Qiu, M.-H. (2018). Progress on the Chemical Constituents Derived from Glucosinolates in Maca (*Lepidium meyenii*). *Natural Products and Bioprospecting*, *8*(6), 405–412.

Hwang, I. M., Park, B., Dang, Y. M., Kim, S. Y., & Seo, H. Y. (2019). Simultaneous direct determination of 15 glucosinolates in eight Brassica species by UHPLC-Q-Orbitrap-MS. *Food Chemistry*, *282*(July 2018), 127–133.

Jiménez, J., Riverón-Negrete, L., Abdullaev, F., Espinosa-Aguirre, J., & Rodríguez-Arnaiz, R. (2008). Cytotoxicity of the β-carboline alkaloids harmine and harmaline in human cell assays in vitro. *Experimental and Toxicologic Pathology*, *60*(4–5), 381–389.

Le, H. T. N., Van Roy, E., Dendooven, E., Peeters, L., Theunis, M., Foubert, K., Pieters, L., & Tuenter, E. (2021). Alkaloids from *Lepidium meyenii* (Maca), structural revision of macaridine and UPLC-MS/MS feature-based molecular networking. *Phytochemistry*, *190*, 112863.

Liu, J. H., Zhang, R. R., Peng, X. R., Ding, Z. T., & Qiu, M. H. (2021). Lepipyrrolins A-B, two new dimeric pyrrole 2-carbaldehyde alkaloids from the tubers of *Lepidium meyenii*. *Bioorganic Chemistry*, *112* (December 2020), 104834.

Muhammad, I., Zhao, J., Dunbar, D. C., & Khan, I. A. (2002). Constituents of *Lepidium meyenii* 'maca.' *Phytochemistry*, *59*(1), 105–110.

Tafuri, S., Cocchia, N., Vassetti, A., Carotenuto, D., Esposito, L., Maruccio, L., Avallone, L., & Ciani, F. (2021). *Lepidium meyenii* (Maca) in male reproduction. *Natural Product Research*, *35*(22), 4550–4559.

Theunis, M., Naessens, T., Peeters, L., Brits, M., Foubert, K., & Pieters, L. (2022). Optimization and validation of analytical RP-HPLC methods for the

quantification of glucosinolates and isothiocyanates in *Nasturtium officinale* R. Br and *Brassica oleracea*. *LWT*, *165*, 113668.

Valerio Jr., L. G., Gonzales, G. F., Valerio, L. G., & F, G. G. (2005). Toxicological Aspects of the South American. *Toxicological Reviews*, *24*(I), 25.

Xia, C., Deng, J., Chen, J., Zhu, Y., Song, Y., Zhang, Y., Li, H., & Lin, C. (2019). Simultaneous determination of macaenes and macamides in maca using an HPLC method and analysis using a chemometric method (HCA) to distinguish maca origin. *Revista Brasileira de Farmacognosia*, *29*(6), 702–709.

Xie, Z., Cao, N., & Wang, C. (2021). A review on β-carboline alkaloids and their distribution in foodstuffs: A class of potential functional components or not? *Food Chemistry*, *348*(January), 129067.

Zhou, M., Zhang, R. Q., Chen, Y. J., Liao, L. M., Sun, Y. Q., Ma, Z. H., Yang, Q. F., Li, P., Ye, Y. Q., & Hu, Q. F. (2018). Three new pyrrole alkaloids from the roots of *Lepidium meyenii*. *Phytochemistry Letters*, *23* (October 2017), 137–140.

CHAPTER 9

General discussion, conclusion and future perspectives

This thesis focused on two different aspects of alkaloid research; the first one dealing with alkaloid-containing plant species from the Amaryllidaceae family as potential sources of antivirally active compounds against SARS-CoV-2; and the second one dealing with quality control of alkaloid-containing food supplements, more in particular preparations based on Maca (*Lepidium meyenii*).

At first, a general structure elucidation workflow was established, combining the power of computer-assisted structure elucidation by DFT calculations with various spectroscopic techniques (**chapter 3**). The workflow was successfully applied to determine 2D and 3D structures of all single compounds in this project, the epimeric mixture of 6-hydroxyhippeastidine and the racemic mixtures of (±)-2-*epi*-pseudolycorine and (±)-9-*O*-demethyl-11-hydroxygalanthine. The integration of computational methods into the conventional structure elucidation workflow allows an objective evaluation of all structural information of a compound, which is highly recommended as a routine practice to avoid the occurrence of misassignment, especially in case of compounds possessing unknown and unprecedented scaffolds.

The phytochemical investigations of the three selected plant species from the Amaryllidaceae family (*Hymenocallis littoralis*, *Pancratium maritimum*, *Scadoxus multiflorus*) by means of extensive chromatographic, spectroscopic and computational techniques led to the discovery of six novel alkaloids (*O*-demethyl-norlycoramine, 2β,10bα-dihydroxy-9-*O*-demethylhomolycorine, (±)-2-*epi*-pseudolycorine, and (±)-9-*O*-demethyl-11-hydroxygalanthine). Our studies also reveals new insights into the phytochemical profiles of the plant species:

- Three compounds: 2-*epi*-lycorine, ungeremine and zephyranthine were reported from *H. littoralis* for the first time (**chapter 4**).

- Three compounds (1-*O*-acetyl-norpluviine, 1-*O*-acetyl-10-*O*-methylpseudolycorine and 4'-*O*-methylnorbelladine) were isolated from *P. maritimum* for the first time (**chapter 5**).

- Nine alkaloids 2-*O*-methylpseudolycorine, ungiminorine, ungiminorine *N*-oxide, narcissidine, *O*-demethylmaritidine, 8-*O*-demethyloxomaritidine, sanguinine, 9-de-*O*-methyl-11β-hydroxygalanthamine and 2-hydroxy-*O,N*-dimethylnorbelladine were reported for the first time from *S. multiflorus* (**chapter 6**).

In total, a library of 37 Amaryllidaceae alkaloids was built from three plant materials (*H. littoralis*, *P. maritimum*, *S. multiflorus*), belonging to 7 skeletal subclasses: 14 lycorine-type, 3 lycorenine-type, 3 narciclasine-type, 8 crinine/ haemanthamine-type, 6 galanthamine-type, 1 montanine-type and 2 norbelladine-type compounds. The library was screened *in vitro* against the emerging SARS-CoV-2 virus and cytotoxicity on Vero-E6 cell line. As a result, thirteen compounds displayed moderate inhibition of SARS-CoV-2 replication (EC_{50} = 39 – 100 µM); ten compounds were devoid of activity and cytotoxicity; and twelve compounds lacked selective activity, since they were cytotoxic (**chapters 4 – 6**). Preliminary evaluation of structure-activity relationships revealed four skeleton-types (lycorine, homolycorine, crinine, and galanthamine) as promising scaffolds for further research. Future research on a larger library of compounds and chemical modification via semi-synthesis may lead to more active compounds.

In addition, as a future perspective, it is interesting to screen our library of compounds against other viruses and strains to further discover the antiviral potency of Amaryllidaceae alkaloids. As a promising result, screening against dengue virus type 2 revealed potent activities in the nanomolar range for some compounds. *In silico* study and functional bioassays are currently being performed to identify possible biological targets.

The phytochemical investigation of *Lepidium meyenii* (Maca) revealed three new compounds (lepidiline E, lepidiline F and lepidiline G) (**chapter 7**). The existence of amidine alkaloids was also described for the first time from Maca (N,N'-dibenzylacetamidine and N,N'-dibenzylformamidine). Furthermore, metabolomics

study by feature-based molecular networking indicated a cluster of quaternary imidazole alkaloids. By MRM analyses to confirm key fragments, other new compounds were tentatively identified. Besides, the structure of a mistakenly assigned compound named “macaridine” was revised to macapyrrolin C by NMR analysis and DFT calculation. The structure revision of “macaridine” to macapyrrolin C is crucial as macapyrrolin C is a major alkaloid in Maca, as evidenced by our analytical work described in **chapter 8**, and is thus a main target compound for quality control.

Using standards obtained from the isolation work reported in **chapter 7** (lepidilines A – E), synthesized standards (macapyrrolins A, C, G) and commercially available standards (BMI, DMI, MTCA, TCCA, ETCA), a new UPLC-TQD-MS/MS method was developed for quality control of alkaloids in Maca-containing food supplements (**chapter 8**). This represents the first analytical method for this purpose. Six alkaloids were found as major compounds, including lepidilines A – E and macapyrrolin C. Results obtained from analyses of commercial products indicate the total alkaloid content of 500 – 600 ppm in pure Maca powder. From a regulatory point of view, and based on safety concerns about the alkaloidal constituents of *L. meyenii*, it may be advisable to establish a maximum level of particular alkaloids or alkaloid classes in food supplements containing Maca, rather than to request their total absence. In addition, although the content of β-carbolines was found to be low in Maca (20-30 ppm), attention should also be drawn since these compounds were proven to exhibit many pharmacological and toxicological effects. In the future, analyses of more commercial products and raw materials are expected.

Last but not least, we recommend the use of lepidilines as chemical markers for characterization of Maca in all types of commercial preparation due to their high content, exclusive existence in Maca, and ultra-sensitivity to MS detection. This allows a rapid and sensitive verification of Maca adulteration in food supplements.

SUMMARY

This PhD project focuses on the phytochemical, antiviral and analytical investigation of selected classes of alkaloids. The thesis contains two distinct parts; the first one dealing with Amaryllidaceae alkaloids and potential antiviral agents, more in particular against SARS-CoV-2; and the second one dealing with quality control of alkaloid-containing food supplements based on Maca (*Lepidium meyenii*).

A general structure elucidation workflow was established, combining the power of computer-assisted structure elucidation (CASE) by DFT (density functional theory) calculations with various chromatographic and spectroscopic techniques. The workflow was successfully applied to determine 2D and 3D structures of all compounds in this study. In addition, the structure of a mistakenly assigned compound named “macaridine” was revised to macapyrrolin C by DFT calculation of chemical shifts.

The phytochemical investigations of the four plant species (*Hymenocallis littoralis*, *Pancratium maritimum*, *Scadoxus multiflorus*, *Lepidium meyenii*) led to the discovery of nine novel alkaloids as well as extended the phytochemical profiles of the plant species. A library of Amaryllidaceae alkaloids was built and screened *in vitro* against the emerging SARS-CoV-2 and cytotoxicity on Vero-E6 cell line. As a result, thirteen compounds displayed moderate inhibition of SARS-CoV-2 replication (EC_{50} = 39 – 100 µM); ten compounds were devoid of activity and cytotoxicity, and twelve compounds lacked selective activity, since they were cytotoxic. Preliminary evaluation of structure-activity relationships revealed four skeleton-types (lycorine, homolycorine, crinine, and galanthamine) as promising scaffolds for further research.

A new UPLC-TQD-MS/MS method was developed and validated for quality control of alkaloids in Maca-containing food supplements, because of their suspected toxicity. This represents the first analytical method for this purpose. Results obtained from analyses of commercial products indicate the alkaloid content of

500 – 600 ppm in pure Maca powder. From a regulatory point of view, and based on safety concerns about the alkaloidal constituents of *L. meyenii*, it may be advisable to establish a maximum level of particular alkaloids or alkaloid classes in food supplements containing Maca, rather than to request their total absence. Also, we recommend the use of lepidilines as chemical markers for characterization of Maca in all types of commercial preparation due to their high content, exclusive existence in Maca and ultra-sensitivity to MS detection.

SAMENVATTING

Dit doctoraatsproject omvat het fytochemisch, antiviraal en analytisch onderzoek van enkele geselecteerde klassen van alkaloïden. De thesis bestaat uit twee grote delen; het eerste handelt over Amaryllidaceae alkaloïden als mogelijke antivirale middelen, meer bepaald tegen SARS-CoV-2; en het tweede over kwaliteitscontrole van alkaloïden-bevattende voedinssupplementen op basis van Maca (*Lepidium meyenii*).

Een algemene strategie voor structuuropheldering werd opgesteld, waarin de kracht van computer-geassisteerde structuuropheldering (computer-assisted structure elucidation, CASE) met density functional theory (DFT) berekeningen gecombineerd wordt met chromatografische en spectroscopische technieken. Deze benadering werd succesvol toegepast om de twee- en drie-dimensionale structuren te bepalen van alle in dit werk verkregen producten. Bovendien werd de foutieve structuur van een product genoemd “macaridine” herzien tot macapyrroline C op basis van DFT berekeningen van de chemische verschuivingen in NMR.

Het fytochemisch onderzoek van vier plantensoorten, meer bepaald *Hymenocallis littoralis*, *Pancratium maritimum* en *Scadoxus multiflorus* behorend tot de Amaryllidaceae, en *Lepidium meyenii* leidde tot de ontdekking van 9 nieuwe alkaloïden, en een beter inzicht van het fytochemisch profiel van deze soorten. Een bibliotheek van Amaryllidaceae alkaloïden werd opgebouwd, en *in vitro* getest voor activiteit tegen SARS-CoV-2 en voor cytotoxiciteit op de Vero-E6 cellijn. Hieruit bleek dat 13 producten een matige inhibitie vertoonden van de SARS-CoV-2 replicatie (EC_{50} = 39 – 100 µM); 10 producten waren niet actief of cytotoxisch; en 12 producten vertoonden geen selectieve activiteit omwille van cytotoxiciteit. Preliminair onderzoek van de structuur-activiteitsrelatie toonde aan dat 4 skelet-types, namelijk het lycorine, homolycorine, crinine en galanthamine skelet, als veelbelovend voor verder onderzoek kunnen beschouwd worden.

Een nieuwe UPLC-TQD-MS/MS methode werd ontwikkeld en gevalideerd voor de kwaliteitscontrole van alkaloïden in Maca-bevattende voedingssupplementen, omwille van hun vermeende toxiciteit. Dit is de eerste analytische methode voor dit doel. De resultaten van de analyse van een aantal commerciële producten toonden aan dat het alkaloïden gehalte in Maca poeder 500 – 600 ppm bedroeg. Op regulatoir gebied, en omwille van veiligheidsoverwegingen, lijkt het wenselijk om eerder een maximumgehalte aan alkaloïden op te leggen, dan hun volledige afwezigheid te vereisen. Ook wordt het gebruik van lepidilines als chemische merkers voor de karakterisatie van Maca in commerciële stalen aanbevolen, omwille van hun relatief hoog gehalte, exclusief voorkomen in *Lepidium meyenii*, en de zeer hoge gevoeligheid van massa-spectrometrische detectie.

ACKNOWLEDGEMENT

Since my master stage, I have spent 4.5 years in the NatuRA research lab. There are numerous unforgettable memories that I will remember for the rest of my life. Looking back the past few years, I got the bitter-sweet feeling writing these words. I am grateful to all of those with whom I have had the pleasure to meet, talk and work with. I would like to thank the following people, without whom I would not have been able to complete this journey:

My supervisors Prof. Luc Pieters and Prof. Emmy Tuenter, for your consistent support, guidance, encouragement and patience. At the end, what we have done is simply for the sake of great science. Among all things, it is my blessing learning NMR expertise from you two – and then other good things come to me following this great stepping stone. I am deeply grateful to Prof. Luc Pieters for giving me the opportunity to conduct this PhD thesis in the lab, for always supporting me learning new knowledge, for always being available and approachable, and for always giving me the freedom to explore the PhD project myself. With that all being said, sometimes I wish I could have been a better student and could have made you feel more proud of me as your PhD student.

Dr. Kenn Foubert, for your support with the analytical method development and validation of course, but most importantly, for your bloodily smart sense of humor.

Mart and Tania, for your help with many things, from big to small, always with a warm welcome. Sometimes the warmth I felt made me behave childishly.

Prof. Nina Hermans and Dr. Tess De Bruyne. I did not have chances to work with you, but you are always nice to me and make me feel welcomed in the lab. Thank you for that!

Labmates and Friends: Stéfi, Andrés, Baldé, Laura, Anastasia, Maxime, Yunita, Stef, Anne-Sophie, Maxim, Sophie, Laurens, Tom, Roy, Murat. It's my pleasure to meet you all in my life. Our talks, gossips and discussions are always an important part of this great journey.

The out-of-this-world collaborators: Prof. Wouter A. Herrebout, Prof. Johannes Kirchmair, Prof. Pieter van Der Veken, Prof. Christophe Pannecouque, Prof. Johan Neyst, Prof. Suzanne Kaptein.

MICA: Prof. Christine Beemelmanns, Sven, Martinus, Marie, Rebecca, Tanya, Min. Thank you for hosting me in your lab! The research stay was a breath of fresh air that I'm fortunate enough to have during this PhD journey. I greatly appreciate your hospitality, professionalism and expertise.

Dr. Julian Hegemann. I really enjoyed our discussions. You helped me find myself in times of uncertainty and self-doubting. The book "Surely You're Joking, Mr. Feynman!" is definitely the best gift I received. Thank you!

Prof. Boutayna Rhourri-Frih, for your trust and support. Without you, my journey in Europe would have not started.

My former Vietnamese advisors and teachers: Cô Võ Thị Bạch Huệ, thầy Phan Văn Hồ Nam, cô Lan, thầy Nghĩa, cô Mai, cô Thư, cô Liên, thầy Lợi, cô Bãng Thu.

My Vietnamese friends: chị Trinh, chị Giang, chị Hoà, anh Vinh, anh Tín, anh Thư, anh Nhàn, anh Huỳnh, anh Thành, anh Dũng, chị Lý, chị Trúc, chị Trúc Linh, Như Ý, Lê Uyên, Hà Phương, Diệu Linh, Huỳnh Kim, Quỳnh Anh, Duy, Hiển, Nhân, Hiệp, Lam Giang, em Trang, em Huyền, em Lượng, em Quân, em Tuân, em Quốc.

The special one. Thank you for coming to my life breaking my every single rule. That's how my life has become colorful.

From the bottom of my heart, nobody has been more important to me in the pursuit of this journey than my beloved family. I am forever indebted to my parents, my brother and my sister, whose love are with me in whatever I pursue.

Long story short, memories of old times came flooding back one after another and I'm blessed. This great chapter of my life is now closed and I cannot be more excited to write the next one! Thank you all again for being there!

SCIENTIFIC CURRICULUM VITAE

1. Personal information

Name: Lê Ngọc Thảo Hiền (Le Ngoc-Thao-Hien)

Date of birth: March 22nd, 1995

Place of birth: Ho Chi Minh City, Vietnam

E-mail: NgocThaoHien.Le@uantwerpen.be / lengocthaohien@gmail.com

2. Education

2020-present	**PhD candidate** Natural Products and Food - Research and Analysis (NatuRA) University of Antwerp, Antwerp, Belgium
2018-2019	**Master 2 in Analytical chemistry for Drugs and Natural products** University of Bordeaux, Bordeaux, France
2013-2018	**Degree of Pharmacist** University of Medicine and Pharmacy Ho Chi Minh City Ho Chi Minh City, Viet Nam
2010-2013	**High school for the gifted** Ho Chi Minh City, Viet Nam

3. Scientific experience

Dec 2019-present	PhD thesis - "Analytical, antiviral and chemical investigations of selected classes of alkaloids" at the research group "Natural Products and Food - Research and Analysis (NatuRA)", Antwerp University.
June – Aug 2023	Research stay at the "Anti-Infectives From Microbiota" research group of Prof. Christine Beemelmanns at Helmholtz

Institute for Pharmaceutical Research Saarland: "Establishment of computational pipeline for structure elucidation of natural products – application on fungal metabolites isolated from *Pseudoxylaria* sp. X187".

Mar – Aug 2019 Master thesis – "Isolation and structure elucidation of alkaloids from *Hymenocallis littoralis* (Jacq.) Salisb.'' at the research group "Natural Products and Food - Research and Analysis (NatuRA)", Antwerp University.

Feb – Jul 2018 Bachelor thesis - ''Quantitative determination of alkaloids in *Vernonia amygdalina* Del. Asteraceae leave'' at Department of Analytical chemistry and Drug quality control, University of Medicine and Pharmacy Ho Chi Minh City.

4. Publications as first author

2021 Le Ngoc-Thao-Hien, Elias Van Roy, Ella Dendooven, Laura Peeters, Mart Theunis, Kenn Foubert, Luc Pieters, Emmy Tuenter. "Alkaloids from *Lepidium meyenii* (Maca), structural revision of macaridine and UPLC-MS/MS feature-based molecular networking". Phytochemistry, Volume 190, 2021, 112863, ISSN 0031-9422. https://doi.org/10.1016/j.phytochem.2021.112863

2023 Le Ngoc-Thao-Hien, Tom Vermeyen, Roy Aerts, Wouter A. Herrebout, Luc Pieters, and Emmy Tuenter. "Epimeric Mixture Analysis and Absolute Configuration Determination Using an Integrated Spectroscopic and Computational Approach—A Case Study of Two Epimers of 6-Hydroxyhippeastidine". Molecules 28, 2023, no. 1: 214. https://doi.org/10.3390/molecules28010214

2023 Le Ngoc-Thao-Hien, Steven De Jonghe, Kristien Erven, Tom Vermeyen, Aliou M. Baldé, Wouter A. Herrebout, Johan Neyts, Christophe Pannecouque, Luc Pieters, and Emmy Tuenter. "Anti-SARS-CoV-2 Activity and Cytotoxicity of Amaryllidaceae Alkaloids from *Hymenocallis littoralis*". Molecules 28, 2023, no. 7: 3222. https://doi.org/10.3390/molecules28073222

2023 Le Ngoc-Thao-Hien, Steven De Jonghe, Kristien Erven, Tom Vermeyen, Wouter A. Herrebout, Johan Neyts, Christophe Pannecouque, Luc Pieters, and Emmy Tuenter. "A new alkaloid from *Pancratium maritimum* - Structure elucidation using computer-assisted structure elucidation (CASE) and evaluation of anti-SARS-CoV-2 activity". Phytochemistry letters, 2023, 58, 1-7. https://doi.org/10.1016/j.phytol.2023.09.006

2023 Le Ngoc-Thao-Hien, Steven De Jonghe, Kristien Erven, Tom Vermeyen, Wouter A. Herrebout, Johan Neyts, Christophe Pannecouque, Luc Pieters, and Emmy Tuenter. "Comprehensive study of alkaloids from *Scadoxus multiflorus* by HPLC-PDA-SPE-NMR and evaluation of their anti-SARS-CoV-2 activity". Phytochemistry Letters, 2023, 57, 156-162. https://doi.org/10.1016/j.phytol.2023.08.015

202x Le Ngoc-Thao-Hien, Kenn Foubert, Mart Theunis, Tania Naessens, Murat Bozdag, Pieter Van Der Veken, Luc Pieters, Emmy Tuenter. “Method validation and quality control of alkaloid content in *Lepidium meyenii* (Maca) containing food and food supplements” (manuscript under review)

5. Scientific communication

- **Oral presentation**: “Phytochemical and analytical investigations on *Lepidium meyenii* (Maca)” at the research day of Faculty of Pharmaceutical, Biomedical and Veterinary Sciences (October 27th 2023 University of Antwerp, Belgium)

- **Oral and poster presentations**: “Application of DFT calculation for structure elucidation of selected alkaloids” at the 71th International

Congress and Annual Meeting of the Society for Medicinal Plant and Natural Product Research (July 1-5th 2023, Dublin, Ireland)

- **Oral presentation**: "Antiviral activities of Amaryllidaceae alkaloids" at the research day of Department of Pharmaceutical Sciences (May 26th 2023 University of Antwerp, Belgium)

- **Poster presentation**: "Application of DFT calculation for structure elucidation of selected alkaloids" at Helmholtz Institute for Pharmaceutical Research Saarland (HIPS) Symposium 2023 (May 4th 2023, Saarbrücken, Germany)

- **Oral presentation**: "Antiviral activities of Amaryllidaceae alkaloids" at the 21st Forum of Pharmaceutical Sciences (April 27-28th 2023, Floreal-Blankenberge, Belgium)

- **Oral presentation**: "Simultaneous quantitation of 11 major alkaloids using UPLC-TQD-MS/MS for quality control of *Lepidium meyenii* (Maca)-containing food and food supplements" at the 12th International Symposium on Drug Analysis and 32nd International Symposium on Pharmaceutical and Biomedical Analysis (September 11-14th, Mons, Belgium)

- **Poster presentation**: "A new alkaloid from *Scadoxus multiflorus*" at the 70th International Congress and Annual Meeting of the Society for Medicinal Plant and Natural Product Research – Planta Medica, DOI: 10.1055/s-0042-1759107 (August 28-31st 2022, Thessaloniki, Greece)

- **Poster presentation**: "A new alkaloid from *Pancratium maritimum* L." at the 69th International Congress and Annual Meeting of the Society for Medicinal Plant and Natural Product Research – Planta Medica, DOI: 10.1055/s-0041-1736841 (September 5-9th 2021, online)