Integrated analytical workflow for chromatographic profiling and metabolite annotation of a cytotoxic Phorbas amaranthus extract

Bruno S. do Amaral a,b,1, Fernanda B. da Silva c,*,1, Gabriel Mazzi Leme a, Letícia S.S. Schmitz d, Paula C. Jimenez d, Roberto Carlos Campos Martins d, Quezia B. Cass a,b,*, Alessandra L. Valverde a,b,*

a Separare, Departamento de Química, Universidade Federal de São Carlos, Rodovia Washington Luís, km 235, 13565-905 São Carlos, SP, Brazil
b Instituto Federal de Educação, Ciência e Tecnologia de São Paulo, Campus Piratiba, 05110-000 São Paulo, SP, Brazil
c Instituto de Pesquisas de Produtos Naturais, Universidade Federal do Rio de Janeiro, CEP 21941-599 Rio de Janeiro, RJ, Brazil
d Departamento de Ciências do Mar, Instituto do Mar, Universidade Federal de São Paulo, CEP 11070-100 Santos, SP, Brazil
e Laboratório de Produtos Naturais, Instituto de Química, Universidade Federal Fluminense, Niterói, RJ CEP 24020-141, Brazil

ARTICLE INFO

Keywords:
Bioactivity
Multi-column screening
Chromatographic modeling and simulation
DryLab®
Marine natural products

ABSTRACT

Phorbas is a widely studied genus of marine sponge and produce structurally rich cytotoxic metabolites. Still, only few studies have assessed metabolites present in Brazilian species. To circumvent redundancy, in this work, we applied and herein report the use of a scouting liquid chromatographic system associate to the design of experiment produced by the DryLab® software to obtain a fast and efficient chromatographic separation of the active hexane fraction, further enabling untargeted high-resolution mass spectrometry (HRMS) data. To this end, a crude hydroalcoholic extract of the sponge Phorbas amaranthus collected in Brazilian coast was prepared and partitioned. The cytotoxicity of the crude extract and the fractions was evaluated using tumor cell culture. The cytotoxicity of the crude extract and the fractions was evaluated using tumor cell culture models. Fragmentation pathways assembled from HRMS data allowed the annotation of 18 known Phorbas metabolites, while 17 metabolites were inferred based on Global Natural Product Social Molecular Networking (GNPS), matching with a further 29 metabolites annotated through molecular subnetwork. The workflow employed demonstrates that chromatographic method development can be accelerated by the use of automated scouting systems and DryLab®, which is useful for profiling natural product libraries, as well as data curation by molecular clusters and should be incorporated to the tools of natural product chemists.

1. Introduction

Marine sponges produce numerous bioactive metabolites and Phorbas amaranthus (class Demospongiae, Order Poecilosclerida, family Anchnidoidea), in particular, is known to be a rich source of structurally diverse and biologically active secondary metabolites [1,2]. Among the substances described for Phorbas amaranthus there are the anthosterones A and B and phorbasterones A–D (oxidized steroids with contracted A-rings) [1], amaranzoles A–F (steroidal alkaloids) [3,4], and amaroxocanes A and B (dimeric steroids bridged by the side-chain), which are considered the main active principles for the antifeedant activity [5]. Other substances have also been identified for the genus Phorbas, such as alotaketals C–E [6,7], anchiopeptolides A–E [8,9], ansellone A [10], B [6] and D–G [7], avilone A and B [7], cycloanchinopeptolide C [9], gagunins A–Q [11,12], gulkanolins A–F [13,14], hemi-phorboxazole A [15], isophorbasone A [16], muironolide A [17], phorbadione [6], phorbaketals A–C [18] and L–N [19], phorbasides A–E [20], F [21] and G–I [22], phorbasin A–K [23–27], phorbasones A and B [28], phorbasone A acetate [16], phorbatopsins A–C [29], phorbazoles A–D [30], phorbazoles A and B [31,32], phorone A [16], and secoepoxysellolone A [6].

Due to the recognized cytotoxic activity against tumor cells in extracts of these sponges [1,2,11,13] and the diversity of metabolites already known, P. amaranthus, collected in the Brazilian sea – an area which remains scarcely studied considering its dimension, seemed an excellent library for profiling metabolites. In this regard, liquid

* Corresponding authors at: Separare, Departamento de Química, Universidade Federal de São Carlos, Rodovia Washington Luís, km 235, 13565-905 São Carlos, SP, Brazil (Q.B. Cass) and Laboratório de Produtos Naturais, Instituto de Química, Universidade Federal Fluminense, 24030-141, Niterói, RJ, Brazil (A.L. Valverde).
1 These authors contributed equally to this work.

https://doi.org/10.1016/j.jchromb.2021.122720

Received 31 January 2021; Received in revised form 9 April 2021; Accepted 12 April 2021
Available online 21 April 2021
1570-0232/© 2021 Elsevier B.V. All rights reserved.
chromatography hyphenated with high resolution mass spectrometry (LC-HRMS) turns out as an important tool for avoiding re-isolation of known compounds.

In order to characterize the constituents of a sample in an untargeted profile, the goal is to obtain chromatograms that can provide the highest number of peaks. Due to the metabolites’ diversity in hydrophobicity, gradient elution is usually carried out in the reverse-phase mode (RPLC). In this scenario, selection of chromatographic conditions are usually achieved by trial-and-error approaches, which, therefore, underscores the need for a systematic workflow. To meet this end, selection of orthogonal columns will provide significant differences in selectivity and retention [33], especially when associated to other chromatographic parameters such as organic modifier, strength range, gradient elution time, pH and buffer concentration. To minimize costs in method development, a column and eluent scouting LC systems is useful [34]. The advantages of these systems include automation, software to assist in method setup and data analysis, stable solvent delivery, injection reproducibility and low carryover. The type and assembly of the valve system, as well as the number of scouting columns and mobile phases varies according to the manufacturers [35,36]. To foster productivity and meet green chemistry principles, the use of DryLab® software in association with scouting systems is an important asset. Based on its Design of Experiments (DoE), the software produces multi-dimensional resolution maps that gives the optimal experimental conditions for a chromatographic separation [37,38], thus leading to reductions in solvent use and analysis time, especially when compared to results obtained by conventional trial-and-error procedures [38].

Although DryLab® has been cited since the early 1990s in the field of natural products, such as for the analysis of eight flavonoids in some species of the genus Althaea [39], it is underused for method development in this area. Over the past five years, the use of the software has been reported for a number of RPLC methods related to analysis of extracts of Cannabis sativa [40], Piper guineense [41], Corchorus olitorius and Vitis vinifera [42], as well as for five species of laxative herbs [43] and for analysis by hydrophilic interaction liquid chromatography (HILIC) of hallucinogenic mushroom extracts [44]. In these cases, however, DryLab® was used for assisting in the resolution of critical pairs and not for optimization of a chromatographic profile.

Herein, we report the use of a scouting system and DryLab® for selecting the DoE parameters and predicting the optimized chromatographic conditions, respectively. For metabolic profiling, a hexane fraction obtained by partition of the hydroalcoholic extract of P. amaranthus was analyzed. This fraction was selected since it showed high cytotoxicity against in vitro tumor cells models.

To access the metabolite diversity, structural database was used to provide data curation. Comparison of the acquired spectra with reference spectra by means of organized MS/MS clusters in the Global Natural Product Social Molecular Networking (GNPS) [45] provided annotation to unknown molecules. The use of fragmentation patterns to dereplicate some known classes of marine metabolites were also used as analytical tools, affording 64 metabolite annotation for the Brazilian marine metabolites were also used as a workflow for prospecting metabolites is fully discussed herein.

2. Material and methods

2.1. Chemicals

All reagents and solvents obtained from commercial suppliers were used without further purification. Methanol (MeOH) and acetonitrile (ACN) HPLC grade were purchased from J.T. Baker (Phillipsburg, USA). Hexane HPLC grade, tetradhydrofuran (THF) HPLC grade and 3-(4,5-Dimethyl-2-thiazolyl)-2,5-diphenyl-2H-tetrazolium bromide (MTT) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Formic acid (p.a. grade) was acquired from Fluka (Buchs, Switzerland). Roswell Park Memorial Institute (RPMI) 1640 Medium, fetal bovine serum (FBS) and penicillin/streptomycin antibiotic solution were purchased from Gibco® by Life Technologies (Thermo Fisher Scientific, Waltham, MA, USA). Water was purified in a Milli-Q system (Millipore, Sao Paulo, Brazil).

2.2. Collection and extraction

The marine sponge Phorbas amaranthus Duchassaing & Michelotti was collected on March 3rd, 1998, at a 15 m depth at Ilha do Frade (Fernando de Noronha/PE, Brazil), then frozen and kept at ~20 °C until use. The project is registered in the SISGEN platform under the number AB724BB.

For the extraction, an aliquot of the collected sponge was cut and weighed. The material was left for 6 h in a desiccator containing silica gel to remove part of the water from the sponge then the material was weighed again. Afterwards, a static maceration in ethanol was carried out for seven days, followed by crushing in a blender with ethanol, and filtration. Then, two static macerations were performed in methanol/ethyl acetate 1:1 (v/v) for seven days each. The three organic fractions were pooled, solvent was eliminated in a rotary evaporator and traces of water removed in Savant SpeedVac® (Thermo Scientific) to obtain the crude dry extract of P. amaranthus (PaCE). For micropartition, 1 mg of PaCE was solubilized in a mixture of 4 mL of MeOH + 3 mL of H2O in a sonicator. Then, 3 mL of hexane was added and sonicated again until maximum solubility. The solution was transferred to a threaded test tube, added with 4 mL of hexane and vortexed for 20 s. The samples were centrifuged for 5 min (~250 rpm). The hexane phase was removed, and the extraction was repeated with 7 mL of hexane. The fractions were combined, solvent was eliminated by rotary evaporator and dried in Savant SpeedVac®. Finally, 3 mL of H2O were added to the aqueous phase to repeat the extraction process with butanol (2 × 3 mL), thus yielding the butanolic fraction (PaB) and the aqueous fraction (PaA).

2.3. Evaluation of cytotoxic activity in vitro

2.3.1. Cell lines

The human colon carcinoma cell line HCT 116 (ATCC CCL 247) was used as cell model to detect cytotoxic activity. The cells were grown in flasks with RPMI 1640 culture medium added with 10% FBS and 1% antibiotics (penicillin/streptomycin) at 37 °C in a humidified atmosphere with 5% CO2. When necessary, cultures were split in a 1:6 ratio to keep them growing exponentially. For the tests, 1 × 104 cells in 0.2 mL of culture medium were seeded in 96-well plates and allowed to adhere for 24 h.

2.3.2. In vitro cytotoxicity assay

For the quantitative evaluation of cytotoxicity, HCT 116 cells were exposed for 72 h to seven concentrations ranging from 0.003 to 50 μg/mL of the PaCE. To compare the samples, the cells were exposed for 72 h to two concentrations – 5 and 50 μg mL−1 of PaCE and fractions thereof (PaB, PaA, PaB, PaA). All samples were previously dissolved in DMSO. This dilution vehicle and doxorubicin were used as negative and positive controls, respectively. After the exposure time, the cytotoxicity of each sample was measured by the MTT assay [46]. Briefly, the culture medium was aspirated from each well and replaced with fresh medium containing 0.5 mg mL−1 MTT, then incubated for 3 h. At the end, all the media was again removed from each well and the precipitates formed were resuspended in DMSO to attain absorbance readings in a plate spectrophotometer at 570 nm. All experiments were carried out in triplicate. The percentage of cell growth inhibition induced by each treatment was calculated in relation to the absorbance of the negative control. Experiments carried out with PaCE in a concentration range had the IC50 values (half-maximum inhibitory
concentration), CI 95% (95% confidence interval) and R² estimated using GraphPad Prism 8.0 software by nonlinear regression analysis.

### 2.4. LC-DAD method development for the PaHex

The scouting UHPLC system used (Nexera®, Shimadzu) consisted of two quaternary LC-30AD pumps, a DGU-20A degasser, a SIL-30A autosampler, a CTO-20AC column oven, a six-column selector valve, diode array detector SPD-M30A, and CBM 20A interface. The Method Scouting Solution® system was used to develop the analysis sequence and LabSolutions® workstation software to control all modules and for data processing.

The PaHex was resuspended in THF to 2.5 mg mL⁻¹ and analyzed (10 μL) by LC-DAD using the Method Scouting Solution® with five different columns (100 × 2.1 mm): Ascentis® Express C₁₈ (2.7 μm, Supelco), Ascentis® Express F5 (2.7 μm, Supelco), Acquity UPLC BEH Phenyl (1.7 μm, Waters), Ascentis® Express Phenyl-Hexyl (2.7 μm, Supelco) and XSelect® HSS T3 (3.5 μm, Waters). The THF was also injected (10 μL) to provide background chromatograms.

The separations were carried out using water as mobile phase A and two different organic modifiers as mobile phase B (ACN or MeOH), both containing 0.1% formic acid, in a flow rate of 0.4 mL min⁻¹. The total run time was 30 min, in a linear gradient elution mode under the following conditions: 0 min, 5% B; 0–20 min, 5–100% B; 20–25 min, 100% B and a reconditioning cycle of 5 min with the same initial conditions of 5% B. All columns were maintained at 40 °C. DAD absorption was acquired in a range of 190–700 nm. The combination of stationary and mobile phases led to 10 different experiments (20 runs) which were programmed in a single batch using Method Scouting Solution®.

The number of chromatographic peaks observed within 25 min of analysis and detected at 220 and 254 nm were used as a parameter for column and organic modifier selection for subsequent optimization. This number was obtained through the automatic integration performed by the software using the same criteria for all chromatograms (slope, 10,000 μV min⁻¹; minimum width at half-height, 5 sec; minimum peak area, 1000 counts). Prior to integration, the background chromatograms were subtracted from the sample chromatograms.

The Ascentis® Express Phenyl-Hexyl (100 × 2.1 mm, 2.7 μm) column was selected for further optimization of the method. For that, LC-DAD analysis was carried out using water (A) and ACN (B), both containing 0.1% formic acid as mobile phase, with gradient times (t₀) of 20 and 60 min to 5–100% B, with the same 100% B (3 min) and reconditioning (5 min) cycles of the column, in a total run time of 28 and 68 min, respectively. The chromatograms were exported to DryLab® 4 Software for Chromatography Modelling (Version 4.2.1.9, Molnár-Institute, Berlin, Germany) for optimization of the gradient range (∆B) and t₀.

### 2.5. LC-HRMS analysis of the PaHex

The UHPLC system (Agilent 1290 Infinity II, Agilent Technologies, Santa Clara, CA, USA) used consisted of a binary pump (G7120A – High speed Pump) and an autosampler and column oven (G7129B – 1290 Vial sampler). The UHPLC system was coupled to a high-resolution mass spectrometer (HRMS) containing a quadrupole time-of-flight mass analyzer (QqTOF). The HRMS analysis were performed using an Impact HD QTOF mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with an electrospray ionization source (ESI). Data acquisition were carried out using the Data Analysis 4.0 software (Bruker Daltonics GmbH, Bremen, Germany).

For the LC-HRMS analysis, the conditions were Ascentis® Express Phenyl-Hexyl (100 × 2.1 mm, 2.7 μm) column, in a gradient elution using water (A) and ACN (B), both containing 0.1% formic acid as mobile phase at a flow rate 0.4 mL min⁻¹, at 40 °C. The total run time was 18 min using the following linear gradient: 0 min, 45% B; 0–10 min, 45–100% B; 10.1–13 min, 100% B and a conditioning cycle time of 5 min with the same initial conditions of 45% B. The ESI source operated in positive and negative ion mode and the QqTOF parameters were set as follows: capillary voltage, 4500 V; end plate offset, 500 V; nebulizer (N₂), 1 bar; dry heater temperature, 200 °C; dry gas flow, 8 L min⁻¹; collision cell energy, 8 eV; transfer time, 80 μs; pre-pulse, 3.0 μs. The MS was programmed to perform acquisition in auto MS/MS mode (number of precursors 3), with absolute threshold of 360 counts; spectra rate, 8 Hz; total time cycle range, 0.5 s; in experiments with different collision energy of 18–45 eV for all m/z range analyzed.

External mass spectrometer calibration was performed with sodium formate clusters (1 mmol L⁻¹ sodium formate in water/acetonitrile 1:9 (v/v)) in quadratic high-precision calibration (HPC) regression mode. The calibration solution (3 μL) was injected at the end of each analytical run and all the spectra were calibrated prior to secondary metabolites’ identification.

To identify the substances present in PaHex, the m/z observed were compared with the theoretical m/z calculated to the 98 substances found in the literature reported for P. amaranthus and the genus Phorbas [1–32]. For that, the acceptable error limit was ± 5 ppm and the substances were correlated with their fragment ions. Additionally, the acquired data were converted to mzXML format using Data Analysis 4.0 software and submitted to Global Natural Products Social Molecular Network (GNPS [45]; http://gnps.ucsd.edu) online system. The molecular network calculations and database matching were constructed using 0.02 Da as precursor ion mass tolerance and 0.05 Da as fragment ion mass tolerance; 0.7 as minimum cosine score and 3 as minimum matched fragment ions for edge linkage. Finally, GNPS data was then imported and visualized using the Cystoscope software (version 3.8.0) to find the subnetworks portions.

### 3. Results and discussion

#### 3.1. Evaluation of the cytotoxicity of the crude extract and fractions

To comprehensively explore the compounds in P. amaranthus, methanol with ethyl acetate were used as solvents to prepare the crude extract (PaC₅). Next, a liquid–liquid partition was carried out to fractionate substances with similar physical–chemical characteristics and to remove interferents, such as primary metabolites. Then, the potential cytotoxicity of the PaC₅ was evaluated against HCT 116 cells in culture through a quantitative approach that showed an IC₅₀ value of 17.69 μg mL⁻¹ (Fig. 1A).

In order to compare the bioactivity of PaC₅ with the different partitions (PaC₁₂, PaC₁₆, PaB₈ and PaA₇), the inhibitory capacity on cell growth was accessed for the samples through a qualitative approach (Fig. 1B). The aqueous fraction (PaC₁₂) showed the lowest cytotoxicity against HCT 116 cells, inhibiting nearly 30% of cell growth at 50 μg mL⁻¹, suggesting polar compounds may not strongly contribute to the bioactivity observed in the crude extract. The acetate fraction (PaA₇), compared to hexane (PaC₁₂) and butanol (PaB₈) fractions, displayed a lower cytotoxicity at 50 μg mL⁻¹, however, at 5 μg mL⁻¹, differently from the other partitions, it still revealed a residual bioactivity, inhibiting approximately 20% of cell growth, suggesting the presence of low amounts of cytotoxic or even the occurrence of cytostatic compounds. Indeed, PaHex at the highest concentration tested, followed closely by PaB₈, showed the strongest cytotoxicity towards HCT 116, revealing an inhibitory potential similar to that of the crude extract (PaC₅). Thus, PaHex was selected for the dereplication step to assess the metabolites putatively responsible for such bioactivity.

#### 3.2. Development of the chromatographic conditions for the PaHex extract

To select the appropriate column and organic modifier for reversed-phase elution mode, an automated column and eluent scouting protocol was carried out. Five orthogonal columns were used [47] selected based on Snyder’s hydrophobic-subtraction model, in accordance to their Fs at
pH 2.8 [48]. For the organic modifier, either methanol or acetonitrile containing 0.1% formic acid were evaluated in an aqueous linear gradient of 5–100% of 30 min. Thus, 10 different chromatograms were obtained (Fig. S1 – Fig. S10).

In all cases, ACN showed better selectivity than MeOH as an organic modifier. All the fused-core columns of the Ascentis® family (2.7 μm) presented better separations than the sub-2 μm column Acquity® UPLC BEH Phenyl. Between the two porous columns, the C18 phase of XSelect® HSS T3 (3.5 μm) exhibited higher efficiency than the column Acquity® UPLC BEH Phenyl, which may be related to its greater loading capacity. The highest number of peaks was obtained with Ascentis® Express Phenyl-Hexyl column. This phase, overall, exhibits a balanced aromatic and hydrophobic interaction that relates to the hydrophobic components of the sample with conjugated double-bond system. Although ACN tends to decrease the π−π interactions between phenyl stationary phases and aromatic analytes, these compounds were not identified in PHex (Section 3.3). The phase also contained the hexyl group which can give increased retention and changes in selectivity.

To optimize tG and ΔB of the gradient elution by the DryLab® software, two extra experiments were carried out using the selected column and organic modifier. For that, gradient elution of 5–100% ACN in 20 min (Fig. 2A) and 60 min (Fig. 2B) were run. The chromatograms registered at 254 nm were exported to DryLab® (Fig. S11) and, after correlating the chromatograms peaks of the two gradients, the ΔB was adjusted to 45–100% and tG to 10 min (Fig. 2C). The in silico optimized conditions were then run and the registered chromatogram (Fig. 2D) revealed a remarkable similarity with that predicted by the software.

One of the greatest DryLab® software limitations is the need for peak tracking in different chromatographic conditions using only the peak areas, since the software does not use any qualitative data. The correlation can be tedious and challenging depending on the complexity of the sample. Nevertheless, the association of scouting for selecting the column and the organic modifier, followed by DryLab® optimization of the gradient conditions, produced a chromatographic profile with selectivity and short analysis time for annotation of the metabolites. To this end, 13 experiments were carried out (26 chromatographic runs), which represented 13.8 h of total analysis time and 331.2 mL of mobile phase.

The optimized chromatographic conditions of LC-DAD were transfer to LC-HRMS system without affecting the gradient elution and analyses were applied in both positive- and negative-ion modes.

3.3. Dereplication of the PHex extract

A total of 64 compounds, including diterpenes, sesterterpenes, sterols and lysosphospholipids were identified. The molecular formulas were accurately assigned within mass errors of 5 ppm. The fragment ions were then used to further confirm the chemical structure. Information including compound name, retention time, formula, precursor ion, fragment ions and literature references of the rest of these compositions can be found in Table S2 and S3. Mass spectra (MS/MS), as well as the fragmentation proposals that led to the suggested structures, are shown in Supplementary Material (Fig. S12 – S73). All structures are shown in Figs. 3, 4 and 6 along with the stereochemistry reported in the literature for the respective metabolites. Global network generated from LC-HRMS/MS are shown in Fig. 5.

All components were identified based on the existing literature [1,6,7,16,18,19,23,24,26,27]. Furthermore, fragmentation pathways of some compounds were proposed in order to facilitate structural identification.
Initially, 18 metabolites (1–18, Fig. 3, Table S2) were annotated by comparison of their accurate mass and MS/MS patterns in literature or databases previously described for Phorbas sp. The known diterpenes phorbasins A, B and H–K were identified (1–6, respectively). Although phorbasin H (3) and I (4) are isomers and differ by the position of one double bond, it was possible to differentiate between them. For differentiation, the diagnostic ions were \( m/z \) 271.1683 for 3 and \( m/z \) 257.1529 for 4 due to neutral loss of 2-methylprop-1-ene and 3-methylbut-1-ene, respectively, through fragmentation pattern (Fig. S17 and S19). These substances have been identified in sponges of the genus Phorbas collected in Australia and phorbasins 2–6 showed cytotoxic activity against non-tumor (NFF) and human tumor cell lines, such as those derived from lung (A549), colon (HT29) and skin (MM96L) cancers [27].

Phorbaketals are sesterterpenoids containing a tricyclic skeleton with a spiral structure and were reported for the first time from sponges of the genus Phorbas collected in South Korea [18]. Here, phorbaketal A, B or C and M were identified (7–9, respectively). Phorbaketals B and C (8) are enantiomers and, therefore, it was not possible to differentiate between them. These metabolites were detected in negative ion mode and had fragmentation of the methyl-dihydropyran ring linked to the alkenyl chain, neutral loss of ethanol for 7 and 8 ([M–H–C\(_{14}H_{22}–\)EtOH], and water for 9 ([M–H\(_2\)O–H–C\(_{14}H_{20}O_2]\]). leading to the fragment ions at \( m/z \) 177.1280 and 175.1524, respectively (Fig. S25, S27 and S29). In relation to their biological activities, phorbaketals A–C exhibited cytotoxicity against human colorectal, hepatoma and lung cancer cell lines [18], while phorbaketal M showed cytotoxicity against a pancreatic cancer cell line, with IC\(_{50} =\)
Isophorbasone A (10) showed unprecedented tricyclic “isophorbasane” carbon skeletons. The ansellone B (11) and 10 share the same biosynthetic pathway related to phorbaketals, phorbasones, and ansellone. Compound 10 was detected as [M+H]+ adduct and showed several fragment ions (Fig. S31), while 11 was detected as [M−H]− and its fragment ions were generated by neutral loss of CO, CO2, H2O and CH4. Metabolite 11 has exhibited potent inhibitory activity on cellular production of nitric oxide (a target for treatment of inflammatory diseases), with IC50 values of 4.5 µmol L−1 [16].

Anvilone A (12), a sesterterpenoid with an unprecedented skeletal class, is biosynthetically associated with other substances from the genus *Phorbas*, such as alloketales and ansellones [7]. Compound 12 was detected as [M+Na−2H]+ adduct and its fragment ions were generated by neutral loss, such as 11. This metabolite has been shown to have pro-viral activity, inducing HIV gene expression [7].

Phorbastereones steroids A–D were identified (13–16, respectively) in the present study. These compounds have the A-ring contracted,
forming a pentane ring as an unusual structural feature that differs by ramifications and side chain isomerism [1]. Phorbasterones have already been identified in P. amaranthus collected from the Caribbean Sea [1]. Furthermore, the authors reported an antifeedant action, which seemingly protects the sponge against predation by a certain species of fish, along with cytotoxic activity against HCT 116 cells, with IC₅₀ ranging between 1 and 3 μg mL⁻¹ [1]. The steroids antherostereones A (17) and B (18) were also identified and differ from phorbasterones by side-chain alkyl branching, isomerism, or oxidation level [1]. Steroids 13, 15, 16 and 18 were detected in positive ion mode as protonated adduct for 15 and sodiated adduct for the others, while 14 and 17 were detected in negative mode as deprotonated for 14 and [M–Na−2H]⁺ for 17. The different adducts did not lead to a fragmentation pattern, except for the loss of methyl formate (60 Da) in A-ring from 13–16 and 18, forming a diketone (Fig. S37, S39, S41, S43 and S47). In all cases, the largest number of fragments occurred in the side chain. Furthermore, in 13 and 15 occurred C-ring fission and A-ring contraction from 17 (Fig S45).

Using GNPS [45] 46 metabolites were further assigned for their structures (19–64, Table S3), most being lysosphopholipids, steroids and carotenoids. Molecular networking has emerged as a powerful analytical tool for dereplication, from which 17 compounds identified herein were directly annotated (19–35, Fig. 4). The metabolic profile data based on fragmentation similarity (MS/MS) and relative intensity, which form groups of compounds with structural similarity as potential biosynthetic analogues, generated several clusters, showed in Fig. 5, with 4 clusters from which further 29 analogues were inferred (Fig. 6, 36–64).

Among the lysosphopholipids (LPL), the subclasses lysosphosphatidylethanolamines (LPE) and lysosphosphatidylcholines (LPCs) were found. The LPE 19 was identified in both ionization modes and has been reported in sponges, such as Stelletta sp. [49] and Homaxinella sp. [50], in diatoms [51] and seaworms [52]. These substances showed anti-inflammatory [51] and cytotoxic activities [49]. Several LPCs (20–31) were identified in the Ph. amaranthus and high levels of LPCs have already been found in sponges Halichondria japonica, Stelletta sp. [49] and Spirastrella purpurea [53], as well as in other marine organisms, such as mollusks, sea cucumber, seaurchin, star-fish, jellyfish and sea anemone [53].

While for LPEs the fragmentation pattern leads to the neutral loss of phosphoethanolamine [M−H−141]⁻, LPEs were evidenced through the similarity of MS/MS fragmentation patterns with four common ions (Fig. S62) at m/z 104.1070, 124.9998, 184.0733 and 258.1101 [54], a loss of water [M−H₂O]⁻ and the molecular ion [M+H]⁺, completing the mass spectrum of LPCs. As the alkyl chain of LPLs has no influence on fragmentation, the metabolites were proposed with side chain and double bond position according to known metabolites reported for sponges or other marine invertebrates. In addition, the strategies for determination of double bond positions for LPLs using MS, in general, includes lithiated adducts, ozone gas, photodissociation, etc., followed by collision-induced dissociation (CID) [55], which were not herein used.

Compound 26 has also been identified in sponges of the species Spirastrella abata [56,57], Haliclonia fulva and H. mucosa [58]. LPCs 26 and 30 have been found in the sponge Suberites domuncula, in which the authors showed the compounds to be indeed synthesized by the invertebrate, suggesting an evolutionary origin and not an association with other microorganisms, as commonly occurs among sponges [59]. The LPC 26 and analogues have been described with a range of biological activities, such as cytotoxic [49,56], antimicrobial [59–62], anti-inflammatory [63], antioxidant [53] and inhibitory effects of cholesterol biosynthesis [56,57].

Two clusters (A and B) grouped as LPL. Fig. 6A shows cluster A, in which nine LPCs were already identified and grouped by dereplication (20, 22, 23, 26–31). Another five were proposed (36–40) with the same fragmentation pattern (Fig. S71), being 36, 37 and 40 previously isolated from the sponge Spirastrella abata [56], while 38 and 39 are analogues of 29 and 27, respectively, with one more methyl group. The LPC 38 has not yet been reported for a marine organism and 39 appears to be a novel compound.

Cluster B, corresponding to the LPL (Fig. 6B), presented the known LPE 19 and other seven LPEs were proposed (41–47), in which the fragmentation patterns leads to the neutral loss of phosphoethanolamine (Fig. S72A). The compounds 41–43, like LPE 19, led to fragment ion m/z 184.0369 (Fig. S72B) due to the formation of an enol between the phosphoethanolamine moiety and the hydroxyl group at the sn-2 position. This ion led to the formation of m/z 104.0706 after neutral loss of metaphosphoric acid (Fig. S72B). The LPE 41 has already been reported for sea urchins [65] and 42 is analogous to 19 with an additional ethyl group. Compounds 43–46 were proposed with an acyl group in sn-2 position due to the absence of fragment ions m/z 184.0369 and 104.0706, which require a hydroxyl group in this position to be generated. The heptenyl was suggested as acyl group due to the fragment ion m/z 129.0910 that corresponds to the protonated heptenoic acid (Fig. S72C). Compound 47 was proposed with alkenyl and acyl group at the sn-1 position and a hydroxyl group at the sn-2 position due to the presence of all three ions m/z 184.0369, 129.0910 and 104.0706 (Fig. S73). LPLs are important molecules in cell signaling and have interesting biological activities [66]. Several LPCs are attributed to strong antitumor activities (against colon cancer, leukemia, lymphoma, ovarian adenocarcinoma, melanoma and glioma) [66] and may contribute to the cytotoxicity found for Ph. amaranthus.

Polyunsaturated fatty acids such as the arachidonic acid (32) identified herein are substances commonly found in marine invertebrates, which has been previously found in Microciona prolifera, a sponge rich in fatty acids [67], and has also been identified in species of sea slug like Aplysia kurodai, A. juliana [68] and Elysia viridis [69].

Among the metabolites annotated by GNPS, only astaxanthin (33) has already been identified in a sponge of the genus Phorbas. This carotenoid has been isolated from the species P. topsenti collected in France [29], but it has also been reported to other sponges such as Gellidites callista [70], Acanthella acuta and Tethya aurantium [71]. It is a strong antioxidant that is commonly found in other marine organisms, mainly fish and crustaceans [72].

Cluster C grouped the astaxanthin analogue with five other substances, thus allowing their inference (48–52, Fig. 6C); all were described as marine metabolites, but only adonixanthin has been reported as a sponge metabolite, in Gellidites callis [70]. In the case of this cluster, observation of the fragmentation pattern of astaxanthin (direct cleavages of conjugated x-system by retro-ene reaction, and vinyl-allyl fragmentation, Fig. S66) was decisive in all cases for the proposition of analogues. The exact masses of substances that led to carotenoids have already been reported for sponges. These substances, however, did not present the same fragmentation patterns of astaxanthin due to structural differences, especially in their hydrocarbon rings [73]. Thus, as well as 33, metabolites 48–52 presented the same fragment ions at m/z 219.1380; 173.1325; 147.0804 and 119.0805 (Fig. S66).

The steroids cholesta-4,6-dien-3-one (34) has been identified in the sponge Condrosia reniformes and its extract has showed toxic effects to mammalian cells (LLCMK² cell lines) with CC₅₀ of 79.8 μg mL⁻¹ [74]. Cholesta-4-en-3-one (35) has been isolated from the sponges Geodia cyclonod (75), Chroella australiensis [76] and Cinachyrella australiensis [77].

Finally, steroids 34 and 35 were grouped in cluster D and another 12 compounds (53–65) could be suggested (Fig. 6D). Although these substances have several fragment ions, most of them are related to the basic scaffold. The steroids in cluster D share fragment ions at m/z 173.0966, 161.1330, 161.0966, 123.174, 109.1017 and 109.0653 through the same fragmentation patterns of 34 and 35 (as shown in Fig. S68 and Fig. S70). Among the proposed steroids, 11 (53–64) have been reported as sponge metabolites.Sterol 59 was suggested as an analogue of 57, but with an additional double bond.

Although other clusters were obtained, none of them presented a known representative metabolite. Even though, the clusters that were not studied disclosed other biosynthetic classes to be further explored.
Moreover, the molecular network expanded knowledge on the substances present in the PetCE of *P. amaranthus*.

### 4. Conclusions

The crude extract obtained from the marine sponge *P. amaranthus* displayed cytotoxicity against the tumor cell line HCT 116 and a partition in hexane (PetHex) revealed comparable bioactivity to that of the crude extract. The automation in column and eluent scouting associated with DryLab® optimization furnished, after only 12 experiments (24 chromatographic runs), a metabolite profile with selectivity and short analysis time that enabled the dereplication by LC-HRMS of 64 annotated metabolites from the PetHex. To the best of our knowledge, eight of those are yet unknown metabolites (39, 42–47 and 64) and ten have never been reported for sponges (19, 22, 24, 25, 29, 38, 41, 48, 49, 51 and 52). Besides the analytical workflow proposed herein, this study has also contributed to the metabolic profiling of *P. amaranthus* and can be employed to guide the isolation of cytotoxic metabolites in complex natural extracts.

### CRediT authorship contribution statement

**Bruno S. Amaral:** Validation, Formal analysis, Writing - original draft, Writing - review & editing. **Fernanda B. Silva:** Investigation, Formal analysis. **Gabriel Mazzi Leme:** Methodology. **leticia S.S. Schmitz:** Formal analysis, Investigation. **Paula C. Jimenez:** Supervision, Writing - review & editing. **Roberto Carlos Campos Martins:** Supervision. **Quezia B. Cass:** Funding acquisition, Project administration, Resources, Supervision, Validation, Visualization, Writing - review & editing. **Alessandra L. Valverde:** Investigation, Methodology, Conceptualization, Funding acquisition, Resources, Supervision, Writing - review & editing.

### Declaration of Competing Interest

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

### Acknowledgments

The authors are thankful to Professor R. A. Epifanio (in memoriam) who collected the material and made the marine invertebrate extracts bank LaProMar. To Dr. Guilherme Muricy of the Laboratory of Porifera at UFRJ for the identification of the species. To Prof. L.V. Costa-Lotufo of ICB-USP (FAPESP No 2015/17177-6) for the generous assignment of their laboratory, reagents, and cell lines for carrying out cytotoxicity assays, as well as to H.V. Forastieri for assistance with cells culture. To MSc. Larissa Ramos Guimarães da Silva and MSc. Luciano da Silva Pinto for helping with GNPS and Cytoscape®.

### Funding

The authors thank for financial support of: Agency for the Improvement of Higher Education Personnel (CAPES) [Finance Code 001], National Council for Scientific and Technological Development (CNPq) [grant number 406064/2018-0 and the scholarship 108535/2017-5], São Paulo Research Foundation (FAPESP) [grant numbers 2014/50244-6 and 2014/50249-8 and the scholarship 2016/24957-0]. GSK is also acknowledged.

### Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jchrohb.2021.122270.


