ACCEPTED MANUSCRIPT

Isolation of anticancer and anti-trypanosome secondary metabolites from the endophytic fungus *Aspergillus flocculus via* bioactivity guided isolation and MS based metabolomics

Ahmed F Tawfike[1,2,a♯*], Muhammad Romli[1], Carol Clements[1], Gráinne Abbott[1], Louise Young[1], Marc Schumacher[3], Marc Diederich[4], Mohamed Farag[5,6] and RuAngelie Edrada-Ebel[1*]

Affiliations

[1]Strathclyde Institute of Pharmacy and Biomedical sciences, University of Strathclyde, Glasgow, G4 0RE, UK

[2]Department of Pharmacognosy, Faculty of Pharmacy, Helwan University, Cairo, 11795, Egypt

[3]Laboratoire de Biologie Moleculaire et Cellulaire du Cancer, Fondation de Recherche Cancer et Sang, Hopital Kirchberg, 9 rue Edward Steichen, L-2540 Luxembourg

[4]Department of Pharmacy, College of Pharmacy, Seoul National University, 1 Gwanak-ro, Gwanak-gu, Seoul, 08826, Korea

[5]Department of Chemistry, School of Sciences & Engineering, The American University in Cairo, New Cairo, Egypt.

[6]Department of Pharmacognosy, Faculty of Pharmacy, Cairo University, Cairo, 11562, Egypt

Correspondence

Dr RuAngelie Edrada-Ebel, Senior lecturer, Strathclyde Institute of Pharmacy and Biomedical sciences, University of Strathclyde, Glasgow, G4 0RE, UK
Email: ruangelie.edrada-ebel@strath.ac.uk, Phone: +44(0)1415485968

Dr Ahmed F Tawfike, Lecturer, Department of Pharmacognosy, Faculty of Pharmacy, Helwan University, Cairo, 11795, Egypt.

[a♯] Current address: Molecular discovery group, Department of Computational and Analytical Science, Rothamsted Research, Harpenden, AL5 2JQ, England, UK.
Email: ahmed.tawfike@rothamsted.ac.uk, Phone: +44(0)1582938120

ACCEPTED MANUSCRIPT

Abstract

This study aims to identify bioactive anticancer and anti-trypanosome secondary metabolites from the fermentation culture of *Aspergillus flocculus* endophyte assisted by modern metabolomics technologies. The endophyte was isolated from the stem of the medicinal plant *Markhamia platycalyx* and identified using phylogenetics. Principle component analysis was employed to screen for the optimum growth endophyte culturing conditions and revealing that the 30-days rice culture (RC-30d) provided the highest levels of the bioactive agents. To pinpoint for active chemicals in endophyte crude extracts and successive fractions, a new application of molecular interaction network is implemented to correlate the chemical and biological profiles of the anti-trypanosome active fractions to highlight the metabolites mediating for bioactivity prior to purification trials. Multivariate data analysis (MVDA), with the aid of dereplication studies, efficiently annotated the putatively active anticancer molecules. The small-scale RC-30d fungal culture was purified using high-throughput chromatographic techniques to yield compound **1**, a novel polyketide molecule though inactive. Whereas, active fractions revealed from the bioactivity guided fractionation of medium scale RC-30d culture were further purified to yield 7 metabolites, 5 of which namely *cis*-4-hydroxymellein, 5-hydroxymellein, diorcinol, botryoisocoumarin A and mellein, inhibited the growth of chronic myelogenous leukaemia cell line K562 at 30 μM. 3-hydroxymellein and diorcinol exhibited a respective inhibition of 56% and 97% to the sleeping sickness causing parasite *Trypanosoma brucei brucei*. More interestingly, the anti-trypanosomal activity of *A. flocculus* extract appeared to be mediated by the synergistic effect of the active steroidal compounds i.e. ergosterol peroxide, ergosterol and campesterol. The isolated structures were elucidated by using 1D, 2D NMR and HR-ESIMS.

Keywords:

Endophytic fungi, *Aspergillus flocculus*, Metabolomics, LC-MS, Anti-trypanosome, Anti-cancer

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

1. Introduction

Endophytes are microbes that are harboured inside plant internal tissues without triggering any immediate, apparent negative effects [1]. Piling evidence points to their possible involvement in the biosynthesis of plant natural products, or even that they might be the sole producers of several other groups of novel pharmacologically active and structurally diverse secondary metabolites [2]. Advanced methods in natural products drug discovery have provided an access to a rich source of novel drug leads having the advantage of optimizing yield production levels *via* the large-scale cultivation of the microorganisms. *Aspergillus* is a fungal genus belonging to Ascomycetes fungi that are found in several terrestrial and marine organisms. The endophytic *Aspergillus flocculus* was first isolated from the stem of the Egyptian medicinal plant *Markhamia platyclayx*. Roots, stems, barks, and leaves of *Markhamia* species are used in traditional medicine to treat various ailments. In Africa, *Markhamia* plant parts were traditionally used for treating microbial and parasitic diseases, anaemia, diarrhoea, backache, sore eyes, pulmonary troubles, gout, scrotal elephantiasis, rheumatoid arthritis, and superficial skin condition [3]. In terms of chemistry, *Markhamia* species produce a myriad of metabolites i.e. antitumor naphthofurandione and naphthoquinones from *M. tomentosa* and *M. hildebrandtii resepectively* [4], cytotoxic sterols from *M. zanzibarica* [5], anti-parasitic triterpenoids musambins A-C and their glycosides in *M. lutea* [6] in addition to polyphenols [7, 8]. For thorough review on bioactive metabolites from the terrestrial *Aspergillus* endophyte, our previous publication ought to be consulted [9]. Continuing our interest in the metabolites profiling of plant fungal endophytes, we report herein on the isolation of bioactive anticancer and antitrypanosomal metabolites from *A. flocculus* guided by metabolomics modern metabolomics tools.

In a preliminary screening for anticancer and antimicrobial active agents from endophytic fungi, *A. flocculus* an endophyte associated with the stem of *M. platycalyx* (Bignoniaceae) exhibited toxicity against prostate cancer cell line (PC3) and chronic myelogenous leukemia (K562) (Fig.1-3S) [9]. Cancer Research UK recorded 356,860 cancer cases in 2014 with estimated deaths of 163,444 which is more than 45% mortalities. Prostate cancer was the fourth most common cause of cancer death in the UK with more than 11.000 deaths recorded in 2016 [10]. According to the International Agency of Research on Cancer, an estimated 1.1 million men worldwide were diagnosed with prostate cancer in 2012, considered as 15% of the cancers diagnosed in men. Between 2009 and 2013, leukaemia was the fifth most common cause of cancer death in men and the sixth in women. In 2017, 24,500 people in US are estimated to die from leukaemia (14,300 males and 10,200 females). In addition, there is

ACCEPTED MANUSCRIPT

approximately more than 380,000 people living with, or in remission from, leukaemia [11]. Human African Trypanosomiasis (HAT), or sleeping sickness, is a vector-borne parasitic disease, caused by the parasitic protozoan *T. brucei.* Trypanosomiasis is one of the most neglected tropical diseases that occurs in sub-Saharan Africa and is regarded as being life threatening if left untreated. HAT affects remote rural communities in isolated regions with inadequate access to suitable health services, with many cases that could not be diagnosed or reported and ultimately affecting the true statistics of disease prevalence in Africa [12]. All available medications used for HAT treatment must be taken by injection over a long-time thus requiring medical facilities and specialized staff that unfortunately often do not exist in rural infected areas. Adverse effects are severe and sometimes fatal [13]. Eradication of HAT is considered though possible by World Health Organization (WHO) [14]. Consequently, reliable methods for diagnosis, novel, safe, effective, and easy-to-apply drugs are still warranted [15]. Historically, natural products have been recognized as a rich source of compounds that have played a potential role in ailments treatment and for maintaining a better health status. The large structural diversity of natural products along with their myriad of biological effects has served as "lead" compounds for drug design programs. In fact, Food and Drug Administration (FDA) reported that 34% of discovered drugs between 1981-2010 were derived from natural products [16]. Moreover, the recent technological advances in high throughput screening along with metabolomics and dereplication studies has led to a paradigm shift in natural products drug discovery. Metabolomics is the technology designed to provide general qualitative and quantitative profile of metabolites in biological systems at different status conditions, with many applications in aspects related to drug discoveries, particularly when coupled to bioactivity assay to expedite the identification of bioactive agents. Particularly of value among the different metabolomics platforms, liquid chromatography coupled to mass spectrometry LCMS can generate an informative rich data set that can assist in the tentative identification of natural products classes present in crude extracts prior to an intensive isolation attempts. For pinpointing active components, multivariate data analysis is normally employed to correlate bioactivity and LCMS profiles [17]. In this study, we used a molecular interaction correlation network that uses the Pearson correlation coefficient to link the chemical and biological profile of either active extracts or fractions. Thus, the metabolites contributing to the bioactivity can be putatively identified before extensive isolation attempts.

2. Materials and Methods

ACCEPTED MANUSCRIPT

2.1. Fungal Material

The endophytic fungus was isolated from the fresh stems of *Markhamia platycalyx* family (Bignoniaceae) collected in October 2010 from Al-Zohriya gardens (Al-Zamalek, Giza, Egypt). Plant material was authenticated by Dr. Therese L. Yousef, senior expert at Orman Garden and with the voucher specimen (No. 633) deposited. Plant material was cut into small pieces, washed with sterilized demineralized water, then thoroughly surface sterilized with 70% isopropanol for 1-2 seconds and ultimately air dried under a laminar flow hood. Such procedure is done to eliminate surface contaminating microbes. With a sterile scalpel, outer tissues were removed from the plant and the inner tissues were carefully dissected under sterile conditions and laid over malt agar (MA) plates containing chloramphenicol to prevent bacterial contamination. Post 4 weeks of incubation at 30°C, hyphal tips of the fungi were removed and transferred to a fresh MA medium. Plates were prepared in duplicates to eliminate the possibility of contamination. Pure strains were isolated by repeated inoculation and with the purified fungus later transferred to the rice solid medium for scaling up.

2.2. Identification of fungal strain

The fungus was annotated as as *Aspergillus flocculus* on the basis of sequence similarity of the ITS region as described previously [18]. DNA extraction and gene amplification was performed using (RED Extract-N-Amp™ Plant PCR Kit, Sigma-Aldrich, UK) following the exact protocol described by Debbab 2009 [19]. BLAST search of the FASTA sequence was performed with the option “nr”, including GenBank, RefSeq Nucleotides, EMBL, DDBJ and PDB sequences on the BLAST homepage, NCBI, Bethesda, USA. The GeneBank accession number was FJ491571 and fungal strain MS-F2 was archived in the microbial culture collection of the Edrada-Ebel Research Laboratory, SIPBS, University of Strathclyde, UK.

2.3. Extractions of endophytic metabolites for screening and metabolites profiling analysis

A plate of each fungal species was transferred into a 250-ml flask, then macerated with ethyl acetate and left overnight followed by homogenization and filtration. The mycelia was further macerated for three more times with 200 mL ethyl acetate and filtered using glass Buchner filter funnel with sintered glass disc. The filtrate was then combined and dried under vacuum using BUCHI Rotavapor R-3 at 45°C. Dried obtained residue was resuspended in 200 mL distilled H_2O and partitioned by adding EtOAc (3 × 200 mL) in a separating funnel. To remove the water-soluble content of the media, water soluble

ACCEPTED MANUSCRIPT

portion was concentrated and then passed over HP-20 column (Approximately 400 cm^3 of Diaion HP-20) using methanol as an eluent till exhaustion. Methanol and ethyl acetate soluble portions were concentrated via rotary evaporator and 1 mg of each extract was subjected to HRMS analysis, whereas 8-10 mg was aliquoted for NMR analysis.

2.4. Small-Scale Fermentation, Extraction and Metabolites Isolation

A small-scale fermentation was carried out in two Erlenmeyer flasks (1L each) on rice medium, which was prepared with 100 g of rice powder and approximately 100 mL of demineralized water just enough to cover the rice layer. The rice media was autoclaved prior to fungal inoculation. A 15-day fungal inoculum grown on petri dish was inoculated on the sterile rice medium and was allowed to grow at room temperature under static condition for 30 days. The fermentation was stopped by adding 500 mL of EtOAc to each flask. Culture media were then cut into pieces to allow complete maceration and left for three days. Then filtration was done followed by repeated extraction with EtOAc until exhaustion. The combined EtOAc extracts were evaporated under vacuum, suspended in 200 mL H_2O and partitioned by adding EtOAc (3x200 mL) in a separating funnel. The pooled EtOAc extracts were then taken to dryness under vacuum. A portion of EtOAc extract (2.6 g) was subjected to size exclusion chromatographic separation on Sephadex LH-20 column (25mm x 510mm) with 100% methanol as mobile phase. 11 fractions resulted, evaporated and subjected to TLC. Fraction 5 (code: Seph-5), has the highest yield (500 mg), was selected for further fractionation work on MPLC using VersaPack C18 silica gel column (23 x 53 mm). Flow rate was at 10 mL/min and H_2O/MeOH was used as mobile phase at gradient elution starting at 100% H_2O reaching to 100% MeOH for 60 minutes, which led to the isolation of 12 sub-fractions. The sub-fractions were evaporated and subjected to LC-HRMS and NMR analysis. Sub-fraction M6 (12 mg) eluted with 70% MeOH was corresponding to compound **1** (new polyketide). In order to optimize the best elution system, another portion of EtOAc extract (2.3 g) was subjected to MPLC fractionation using VersaPack silica gel column (23 x 110 mL) at a flow rate of 15 mL/min and n-hexane/EtOAc as mobile phase with gradient elution starting at 100% n-hexane and reaching to 100% EtOAc in 2.1 hours run. The MPLC fractionation led to isolation of 253 fraction 50 mL each. Fractions were subjected to TLC where similar fractions were pooled and dried for further analysis. Fraction MPLC-24 (16 mg) eluted with 80% EtOAc was equivalent for compound **2** (dihydroaspyrone).

ACCEPTED MANUSCRIPT

2.5. Bioactivity guided isolation of medium-scale 30-day rice culture (RC-30d)

A medium scale fermentation was performed in 10 Erlenmeyer flasks (1L each) on rice solid medium for 30 days under same condition applied to small scale culture. The EtOAc extract of the medium scale batch of 30-day rice culture (30 g) was dried and reconstituted in MeOH which was slowly evaporated that led for the precipitation of Kojic acid crystals (**3**) (4.8 g) which was further purified by decantation. The rest of the MeOH liquor (25 g) was dried under vacuum, dissolved in 10% aqueous MeOH and partitioned with n-hexane in a separating funnel as a defatting step. The *n*-hexane soluble layer was concentrated and subjected to further fractionation on silica gel open column (19 mm x 46 mm) using DCM/EtOAc (90:10) as eluent to yield compounds **4** (ergosterol) 8 mg, **5** (ergosterol peroxide) 7 mg and **6** (campesterol) 9 mg. While the methanol soluble portion (5.1 g) was fractionated on a Grace flash chromatography instrument using silica gel cartridge 80g (186 mm (L) x 32 mm (ID), 40µm particle size) with *n*-hexane and EtOAc as mobile phase on gradient elution started at 100% n-hexane reaching to 100% EtOAc in 2 hours 15 minutes. This was followed by another run, to elute the highly polar compounds, employing EtOAc/MeOH as a gradient elution system started at 0% MeOH reaching 30% MeOH in 20 minutes. The fractions were collected in 20 mL volume each, evaporated under nitrogen and pooled in accordance to the detector (UV and LSD) chromatogram and TLC. Pooled-fractions resulted from both runs, evaporated and subjected for further analysis. Fractions were tested for their ability to inhibit trypanosomal viability, potentially active fraction F18-29 was pooled with moderately active F11-18 to increase the yield that will be subjected to further purifications. The pooled fraction F11-29 (216 mg) was then further fractionated using Grace Silica cartridge 12g (82 mm x 22 mm) with isocratic elution system employing *n*-hexane/CH_2Cl_2 at ratio (65:35) in 55 minutes run. This resulted in isolation of compounds **7a** and **7b** in a mixture (11mg) which equivalent for botryoisocoumarin A and mellein respectively. Similarly, fractions F53-55, F56-64 and F65-71 were pooled into F53-71 (157 mg) that was chromtographed on Grace silica cartridge 12g (82 mm x 22 mm) with isocratic elution system employing n-hexane/CH_2Cl_2/EtOAc at ratio (70:20:10) in 65 minutes run which gave instantaneously 71 sub-fractions. The sub-fractions have pooled according to TLC and yielded 13 pooled sub-fractions. Sub-fraction F40-53 (40 mg) was further purified on Biotage silica cartridge SNAP 10 g (20 mm x 60 mm) using *n*-hexane/CH_2Cl_2/EtOAc at ratio (70:20:10) in 60 minutes run to give compound **8** (*cis*-4-hydroxymellein) 13 mg and **9** (*trans*-4-hydroxymellein) 15 mg. Fraction F72-89 (479 mg)

ACCEPTED MANUSCRIPT

was fractionated using Grace silica cartridge 12g (82 mm x 22 mm) with isocratic elution system employing *n*-hexane/CH_2Cl_2/EtOAc at ratio (70:20:10) in 62 minutes run which gave instantly 92 sub-fractions. Sub-fractions F41-47 (60 mg) and F32-40 (53 mg) were applied on Biotage silica cartridge SNAP 10 g (20 mm x 60 mm) using *n*-hexane/CH_2Cl_2/EtOAc at ratio (70:20:10) in 45 minutes run to give compounds **10** (10 mg) and **11** (12 mg) equivalent for 3-hydroxymellein and diorcinol respectively, from sub-fraction F41-47. While sub fraction F32-40 yielded compound **12** corresponded to 5-hydroxymellein. The inactive fractions were also further fractionated in search for either new chemistry or other bioactivities. Fraction F90-135 (2.7 g) was subjected to fractionation using Grace Silica cartridge 40 g (122 mm x 27 mm) with isocratic/gradient elution system employing CH_2Cl_2/MeOH commenced with 0% MeOH reaching to 3% MeOH in 5 min that remained for 20 min isocratic elution. Followed by gradient elution started at 3% MeOH to 30% in the next 20 mins. This resulted in 74 fractions, which were evaporated and pooled according to TLC to give 15 sub-fractions. Sub-fraction F16-43 (500 mg) was applied on Biotage silica cartridge SNAP 25g (29 mm x 78 mm) to be fractionated using *n*-hexane/CH_2Cl_2/EtOAc at ratio 70:20:10 for 30 minutes followed by EtOAc/MeOH gradient elution system started at 0% MeOH reaching to 30% in 20 minutes. This resulted in elution of F113-117 (189 mg) corresponded to compound **13** (7-*O*-acetyl kojic acid), F2-19 (12 mg) which corresponded to compound **14** (methyl 2-(4-hydroxyphenyl) acetate) and F108-112 (51 mg). F108-112 was subjected to fractionation using preparative TLC using *n*-hexane/CH_2Cl_2/EtOAc (70:20:10) for three cycles which resulted in isolation of three bands at Rf values 0.8, 0.7 and 0.3 equivalent for compounds **15** (7 mg), **16** (5mg) and **17** (11 mg) corresponded to phomaligol A, phomaligol A1 epimer and dihydropenicillic acid respectively. Sub-fraction F51-66 (1.5 g) was fractionated using Grace Silica cartridge 40 g (122 mm x 27 mm) with gradient elution system employing CH_2Cl_2/EtOAc, started by 0% EtOAc reaching 100% in 55 minutes and followed by another run using gradient elution system EtOAc/MeOH started at 0% MeOH to 30% in 20 minutes. This resulted in elution of F2-8 (61 mg) which has been purified using preparative TLC with *n*-hexane/EtOAc at ratio 70:30 as a mobile phase. This resulted in separation of compound **18** (8 mg) equivalent for *p*-hydroxy benzaldehyde; F17-37 (520 mg) that has been subjected to fractionation using Grace silica cartridge 12 g (82 mm x 22 mm) with isocratic/gradient elution system employing CH_2Cl_2/MeOH. The run commenced with 0% MeOH reaching to 3% in 5 minute, remained at 3% MeOH for another 25 minutes and followed by gradient elution started at

ACCEPTED MANUSCRIPT

3% MeOH reaching to 20 % in 15 minutes run. This led to isolation of F36-37 (15 mg) which, was a mixture of compounds **19a** (2-hydroxyphenyl acetic acid) and **19b** (4-hydroxyphenyl acetic acid); F33-35 (170 mg) was purified using preparative TLC with mobile phase CH_2Cl_2/MeOH at ratio 95:5 three cycles to give a band at R_f value 0.7. The band was corresponded to compound **20** (17 mg) equivalent to 4, 5-dihydroxymellien. Scheme of the isolated compounds is shown in Fig.17S (SI).

2.6. General experimental procedure

1D and 2D-NMR experiments were recorded at 25 °C in DMSO-d_6 and $CDCl_3$ on JEOL JNM-LA400 NMR spectrometer (JEOL, Ltd., Tokyo, Japan). Chemical shifts were referenced to the solvent residual peaks at δ_H 2.50 and 7.26 for ^1H and δ_C 39.5 and 77.16 for ^{13}C for DMSO and $CDCl_3$, respectively. Electrospray ionization high resolution mass spectrometry (ESI-HRMS) was measured using Fourier transform (FTHRMS)-Finnigan LTQ Orbitrap or Exactive mass spectrometer (Thermo Fisher Scientific, Massachusetts, USA). The samples were run in duplicate, the MS detection range was from m/z 100–2000 and scanning was performed under ESI polarity switching mode. The needle voltages were −4.0 kV, 4.5 kV positive and the sheath and auxiliary gases were set at 50 and 17 arbitrary units, respectively. HPLC analysis was performed with Dionex UltiMate3000 (Thermo Fisher Scientific, Massachusetts, USA) coupled to photodiode array detector (DAD3000RS). The samples were chromatographed *via* a C-18 column (ACE) with a length of 75mm, internal diameter of 3.0 mm and silica particle size 5 µm. A mobile phase composed of 0.1% formic acid HPLC-grade water (solvent A) and acetonitrile (solvent B) set at a flow rate of 300µL/min was used for peaks elution. Gradient elution was employed, commencing with 10% B for 5 minutes increased to 100% B over 30 minutes, and kept for another 5 minutes before decreasing to 10% B. The column was then equilibrated with 10% B for 4 minutes until the end of run.

A medium pressure liquid chromatography (MPLC) from BÜCHI was used for the isolation of fungal fractions, MPLC instrument was the Sepacore Purification System with Versaflash column stand (BUCHI UK Ltd, Oldham, UK). The Reveleris® Flash Forward system of Grace Davison Discovery Sciences (Illinois, USA) was also used for further purification, which is characterized of having two detectors, an evaporative light scattering detector (ELSD) and a UV detector (wavelength range: 200-500 nm).

2.7. LC-MS data analysis, processing and visualization

ACCEPTED MANUSCRIPT

LC-MS spectra were processed using Thermo Xcalibur 2.1 (Thermo Fisher Scientific, Massachusetts, USA). To convert the raw data into separate positive and negative ionisation files, msconvert from ProteoWizard was used [20]. The files were then imported to the data mining software MZmine 2.20 (VTT, Espo, Finland) for peak picking, deconvolution, deisotoping, alignment and formula prediction [9, 17]. Macro file of mass peaks abundance was written to an Excel file, used to combine positive and negative MS files and for further clean-up of media components [21]. For metabolites identification via LCMS, the Dictionary of Natural Products (DNP) 2015 was used. MestReNova (MNova) 2.10 (MestrelabResearch, S.L, Santiago de Compostela, Spain) was used to process NMR data. SIMCA 14 (Umetrics AB, Umeå, Sweden) was used for multivariate data analysis using centre scaling. Heatmap was created using http://www.metaboanalyst.ca/.

2.8. Molecular network

A molecular interaction network was created via specific application of the cytoscape software (version 3.4.0)[22]. The ExpressionCorrelation app, implemented by Sander Group, (Computational Biology Center, Memorial Sloan-Kettering Cancer Center, New York City), was used to compute a similarity network from either observation (active fractions) or their corresponding features (m/z) in data matrix. Similarity network is using the Pearson correlation coefficient to link the active fractions (observations correlation network) or their corresponding metabolites (features correlation network). A feature correlation network was created to explore which of the metabolites will be highly correlated with the bioactivity (represented by percentage of viability). The negative correlation threshold was set to 0.7 whereas the positive one was neglected. The network was mapped via organic yfiles layout, a kind of spring-embedded algorithm that combines elements to show the clustered structure of a graph. Features (metabolites) were represented by nodes which are linked by edges (correlation value). The width of the edge is corresponding to the strength of the correlation.

2.9. In vitro Anti-trypanosomal activity via Alamar blue assay

Anti-trypanosomal activity was tested following the protocol reported by (Räz et al., 1997)[23]. Test compounds were prepared as 10 mg/mL stock solutions in 100% DMSO. The samples were initially screened at a concentration of 20µg/mL and then minimum inhibitory concentration (MIC) determinations were determined for the active compounds. MICs were carried out in 96-well microplates. DMSO at a concentration

ACCEPTED MANUSCRIPT

range of 0.001–1% and suramin over a concentration range of 0.8–100 µM were included as negative and positive controls.

2.10. Anticancer activity for crude extracts and fractions

2.10.1. Cell culture and treatment

Crude extracts and fractions were tested at Laboratoire de Biologie Moleculaire et Cellulaire du Cancer (LBMCC). The in vitro growth inhibitory ability of the extracts, fractions and pure compounds were tested on different cell lines including Human Philadelphia chromosome-positive chronic myelogenous leukemia cells (K562) and prostate cancer cells (PC3). The cell lines were purchased from Deutsche SammLung für Mikroorganismen und Zellkulturen (DSMZ, Braunschweig, Germany) and cultured in RPMI 1640 medium (Lonza, Verviers, Belgium) supplemented with 10% fetal calf serum (FCS) (Hyclone, Perbio, Erembodegem, Belgium) and 1% (v/v) antibiotic– antimycotic (Lonza, BioWhittaker™, Verviers, Belgium) at 37℃, in a 5% CO_2, humidified atmosphere. Human recombinant TNFa (PeproTech, Rocky Hill, NJ, USA) was re-suspended in a phosphate buffer salt (PBS) 1X sterile solution containing 0.5% bovine serum albumin (MP Biomedicals, Asse-Relegem, Belgium) to reach a final concentration of 10µg/mL [24].

2.10.2. Transient transfection and luciferase reporter gene assay

Transient transfections of K562 cells were achieved as previously described [25]. As summary, 5 µg of luciferase reporter gene construct containing five repeats of a consensus NF-κB site (Stratagene, Huissen, Netherlands) and 5 µg Renilla luciferase plasmid (Promega, Leiden, Netherlands) was utilized for each pulse. After electroporation, the cells were re-suspended in growth medium (RPMI/FCS 10%) and incubated at 37 °C which followed by 5% CO_2. 20 h after transfection. The cells were then harvested and re-suspended in growth medium (RPMI/FCS 0.1%) to a final concentration of 106 cells/mL and treated for 2 h with or without the natural compound. Successively, the cells were challenged with 20 ng/mL TNFα for 6 h. 75 µl Dual-Glo™ Luciferase reagent (Promega) was added to the cells for 10 min incubation at 22 °C followed by measuring luciferase activity. Then, 75 µl Dual-Glo™ Stop and Glo1 reagent (Promega) were added for 10 min at 22 °C in order to assay Renilla activity. Luciferase and Renilla (Promega) activities were measured using an Orion microplate luminometer (Berthold, Pforzheim, Germany) by integrating light emission for 10 s. The results are expressed as a ratio of arbitrary units of firefly luciferase activity normalized to Renilla

ACCEPTED MANUSCRIPT

luciferase activity [24].

2.10.3. In vitro cytotoxic assay (viability assay)

The cell lines were preserved in continuous culture in a humid atmosphere at 37 °C and 5% CO_2 in RPMI 1640 medium supplemented with penicillin, streptomycin, gentamicin, L-glutamine, and fetal calf serum. Potential mycoplasm contaminations were checked twice per month. For the assay, 24-well plates were seeded with 500 µL of cell suspension containing 2 x 10^5 cells/mL. Cells were treated with natural compounds at 30µM. After 24 h incubation, the cells were transferred to 96-well microplates and assayed for the CellTiter-Glo® Luminescent Cell Viability Assay to calculate the number of viable cells in culture, based on quantification of the ATP present. Each condition was performed in triplicate. The results correspond to an average of three independent experiments ± SD [24].

2.11. Anticancer activity of pure compounds

Purified compounds were tested at Strathclyde Innovations in Drug Research (SIDR) lab. Human Philadelphia chromosome-positive chronic myelogenous leukemia cells (K562) and prostate cancer cells (PC3) were purchased from the Global Bioresource Centre ATCC. The cells were cultured in an incubator in 5% CO_2 at 37 °C and then seeded at a density of 2 x 10^4 cells/well in 100 µL of the growth medium Minimum Essential Medium alpha (MEMα) supplemented with 20% foetal bovine serum (FBS). The growth medium was replaced (after 48h) with 80 µL of new medium that contained the purified compounds at concentrations 30 µM (mixtures are tested at 30 µg/mL). 20 µL of 50% PrestoBlue® (resazurin based, Life Technologies) solution (diluted with growth media) was added to the cells after 48h incubation for viability measurements. This mixture was incubated for 60 min before reading the fluorescence (570 nm excitation/585 nm emission). The percentage of treated cell viability was calculated and compared to untreated cells. Tests were done in triplicate. Additionally, cells were observed and micrographs taken using an EVOS XL Core imaging station (Life Technologies) with phase contrast before treatment, after 24 h and 48 h.

Normal epithelial cells derived from human prostate (PNT2 cells) were obtained from ECACC (Sigma-Aldrich, Dorset, UK). PNT2 cells were cultured in RPMI 1640 media; supplemented with 10% (v/v) foetal bovine serum, 2 mM glutamine and 50 µg/mL penicillin/streptomycin solution (all Invitrogen, Paisley, UK) in a humidified incubator at 37 °C in the presence of 5% CO2. The confluence of cells was 90%–95%. Consequently,

ACCEPTED MANUSCRIPT

cells were seeded at a concentration of 3750 cells/well in clear 96 flat-bottomed plates and allowed to adhere overnight. After that time, purified compounds were added at a final concentration of 30 µM and allowed to incubate for 42 hours. Viability was determined using Alamar Blue® (Thermo Fisher, Paisley, UK), according to the manufacturer's instructions and incubated for a further 6 h. The resulting fluorescence was measured using a Wallac Victor 2 1420 multi-label counter (Perkin Elmer, Beaconsfield, UK), in fluorescence mode: 560/590 nm (ex/em). Vehicle treated control cells (media with 0.3% DMSO) were considered 100% viable against which compound treated cells (at a concentration of 30 µM, n = 3) were compared. All results were confirmed microscopically [26].

3. Results and discussion

Metabolite profiling, dereplication study and multivariate data analysis of the crude extracts

We have previously developed an high-performance liquid chromatography coupled to mass spectrometry (HPLC-MS) method for profiling of fungal endophytes [17] which we apply herein for the profiling of *A. flocculus* culture. To provide a comprehensive coverage of *A. flocculus* metabolome, fungal extract was analysed in both positive and negative ion electrospray ionization (ESI) MS modes as changes in ESI polarity can often circumvent or significantly alter competitive ionization and suppression effects revealing otherwise suppressed metabolite signals [27]. Although different metabolite patterns could be observed by simple inspection of HPLC–MS chromatograms from the different fungal cultures (Fig.1a &b), principal component analysis (PCA) and heatmap multivariate data analyses were employed as a more holistic approach to determine relative variability within the different cultures type and help select one for the large-scale culturing and further detailed isolation attempt. PCA is an unsupervised clustering method requiring no knowledge of the dataset and acts to reduce the dimensionality of multivariate data while preserving most of the variance within [28]. The mass signals extracted by MZmine 2.20 software from the raw LC–MS dataset were subjected to firstly PCA analysis.

Fig. 1 Base peak chromatogram of *A. flocculus* various extracts acquired in negative (a) and positive ionization mode (b). Including malt agar (MA-plate), 7 days liquid culture (LC-7d), 30 days liquid culture (LC-30d), 7 days rice culture (RC-7d) and 30 days rice culture (RC-30d) extracts.

ACCEPTED MANUSCRIPT

The main principal component (Fig. 2A) differentiated between the various culture extracts, *i.e.*, PC1, PC2 accounted for a respective total variance 34%, 25% and revealed for the dispersal of the 30-day rice culture extract (RC-30d) arguing for its distinct chemical profile that was further confirmed from the heatmap plot (Fig. 2C). The unique chemical fingerprints of the RC-30d extract displayed by heatmap led to a corresponding discrimination in PCA score plot. PCA loading plot (Fig.2B) highlighted the metabolites contributing to such segregation. Those metabolites were dereplicated using Dictionary of Natural Products (DNP). The resulted hits were reduced by applying a chemotaxonomic filter and only the predicted molecular formulas were selected from the identified hits. The top discriminatory compounds, corresponding to *m/z* (retention time in min) 475.316 $[M+H]^+$ (35.89), 187.096 $[M+H]^+$ (3.12) and 318.279 $[M+H]^+$ (17.01), were putatively identified as BU-4514N ($C_{27}H_{42}N_2O_5$), dihydroaspyrone ($C_9H_{14}O_4$) and preussin ($C_{21}H_{35}NO$), respectively. BU-4514N, antibiotic related to lydicamycin, was isolated from the fermentation broth of *Microtetraspora* sp. T689-92 and reported for its significant neurogenetic and antibacterial activities against gram positive bacteria [29]. Dihydroaspyrone, a common metabolite found in *Aspergillus* species [30-32], reported for its toxicity against L-1210 (mouse lymphocytic leukaemia cells) and Methicillin-resistant *Staphylococcus aureus* (MRSA) [30]. Preussin, produced by *A. ochraceus* ATCC22947, *Preussia* sp. and *Simplicillium lanosoniveum* TAMA 173, was reported for its anti-fungal and cytotoxic activities [33, 34]. The bioactivities reported for the discriminatory metabolites of RC-30d culture indicated a promising biological and chemical metabolic profile. RC-30d culture extract exhibited the strongest inhibition against NF-κB (Fig.4S). NF-κB is a protein complex that controls transcription of DNA, cytokine production and overall cell survival [35]. NF-κB activation contributes to an inflammatory or immune response and a cellular proliferation. NF-κB inhibition consequently can suppress cancer growth and angiogenesis. Compounds, demonstrating NF-κB inhibition, may exhibit anticancer activity [36] and regulate inflammation-associated with cancer development [37]. Thus, the dereplication and biological investigation results encouraged the further purification of RC-30d fungal extract to confirm the identity of the discriminatory metabolites and/or exploring new chemical structures. Although, a medium scale-up fermentation culture of RC-30d extract was established for the bioactivity guided isolation of active compounds, the small-scale fermentation culture, used for optimizing purposes and metabolomics study, was purified to yield compounds **1** (a new polyketide)

ACCEPTED MANUSCRIPT

and **2** (dihydroaspyrone), that were found to be inactive either against (chronic myologeous leukemia (K562) and prostate cancer (PC3)) cell lines or *T. brucei* (Fig.5S,6S). However, compound **1** is a new polyketide isolated herein for the first time from a natural source.

Fig. 2 a: Principle component analysis (PCA) score plot of various culture extracts of *A. flocculus* b: PCA loading plot of various culture extracts where metabolites are colour coded based on their *m/z* range, c: Heat map of *A. flocculus* culture extracts showing the metabolites pattern responsible for the variation of RC-30d extract

Metabolite profiling, dereplication study and multivariate data analysis of the bioactive fractions

The bioactivity guided fractionation of the medium-scaled RC-30-d culture led to 20 major fractions which showed various anticancer and anti-trypanosomal activities. Fractions showed the strongest inhibition against *T. brucei* were those from 18-29 to 72-89 with 93-100% inhibition at 20 µg/mL comparable to standard positive drug Suramin (Fig.7S). Moreover, F51-52 and F53-55 exhibited a significant cytotoxicity against chronic myelogenous leukemia cell line (K562) whereas, F53-55 was active against prostate cancer cell line PC3 at concentration 10 µg/mL (Fig.8S, 9S).

To investigate which of the metabolites in the anti-trypanosome active fractions are likely to mediate the biological activity prior to purification attempt, a molecular interaction network was created by employing specific application of the cytoscape software. Similarity network used the Pearson correlation coefficient to link either the active fractions (observations correlation network) or their corresponding metabolites (features correlation network). Features correlation network (Fig. 3A) was implemented to assign metabolites significantly correlated to the bioactivity (represented by percentage of cell viability). The network was based only on the negative correlations. Thus, only metabolites having high correlation coefficient will be expected to link with the decreased cell viability in the bioactive fractions. Features are connected by edges (correlation values) where edge's width is corresponding to the strength of this correlation. A network of metabolites linked to the bioactivity is depicted in Fig. 3B, with metabolites mass signals at *m/z* 952.753 $[M-H]^-$, 957.766 $[M+H]^+$, 251.128 $[M+H]^+$, 955.757 $[M+H]^+$, 929.742 $[M+H]^+$ and 778.656 $[M+H]^+$, strongly connected to the bioactivity. Searching for their *m/z* against DNP database after chemotaxonomic filtration and matching the

ACCEPTED MANUSCRIPT

predicted molecular formula, resulted in no known hits except for the molecular ion peak at m/z 251.128 $[M+H]^+$ eluting at 13.6 min equivalent to aspergillumarin B ($C_{14}H_{18}O_4$) as shown in Fig. 3B. Aspergillumarin B was reported from *Aspergillus* sp. for its moderate antimicrobial activity against *Staphylococcus albus* [38].

Fig. 3 a: Feature correlation network of the antitrypanosome active fractions, b: Extracted network of the metabolites strongly correlated to the bioactivity

The dereplication result was further confirmed by NMR analysis of the active fractions. The ABC spin system on the benzene ring of aspergillumarin B was revealed from ^{1}HNMR spectrum of the pooled active fractions (33-52) Fig. 4. In addition, presence of more than one broad peak at 11 ppm, equivalent for phenolic OH, indicates the presence of more than one structural analogue. To validate this work, mass spectral (MS) data for the isolated fractions was subjected to further multivariate data analysis (MVDA) (Fig.10S) to pinpoint metabolites highly correlated to the anti-trypanosomal activity in these fractions. Results for correlated metabolites along with their MS data are presented in Table 1. Hits were found comparable to those detected by molecular correlation network (MN) i.e. molecular ion peaks at m/z 955.757 $[M+H]^+$ and 929.742 $[M+H]^+$ also revealed by the molecular interaction network as being strongly correlated with bioactivity (Fig. 3B). Nevertheless, MN was superior in detecting aspergillumarin B. Molecular network was thus evidenced as a promising metabolomics tool for highlighting the active agents in either fractions or crude extracts prior to chromatographic purification.

Fig. 4 ^{1}HNMR spectrum of the pooled active fractions 33-52 from RC-30d extract of *A. flocculus* showing ABC spin system of aspergillumarine analogues

Table 1 Dereplication of metabolites contributing to the activity of antitrypanosome active fractions of *A. flocculus* as revealed from OPLS analysis (Fig.10S). Metabolites were arranged according to their statistical significance (p-values).

Similarly, to pinpoint the bioactive agents in the anticancer active fraction, samples were classified in 2 class groups: active versus inactive. The OPLS-DA score plot (Fig.5A) showed a clear separation between both sample groups. The OPLS score plot explained

ACCEPTED MANUSCRIPT

87% of the total variance (R^2 = 0.87) with the prediction goodness parameter Q^2 = 0.85. The model was validated by permutation test which revealed a permuted Q^2 value of -0.76 which indicated the good prediction ability of the model. Moreover, cross validation ANOVA (CV-ANOVA) revealed P-value 0.003 and F-test value 7.28 which confirmed the significance difference between the active and inactive fractions. The ANOVA F-test was used to assess whether any of the groups is different to the other versus the null hypothesis that all groups yield the same mean response. F-test statistical value was obtained by dividing the between-group variance by the within-group variance. A particularly useful tool that compares the variable magnitude against its reliability is the S-plot obtained by the OPLS-DA model and represented in Fig.5B, where axes plotted from the predictive component are the covariance p[1] against the correlation p(cor)[1]. S-plot displayed metabolites distinctive for the active fractions and highly correlated to their bioactivity. Those metabolites were dereplicated using DNP and their structure are presented in Fig.5. The highly-correlated metabolites at *m/z* (retention time in min) 169.123 $[M+H]^+$ (12.57), 239.139 $[M+H]^+$ (12.59) and 295.227 $[M+H]^+$ (21.94) were selected after the chemotaxonomic filtration and dereplicated as 5,6-dihydro-6-pentyl-2H-pyran-2-one (from *Simplicillium lamellicola*) [39], *N*-[4-(2-formyl-5-hydroxymethyl-pyrrol-1-yl)-butyl]-acetamide (from *Fusarium incarnatum* HKI00504) [40] and tetrahydro-6-(3-hydroxy-4,7-tridecadienyl)-2*H*-pyran-2-one (from *Aspergillus nidulans*) [41]. 5,6-dihydro-6-pentyl-2*H*-pyran-2-one (known as massoia lactone) was reported for its significant cytotoxicity against MALME-3M human Melanoma tumor cell line [42]. Such results suggest for the crucial structural motifs responsible for inducing a target biological activity without further purification. However, to further confirm the exact chemical structure of these differential masses as revealed from the molecular network and MVDA, additional purification was attempted and to further assess its anticancer and antitrypanosomal activities.

Fig. 5 a: OPLS-DA of anticancer fractions (●) versus the inactive (●). b: S-loading plot showing the putatively active molecules (coloured according to mass range)

Bioactivity guided fractionation of the medium up-scaled 30-day rice solid extract (RC-30d) led to 9 antitrypanosomal active fractions whereas two of them namely 50-52 and 53-55 exhibited also anticancer activity (Fig.8S, 9S). The amount of the active fractions

ACCEPTED MANUSCRIPT

was unfortunately not sufficient for further isolation attempt except for fraction (72-89). Hence, fractions F11-17 and F18-29 were pooled into F11-29 and similarly F53-55, 56-64 & 65-71 were gathered into F53-71 based on their TLC pattern. Bioactivity guided isolation of the three main active fractions including F11-29, F53-71 & F72-89 yielded 7 compounds (Fig.7), of which two compounds namely 3-hydroxymellein (**10**) and diorcinol (**11**) demonstrated significant anti-trypanosomal activity (Fig. 6S) with a respective inhibition 56 % and 97%. Diorcinol (**11**) was found active against *T. brucei* with an MIC of 25 µg/mL (108.7 µM). Other analogues of 3-hydroxymellein including botryoisocoumarin A (**7a**) and mellein (**7b**), *cis*-4-hydroxymellein (**8**), and 5-hydroxymellein (**12**) isolated from the active fractions, inhibited growth of K562 cancer cell line at conc. 30 µM as shown in Fig.5S (compounds **7a** and **7b** were isolated as a mixture and biologically tested at conc. 30 µg/mL). The former compounds (mellein's analogues) share the same skeleton with aspergillumarin B, the compound spotted by the molecular network to be highly anticipated to mediate for cytotoxic activity against *T. brucie*. They are also congeners to massoia lactone, the compound detected by MVDA to be contributing to anticancer activity. Such results led to the compounds are likely to be responsible for the biological properties of RC-30d culture extract. Additionally, diorcinol showed not only a cytotoxic effect against K562 cell line but also inhibited NF-κB.

Although, the bioactivity guided isolation, against either cancer cell lines or *T. brucei,* pointed to fraction (90-135) as inactive, isolation attempts were made considering its highest yield (2.7 g) and to identify other chemical structures which would be screened against other microorganisms i.e *Mycobacterium marinum* (unpublished data). Purification of fraction (90-135) led to the isolation of 9 compounds (Fig.7), two of which namely phomaligol A1 (**16**) and dihydropenicillic acid (**17**) possessed a moderate activity against *T. brucei* with MIC of 25 µg/mL (88 µM and 145.3 µM, respectively). Moreover, purification of hexane layer of RC-30d extract yielded three steroidal compounds; compound 5 (ergosterol peroxide) was found strongly active against *T. brucei* with MIC 3.12 µg/mL (7.3 µM) and compound 4 (ergosterol) recorded an MIC of 12.5 µg/mL (31.6 µM). Such results led to deduction that anti-trypanosomal activity could be due to the synergistic effect of diorcinol, mellein derivatives and the active steroidal compounds (Fig.6S). The results pointed to the importance of using more than one method of analysis i.e. GC-MS or NMR along with LC-MS to achieve the best metabolic coverage. Moreover, the study emphasized the significance of bioactivity guided isolation approach

ACCEPTED MANUSCRIPT

as confirmative for the metabolomics results. Even though the MS based metabolomics did not spot the isolated active compounds, the dereplications study of the detected hits revealed very related structures that helped the identification of the isolated molecules. It should be noted that although compound **1**, isolated from the small-scale extract showed no cytotoxic or anti-trypanosomal effect, it exhibited a novel structure first time to be reported in nature which we detail for its spectral assignment in the next section.

Compound 1 (5,9-dihydroxy-2,4,6,8,10-pentamethyldodeca-2,6,10-trienal): isolated as white powder, $[\alpha]^{20}_{D} = -10$ (*c* 0.05 in MeOH), HR ESIMS *m/z* 281.2110 $[M+H]^+$ (calcd. for $C_{17}H_{29}O_3$). Compound 1 is partly similar to TMC-151s (Fig.11S), an antibiotic from *Gliocladium catenulatum* [43]. Apart from the extra hydrocarbon chain and sugar moiety, compound **1** shares the basic polyketide skeleton with TMC-151s as shown in Fig. 6, Table 2. The ^{1}H NMR data (DMSO, 400 MHz) (Table 2, Fig.12S) showed a downfield resonance at δ_H 9.37 indicative of an aldehydic proton H-1 in addition to three protons in the olefinic region at δ_H 6.56 (d, *J* = 9.5 Hz), 5.33 (q, *J* = 6.6 Hz) and 5.25 (d, *J* = 5.2 Hz). Furthermore, two protons in the oxygenated region at δ_H 3.74 (d, *J* = 7.3 Hz) and 3.58 (d, *J* = 7.2 Hz); two protons at δ_H 2.77 (q, *J* = 7.4 Hz) and 2.46 (m). In addition, three methyl doublets at δ_H 0.85 (3H, d, *J* = 6.6 Hz), δ_H 0.73 (3H, d, *J* = 6.8 Hz) and δ_H 1.53 (3H, brs) and three methyl sharp singlets at δ_H 1.65 (3H, s), δ_H 1.55 (3H, s) and δ_H 1.51 (3H, s). Moreover, the presence of two broad exchangeable protons at δ_H 4.78 and 4.26 confirmed the presence of 2 OH groups. The ^{13}C NMR spectrum (400 MHz-DMSO) (Table 2, Fig.13S) showed 17 carbon signals including six methyl groups at δ_C 11.8 (C-12), 9.7 (C-13), 17.4 (C-14), 12.2 (C-15), 18.6 (C-16) and 13.4 (C-17). Three olefinic quaternary carbons were detected at δ_C 138.4 (C-2), 136.2 (C-6) and 138.1 (C-10), furthermore three olefinic methines at δ_C 160.0 (C-3), 131.4 (C-7) and 120.0 (C-11). In addition, two methines at δ_C 37.9 (C-4) and 36.5 (C-8), two oxygenated methine carbons were assigned at δ_C 81.0 (C-5) and 81.5 (C-9) and the carbonyl of the aldehyde group at δ_C 196.0 (C-1). The ^{1}H-^{1}H COSY spectrum (Fig.14S) showed three spin systems, the first connected δ_H 6.56 with 2.77 and methyl doublet at δ_H 0.85 assigned for H-3, H-4 and Me-14 respectively and further confirmed by correlation of H-4 with the oxygenated proton at δ_H 3.74 and the exchangeable proton δ_H 4.78 which can be assigned as H-5 and OH-5 respectively; along with the allylic coupling between H-3 and the sharp methyl singlet δ_H 1.65 for Me-13 (Bold red, Fig.6A). The second spin system connected δ_H 5.25 with 2.46 and methyl doublet at δ_H 0.73 assigned as H-7, H-8 and Me-16 respectively (Bold black, Fig.6A). Similarly, H-8 was further linked with the oxygenated δ_H 3.58 (H-9) and exchangeable OH-9

ACCEPTED MANUSCRIPT

at δ_H 4.26; H-7 correlated with methyl singlet at δ_H 1.55 (Me-15). The third spin system correlated the third methyl doublet at δ_H 1.53 equivalent to Me-12 with the olefinic proton at δ_H 5.33 assigned for H-11. The structure was further confirmed from the HMBC spectrum (Fig.15S) *via* the key cross peak correlations of the allylic protons of H-3 at δ_H 6.57 with carbons (δ_C) C-13 (9.7), C-14 (17.4), C-4 (37.9), C5 (81.0) and C1 (196.0). Furthermore, H-7 at δ_H 5.25 with carbons (δ_C) C-15 (12.2), C-16 (18.6), C-8 (36.5) and C-5 (81.0); H-11 (δ_H 5.33) with carbons at δ_C 13.4 and 81.5 equivalent for C-17 and C-9 respectively. Moreover, the position of OH-5 and OH-9 was confirmed from the correlations of H-5 (δ_H 3.74) with C-15, C-14, C-4, C-7 (δ_C 131.4) and C-3 (δ_C 160.0); H-9 (δ_H 3.58) with C-12 (δ_C 11.8), C-16, C-8, C-11 (δ_C 120.0), C-7 and C-10 (δ_C 138.1) (Table.2). The relative configuration between the hydroxyl group at C-9 and the methyl group (Me-16) at C-8 was established by 1H–1H coupling constant and ROESY correlations at these stereocenters (Fig.16S). The large coupling constant (7.1 Hz) between H-8 and H-9 defined their *anti* orientation. Since H-8 and H-9 were fixed as *anti*, the ROESY couplings shown in (Fig.15S) demonstrated a *syn* relationship of Me-16 and H-9 which consequently defined an *anti* relationship between the methyl group (Me-16) and the hydroxyl group at C-9 (Fig.6B). By a similar logic, the OH group at C-5 has *anti* orientation with the aliphatic methyl group (Me-14). Searching the databases for compound **1** confirmed it was not described before. The closest structure in the literature described by Kohno et al., 1999 [43] (Fig.16S) was a polyketide molecule (TMC-151s) with an extra substitution at C-1 forming an ester linkage which was found up-field at δ_C 167.2 when compared to the aldehydic C-1 (δ_C 196.0) of compound **1**. Consequently, compound **1** (Fig. 6) is designated as a new compound first time to be reported in nature. The structure was further confirmed by comparison of compound **1** to the closely related structure of polyketide TMC-151s (**Table 2**).

Fig. 6 Structure of novel compound 1 isolated from *A. flocculus* showing a: HMBC correlations (arrows) and H-H spin systems (coloured bold bonds) and b: ROESY correlation assigning stereochemistry of OH groups

Table 2 ^{1}H NMR, ^{13}C NMR data and HMBC correlations of compound 1 in comparison to that reported for the reported analogue ''TMC-151s''

Identification of the known compounds

ACCEPTED MANUSCRIPT

Nineteen known compounds (Fig.7) were identified by the spectroscopic data analysis (Supporting information) and comparison with literature values. They were identified as dihydroaspyrone (**2**) [44], kojic acid (**3**) [45], ergosterol (**4**) [46], ergosterol peroxide (**5**) [47], campesterol (**6**) [48], botryoisocoumarin A (**7a**) [49], mellein (**7b**) [50], *cis*-4-hydroxymellein (**8**) [51], *trans*-4-hydroxymellein (**9**) [51], 3-hydroxymellein (**10**) [49], diorcinol (**11**) [52], 5-hydroxymellein (**12**) [53], 7-*O*-acetyl kojic acid (**13**) [54], methyl 2-(4-hydroxyphenyl) acetate (**14**)**,** phomaligol A (**15**) [55], phomaligol A1 (**16**) [55], dihydropenicillic acid (**17**) [56], *p*-hydroxybenzaldehyde (**18**) [57], 2-hydroxyphenyl acetic acid (**19a**) [58], 4-hydroxyphenyl acetic acid (**19b**) [59] and 4,5-dihydroxymellein (**20**) [60].

Fig. 7 Structures of the isolated known compounds (2-20)

4. Conclusion

In conclusion, results provided a realistic, compound-based rational for the anticancer and anti-trypanosomal principles present in ***Aspergillus flocculus***. Additionally, they pointed to an additional evidence for the efficacy and complementarity of LC-MS metabolites profiling when coupled to bioassays in the field of natural product based drug discovery, to speed up the traditional lengthy processes of identifying an active principle by consecutive isolation from crude extracts. The emerging spectroscopic and informatics technologies can indeed help to narrow the gap that has opened to modern synthesis based drug discovery and with the pressure towards a shorter time in the discovery of bioactive agents from natural resources. Multivariate data analysis viz. PCA revealed the most unique fungal extract in terms of chemical composition. Moreover, the study presented an efficient use of the similarity molecular interaction network along with OPLS-DA to highlight the metabolites strongly correlated to the featured bioactivity prior to purification attempts. Dereplication studies, based on the chemotaxnomic sorting proposed the putative active agents whereas structural assignment of isolated compounds, employed both HR-MS and NMR, confirmed the identified hits. Although, it should be noted that two of the most active antitrypansomal agents (ergosterol and ergosterol peroxide) were not revealed from HPLC-MS analysis which suggest that in case of anti-trypanosomal effect, activity might be due to synergism of more than one component i.e mellein derivatives, diorcinol and compounds undetected using LC-MS. We forecast that the implementation of other universal metabolites profiling

ACCEPTED MANUSCRIPT

technologies namely 1D- and 2D-NMR applied on fungal crude extracts and fractions could provide better coverage of investigated metabolome and result in a gradual shift from the common isolation procedure. The application of metabolomics in drug discovery from endophytes has also yet to be fully examined compared to that reported *in planta*.

Supporting information

1D, 2D NMR spectra of compound **1**, ^{1}H, ^{13}CNMR data of the known compounds and bioactivity charts are available as supporting information.

Acknowledgements

Financial support for Dr Ahmed Tawfike by sector of missions, Ministry of Higher Education, Egypt, is greatly appreciated. Thanks are due to Dr Rothwelle Tate, Strathclyde institute of Pharmacy and Biomedical science for his guidance through the molecular biological procedures used for the phylogenetic identification of endophytic fungi**.**

Conflict of Interest

The authors declare no conflict of interest.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

References

[1] D.J. Newman, Cragg, G.M., Snader, K.M, Natural products as sources of new drugs over the period 1981–2002, J. Nat. Prod. 66 (2003) 1022-1037.
[2] A. Arnold, Understanding the diversity of foliar endophytic fungi: Progress, challenges, and frontiers, Fungal. Biol. Rev. 21 (2007) 51-66.
[3] M.B. Ibrahim, N. Kaushik, A.A. Sowemimo, O.A. Odukoya, Review of the phytochemical and pharmacological studies of the genus *Markhamia*, Pharmacogn. Rev. 10 (2016) 50-59.
[4] F. Tantangmo, B.N. Lenta, F.F. Boyom, S. Ngouela, M. Kaiser, E. Tsamo, B. Weniger, P.J. Rosenthal, C. Vonthron-Senecheau, Antiprotozoal activities of some constituents of *Markhamia tomentosa* (Bignoniaceae), Ann. Trop. Med. Parasitol. 104 (2010) 391-398.
[5] F. Epifano, S. Genovese, S. Fiorito, V. Mathieu, R. Kiss, Lapachol and its congeners as anticancer agents: a review, Phytochem. Rev. 13 (2014) 37-49.
[6] D. Lacroix, S. Prado, A. Deville, S. Krief, V. Dumontet, J. Kasenene, E. Mouray, C. Bories, B. Bodo, Hydroperoxy-cycloartane triterpenoids from the leaves of *Markhamia lutea*, a plant ingested by wild chimpanzees, Phytochemistry 70 (2009) 1239-1245.
[7] R.A. El Dib, A.H. Gaara, S.M. El-Shenawy, J.A. Micky, A.A. Mohammed, M.S. Marzouk, Leaf extract of *Markhamia platycalyx*: polyphenolic profile, acute toxicity, anti-inflammatory, hepatoprotective and in vitro antioxidant activities, Drug Res. 64 (2014) 680-689.
[8] M.R. Kernan, A. Amarquaye, J.L. Chen, J. Chan, D.F. Sesin, N. Parkinson, Z. Ye, M. Barrett, C. Bales, C.A. Stoddart, B. Sloan, P. Blanc, C. Limbach, S. Mrisho, E.J. Rozhon, Antiviral phenylpropanoid glycosides from the medicinal plant *Markhamia lutea*, J. Nat. Prod. 61 (1998) 564-570.
[9] A.F. Tawfike, R. Tate, G. Abbott, L. Young, C. Viegelmann, M. Schumacher, M. Diederich, R.A. Edrada-Ebel, Metabolomic tools to assess the chemistry and bioactivity of endophytic *Aspergillus* Strain, Chem. Biodivers. 14 (2017).
[10] C.R. UK, Prostate cancer statstics available at http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer#heading-One (Accessed November 29, 2018).
[11] T.L.L.S. (LLS), Facts and Statistics available at http://www.lls.org/http%3A/llsorg.prod.acquia-sites.com/facts-and-statistics/facts-and-statistics-overview/facts-and-statistics, (Accessed November 29, 2018).
[12] P.P. Simarro, J. Jannin, P. Cattand, Eliminating human African trypanosomiasis: Where do we stand and what comes next?, PLoS Med. 5 (2008) e55.
[13] R. Brun, J. Blum, F. Chappuis, C. Burri, Human African trypanosomiasis, The Lancet 375 (2010) 148-159.
[14] W.H. Organization, Trypanosomiasis, human African (sleeping sickness) available at http://www.who.int/mediacentre/factsheets/fs259/en/, (Accessed November 29, 2018).
[15] M.P. Barrett, The rise and fall of sleeping sickness, The Lancet 367 (2006) 1377-1378.
[16] D.J. Newman, G.M. Cragg, Natural products as sources of new drugs over the 30 years from 1981 to 2010, J. Nat. Prod. 75 (2012) 311-335.
[17] A. Tawfike, C. Viegelmann, R. Edrada-Ebel, Metabolomics and dereplication strategies in natural products, in: U. Roessner, D.A. Dias (Eds.) Metabolomics tools for natural product discovery, Humana Press, 2013, pp. 227-244.
[18] J. Kjer, A. Debbab, A.H. Aly, P. Proksch, Methods for isolation of marine-derived endophytic fungi and their bioactive secondary products, Nat Protoc. 5 (2010) 479-490.
[19] A. Debbab, A.H. Aly, R. Edrada-Ebel, W.E.G. Mueller, M. Mosaddak, A. Hakikj, R. Ebel, P. Proksch, Bioactive secondary metabolites from the endophytic fungus *Chaetomium*

ACCEPTED MANUSCRIPT

sp. isolated from *Salvia officinalis* growing in Morocco, Biotechnology, Agronomy, Society and Environment (BASE), 13 (2009) 229-234.
[20] M.C. Chambers, B. Maclean, R. Burke, D. Amodei, D.L. Ruderman, S. Neumann, L. Gatto, B. Fischer, B. Pratt, J. Egertson, K. Hoff, D. Kessner, N. Tasman, N. Shulman, B. Frewen, T.A. Baker, M.-Y. Brusniak, C. Paulse, D. Creasy, L. Flashner, K. Kani, C. Moulding, S.L. Seymour, L.M. Nuwaysir, B. Lefebvre, F. Kuhlmann, J. Roark, P. Rainer, S. Detlev, T. Hemenway, A. Huhmer, J. Langridge, B. Connolly, T. Chadick, K. Holly, J. Eckels, E.W. Deutsch, R.L. Moritz, J.E. Katz, D.B. Agus, M. MacCoss, D.L. Tabb, P. Mallick, A cross-platform toolkit for mass spectrometry and proteomics, Nat. Biotech. 30 (2012) 918-920.
[21] L. Macintyre, T. Zhang, C. Viegelmann, I.J. Martinez, C. Cheng, C. Dowdells, R. Edrada-Ebel, U.R. Abdelmohsen, C. Gernert, U. Hentschel, Metabolomic tools for secondary metabolite discovery from marine microbial symbionts, Mar. drugs 12 (2014) 3416-3448.
[22] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: A software environment for integrated models of biomolecular interaction networks, Genome Res. 13 (2003) 2498-2504.
[23] B. Räz, M. Iten, Y. Grether-Bühler, R. Kaminsky, R. Brun, The Alamar Blue® assay to determine drug sensitivity of African trypanosomes (*T.b. rhodesiense* and *T.b. gambiense*) in vitro, Acta Trop. 68 (1997) 139-147.
[24] M. Schumacher, C. Cerella, S. Eifes, S. Chateauvieux, F. Morceau, M. Jaspars, M. Dicato, M. Diederich, Heteronemin, a spongean sesterterpene, inhibits TNFα-induced NF-κB activation through proteasome inhibition and induces apoptotic cell death, Biochem. Pharmacol. 79 (2010) 610-622.
[25] A. Duvoix, S. Delhalle, R. Blasius, M. Schnekenburger, F. Morceau, M. Fougère, E. Henry, M.M. Galteau, M. Dicato, M. Diederich, Effect of chemopreventive agents on glutathione S-transferase P1-1 gene expression mechanisms via activating protein 1 and nuclear factor kappaB inhibition, Biochem. Pharmacol. 68 (2004) 1101-1111.
[26] K. Purves, L. Macintyre, D. Brennan, G.O. Hreggviethsson, E. Kuttner, M.E. Asgeirsdottir, L.C. Young, D.H. Green, R. Edrada-Ebel, K.R. Duncan, Using molecular networking for microbial secondary metabolite bioprospecting, Metabolites 6 (2016).
[27] M.A. Farag, S.H. El-Ahmady, F.S. Elian, L.A. Wessjohann, Metabolomics driven analysis of artichoke leaf and its commercial products via UHPLC-q-TOF-MS and chemometrics, Phytochemistry 95 (2013) 177-187.
[28] M.A. Farag, Comparative mass spectrometry & nuclear magnetic resonance metabolomic approaches for nutraceuticals quality control analysis: a brief review, Recent patents on biotechnology 8 (2014) 17-24.
[29] S. Toda, S. Yamamoto, O. Tenmyo, T. Tsuno, T. Hasegawa, M. Rosser, M. Oka, Y. Sawada, M. Konishi, T. Oki, A new neuritogenetic compound BU-4514N produced by *Microtetraspora* sp, J. Antibiot. 46 (1993) 875-883.
[30] K. Kito, R. Ookura, S. Yoshida, M. Namikoshi, T. Ooi, T. Kusumi, Pentaketides relating to aspinonene and dihydroaspyrone from a marine-derived fungus, *aspergillus ostianus*, J. Nat. Prod. 70 (2007) 2022-2025.
[31] X.W. Chen, C.W. Li, C.B. Cui, W. Hua, T.J. Zhu, Q.Q. Gu, Nine new and five known polyketides derived from a deep sea-sourced *Aspergillus* sp. 16-02-1, Marine drugs, 12 (2014) 3116-3137.
[32] Y. Liu, X.-M. Li, L.-H. Meng, B.-G. Wang, Polyketides from the marine mangrove-derived fungus *Aspergillus ochraceus* MA-15 and their activity against aquatic pathogenic bacteria, Phytochem. Lett. 12 (2015) 232-236.

ACCEPTED MANUSCRIPT

[33] M. Okue, H. Watanabe, K. Kasahara, M. Yoshida, S. Horinouchi, T. Kitahara, Short-step syntheses of all stereoisomers of preussin and their bioactivities, Biosci. Biotechnol. Biochem. 66 (2002) 1093-1096.
[34] C.H. Chao, K.J. Chou, G.H. Wang, Y.C. Wu, L.H. Wang, J.P. Chen, J.H. Sheu, P.J. Sung, Norterpenoids and related peroxides from the formosan marine sponge *Negombata corticata*, J. Nat. Prod. 73 (2010) 1538-1543.
[35] T.D. Gilmore, Introduction to NF-kappaB: players, pathways, perspectives, Oncogene 25 (2006) 6680-6684.
[36] R.O. Escarcega, S. Fuentes-Alexandro, M. Garcia-Carrasco, A. Gatica, A. Zamora, The transcription factor nuclear factor-kappa B and cancer, Clin. Oncol. (R. Coll. Radiol.) 19 (2007) 154-161.
[37] C. Monaco, E. Andreakos, S. Kiriakidis, C. Mauri, C. Bicknell, B. Foxwell, N. Cheshire, E. Paleolog, M. Feldmann, Canonical pathway of nuclear factor kappa B activation selectively regulates proinflammatory and prothrombotic responses in human atherosclerosis, Proc. Natl. Acad. Sci. U.S.A. 101 (2004) 5634-5639.
[38] J. Qi, C.-L. Shao, Z.-Y. Li, L.-S. Gan, X.-M. Fu, W.-T. Bian, H.-Y. Zhao, C.-Y. Wang, Isocoumarin derivatives and benzofurans from a sponge-derived *Penicillium* sp. fungus, J. Nat. Prod. 76 (2013) 571-579.
[39] Q. Le Dang, T.S. Shin, M.S. Park, Y.H. Choi, G.J. Choi, K.S. Jang, I.S. Kim, J.C. Kim, Antimicrobial activities of novel mannosyl lipids isolated from the biocontrol fungus *Simplicillium lamellicola* BCP against phytopathogenic bacteria, J. Agric. Food Chem. 62 (2014) 3363-3370.
[40] L.Y. Li, Y. Ding, I. Groth, K.D. Menzel, G. Peschel, K. Voigt, Z.W. Deng, I. Sattler, W.H. Lin, Pyrrole and indole alkaloids from an endophytic *Fusarium incarnatum* (HKI00504) isolated from the mangrove plant *Aegiceras corniculatum*, J. Asian Nat. Prod. Res. 10 (2008) 775-780.
[41] P. Mazur, H.V. Meyers, K. Nakanishi, A.E.E.-Z. A, S.P. Champe, Structural elucidation of sporogenic fatty acid metabolites from *aspergillus nidulans*, Tetrahedron Lett. 31 (1990) 3837-3840.
[42] N.C.f.B. Information, PubChem Substance Database; SID=535786, https://pubchem.ncbi.nlm.nih.gov/substance/535786, (Accessed November 29, 2018).
[43] J. Kohno, M. Nishio, M. Sakurai, K. Kawano, H. Hiramatsu, N. Kameda, N. Kishi, T. Yamashita, T. Okuda, S. Komatsubara, Isolation and structure determination of TMC-151s: Novel polyketide antibiotics from *Gliocladium catenulatum* Gilman & Abbott TC 1280, Tetrahedron 55 (1999) 7771-7786.
[44] T. Sassa, S. Hayakari, M. Ikeda, Y. Miura, Plant growth inhibitors produced by fungi. I. Isolation and identification of penicillic acid and dihydropenicillic acid, Agr. Biol. Chem. 35 (1971) 2130-2131.
[45] X. Li, J.H. Jeong, K.T. Lee, J.R. Rho, H.D. Choi, J.S. Kang, B.W. Son, γ-pyrone derivatives, kojic acid methyl ethers from a marine-derived fungus *Altenaria* sp, Arch. Pharmacal Res. 26 (2003) 532-534.
[46] V. Chobot, L. Opletal, L. Jáhodář, A.V. Patel, C.G. Dacke, G. Blunden, Ergosta-4,6,8,22-tetraen-3-one from the edible fungus, *Pleurotus ostreatus* (oyster fungus), Phytochemistry 45 (1997) 1669-1671.
[47] D.Y. Lee, S.J. Lee, H.Y. Kwak, J. Lakoon, J. Heo, S. Hong, G.-W. Kim, N.I. Baek, Sterols isolated from Nuruk (*Rhizopus oryzae* KSD-815) inhibit the migration of cancer cells, J. Microbiol. Biotechnol. 19 (2009) 1328-1332.
[48] X. Zhang, A. Cambrai, M. Miesch, S. Roussi, F. Raul, D. Aoude-Werner, E. Marchioni, Separation of Δ5- and Δ7-phytosterols by adsorption chromatography and semipreparative

ACCEPTED MANUSCRIPT

reversed phase high-performance liquid chromatography for quantitative analysis of phytosterols in foods, J. Agric. Food Chem. 54 (2006) 1196-1202.
[49] Y. Xu, C. Lu, Z. Zheng, A new 3,4-dihydroisocoumarin isolated from *Botryosphaeria* sp. F00741, Chem. Nat. Compd. 48 (2012) 205-207.
[50] C. Dimitriadis, M. Gill, M.F. Harte, The first stereospecific approach to both enantiomers of mellein, Tetrahedron: Asymmetry 8 (1997) 2153-2158.
[51] K.N. Asha, R. Chowdhury, C.M. Hasan, M.A. Rashid, Steroids and polyketides from *Uvaria hamiltonii* stem bark, Acta Pharm. (Zagreb, Croatia) 54 (2004) 57-63.
[52] L.J. Fremlin, A.M. Piggott, E. Lacey, R.J. Capon, Cottoquinazoline A and cotteslosins A and B, metabolites from an Australian marine-derived strain of *Aspergillus versicolor*, J. Nat. Prod. 72 (2009) 666-670.
[53] F.D. da Silva Araújo, L.C. de Lima Fávaro, W.L. Araújo, F.L. de Oliveira, R. Aparicio, A.J. Marsaioli, Epicolactone – Natural Product Isolated from the sugarcane endophytic fungus *Epicoccum nigrum*, Eur. J. Org. Chem. 2012 (2012) 5225-5230.
[54] L. Kremnicky, V. Mastihuba, G.L. Cote, *Trichoderma reesei* acetyl esterase catalyzes transesterification in water, J. Mol. Catal. B: Enzym. 30 (2004) 229-239.
[55] M. Elbandy, P.B. Shinde, J. Hong, K.S. Bae, M.A. Kim, S.M. Lee, J.H. Jung, α-pyrones and yellow pigments from the sponge-derived fungus *Paecilomyces lilacinus*, Bull. Korean Chem. Soc. 30 (2009) 188-192.
[56] T. Sassa, S. Hayakari, M. Ikeda, Y. Miura, Plant growth inhibitors produced by fungi. I. Isolation and identification of penicillic acid and dihydropenicillic acid, Agr. Biol. Chem. 35 (1971) 2130-2131.
[57] H. Kim, J. Ralph, F. Lu, S.A. Ralph, A.M. Boudet, J.J. MacKay, R.R. Sederoff, T. Ito, S. Kawai, H. Ohashi, T. Higuchi, NMR analysis of lignins in CAD-deficient plants. Part 1. Incorporation of hydroxycinnamaldehydes and hydroxybenzaldehydes into lignins, Org. Biomol. Chem. 1 (2003) 268-281.
[58] G.B. Zhou, P.F. Zhang, Y.J. Pan, A novel method for synthesis of arylacetic acids from aldehydes, *N*-(2,3,4,6-tetra-*O*-pivaloylated-d-glucopyranosyl)amine and trimethylsilylcyanide, Tetrahedron 61 (2005) 5671-5677.
[59] J.E. Milne, T. Storz, J.T. Colyer, O.R. Thiel, S.M. Dilmeghani, R.D. Larsen, J.A. Murry, Iodide-catalyzed reductions: Development of a synthesis of phenylacetic acids, J. Org. Chem. 76 (2011) 9519-9524.
[60] X.Y. Jian, Y. Chen, C. Huang, Z. She, Y. Lin, A new isochroman derivative from the marine fungus *Phomopsis* sp. (No. ZH-111), Chem. Nat. Compd. 47 (2011) 13-16.

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 3: **Dereplication of metabolites contributing to the activity of antitrypanosome active fractions of *A. flocculus* as revealed from OPLS analysis (Fig.10S). Metabolites were arranged according to their statistical significance p-values.**

m/z	**Retention time**	**Molecular weight**	**Name/formula**	**Biological source**	**Probability**
955.757 $[M+H]^+$	31.89	954.749	Unknown $C_{55}H_{98}O_7N_6$		1.12E-06
929.742 $[M+H]^+$	31.64	928.735	Unknown $C_{53}H_{96}O_7N_6$		2.25E-05
904.761 $[M+H]^+$	31.92	903.754	Unknown $C_{57}H_{95}N_9$		0.000474
909.717 $[M+H]^+$	31.54	908.71	Unknown $C_{55}H_{94}O_7N_3$		0.001501
935.732 $[M+H]^+$	29.79	934.725	Unknown $C_{57}H_{96}O_7N_3$		0.002307
293.212 $[M-H]^-$	20.86	294.219	Tetrahydro-6-(3-hydroxy-4,7-tridecadienyl)-2*H*-pyran-2-one $C_{18}H_{30}O_3$	*Aspergillus nidulans*	0.002518
267.159 $[M+H]^+$	17.06	266.152	Aspterric acid / Avenaciolide / Daldinin C-1"-Deoxy, 2"-hydroxyl $C_{15}H_{22}O_4$	*Aspergillus terreus, A. avenaceus, A. ustus 094102*	0.022793
309.207 $[M-H]^-$	19.31	310.214	Cephalosporolide H / Lachnelluloic acid $C_{18}H_{30}O_4$	*Penicillium sp., Lachnellula fuscosanguinea*	0.025089
169.123 $[M+H]^+$	15.57	168.115	5,6-Dihydro-6-pentyl-2*H*-pyran-2-one $C_{10}H_{16}O_2$	*Simplicillium lamellicola* BCP	0.042953
283.154 $[M+H]^+$	15.92	282.147	1,7-Dihydroxy-1,3,5-bisabolatrien-15-oic acid, 11-hydroxy $C_{15}H_{22}O_5$	*Aspergillus sydowii*	0.046196

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Table 4: **^{1}H NMR, ^{13}C NMR data and HMBC correlations of compound 1 in comparison to that reported for its closest analogue ''TMC-151s''**

Atom No.	Compound 1-*d*-DMSO			TMC-151s-*d*-DMSO (Kohno 1999)	
	δ_H (m, *J* in Hz)	δ_C (m)	HMBC correlations	δ_H (m, *J* in Hz)	δ_C (m)
1	9.37 (s)	196.0 (CH)	C-13, C-2, C-3		167.2 (C)
2		138.4 (C)			126.5 (C)
3	6.57 (d, 7.7 Hz)	160.0 (CH)	C-13, C-14, C-4, C-5, C-1	6.70 (d, 10.1 Hz)	146.8 (CH)
4	2.77 (q, 7.7 Hz)	37.9 (CH)	C-14, C-5, C-3, C-2	2.56 (m)	36.9 (CH)
5	3.74 (d, 7.7 Hz)	81.0 (CH)	C-15, C-14, C-4, C-7, C-3	3.67 (m)	80.9 (CH)
6		136.1 (C)			136.0 (C)
7	5.25 (d, 9.0 Hz)	131.4 (CH)	C-15, C-16, C-8, C-5	5.19 (br. d, 9.0 Hz)	131.6 (CH)
8	2.46 (m)	36.5 (CH)	C-16, C-9, C-7, C-6	2.45 (m)	36.1 (CH)
9	3.58 (d, 7.1 Hz)	81.5 (CH)	C-12, C-16, C-8, C-11, C-7, C-10	3.55 (dd, 8.5, 3.4 Hz)	81.2 (CH)
10		138.1 (C)			134.5 (C)
11	5.33 (q, 6.6 Hz)	120.0 (CH)	C-17, C-9	5.39 (br.d, 9.0 Hz)	130.1 (CH)
12	1.53 (br. s)	11.8 (CH_3)			
13	1.65 (s)	9.7 (CH_3)	C-2, C-3, C-1	1.80 (d, 1.0 Hz)	12.6 (CH_3)
14	0.85 (d, 7.7 Hz)	17.4 (CH_3)	C-4, C-5, C-3	0.76 (d, 6.8 Hz)	16.4 (CH_3)
15	1.55 (s)	12.2 (CH_3)	C-5, C-7	1.56 (br. s)	11.0 (CH_3)
16	0.73 (d, 6.8 Hz)	18.6 (CH_3)	C-8, C-9, C-7	0.70 (d, 6.8 Hz)	17.4 (CH_3)
17	1.51 (s)	13.4 (CH_3)	C-10, C-11, C-9	1.55 (br. s)	11.3 (CH_3)

ACCEPTED MANUSCRIPT

ACCEPTED MANUSCRIPT

Highlights

1. A combination of bioactivity guided-isolation and MS based-metabolomics was adopted for the target analysis of bioactive secondary metabolites
2. Multivariate data analysis and molecular correlation networks were implemented for samples classification and detection of bioactive agents
3. A new polyketide molecule was isolated
4. Antitrypanosomal and anticancer activities of the purified compounds were assessed

ACCEPTED MANUSCRIPT

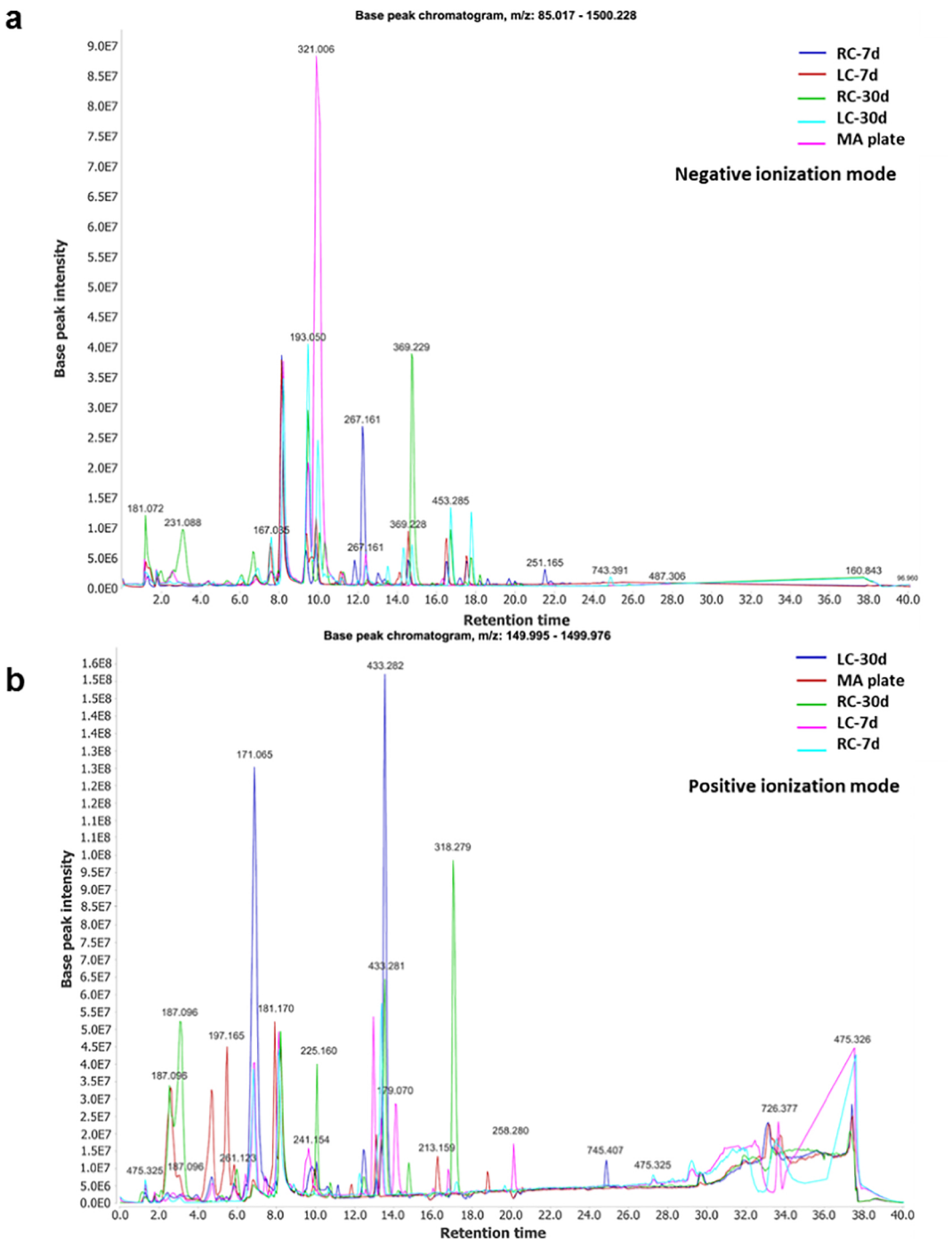

a
Base peak chromatogram, m/z: 85.017 - 1500.228
RC-7d
LC-7d
RC-30d
LC-30d
MA plate
Negative ionization mode
Base peak intensity
Retention time
321.006
193.050
369.229
267.161
453.285
181.072
231.088
167.035
369.228
267.161
251.165
743.391
487.306
160.843
b
Base peak chromatogram, m/z: 149.995 - 1499.976
LC-30d
MA plate
RC-30d
LC-7d
RC-7d
Positive ionization mode
Base peak intensity
Retention time
433.282
171.065
318.279
433.281
187.096
181.170
197.165
225.160
187.096
179.070
726.377
475.326
241.154
258.280
213.159
745.407
475.325
475.325
187.096
261.123

Figure 1

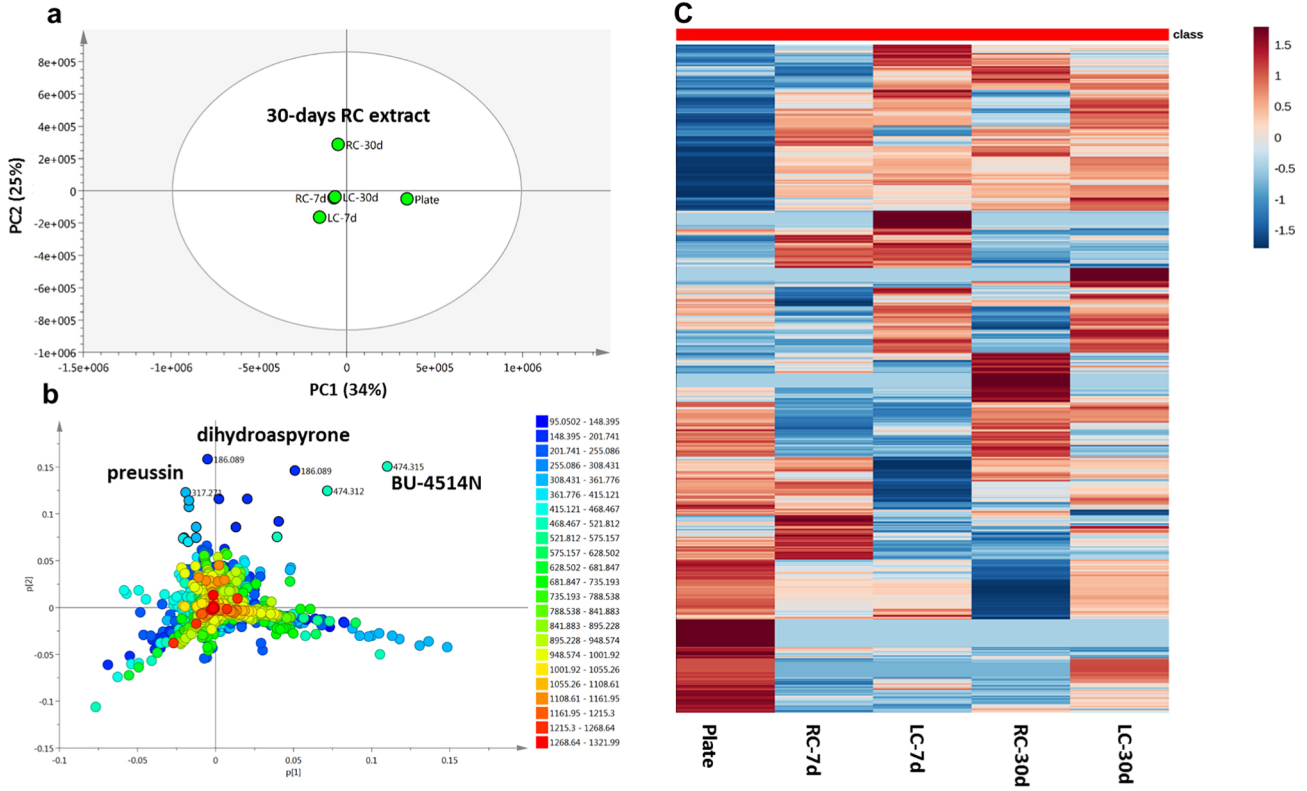

a
30-days RC extract
RC-30d
RC-7d
LC-30d
Plate
LC-7d
PC2 (25%)
PC1 (34%)
b
dihydroaspyrone
preussin
BU-4514N
186.089
186.089
474.315
317.221
474.312
p[2]
p[1]
95.0502 - 148.395
148.395 - 201.741
201.741 - 255.086
255.086 - 308.431
308.431 - 361.776
361.776 - 415.121
415.121 - 468.467
468.467 - 521.812
521.812 - 575.157
575.157 - 628.502
628.502 - 681.847
681.847 - 735.193
735.193 - 788.538
788.538 - 841.883
841.883 - 895.228
895.228 - 948.574
948.574 - 1001.92
1001.92 - 1055.26
1055.26 - 1108.61
1108.61 - 1161.95
1161.95 - 1215.3
1215.3 - 1268.64
1268.64 - 1321.99
C
class
Plate
RC-7d
LC-7d
RC-30d
LC-30d

Figure 2

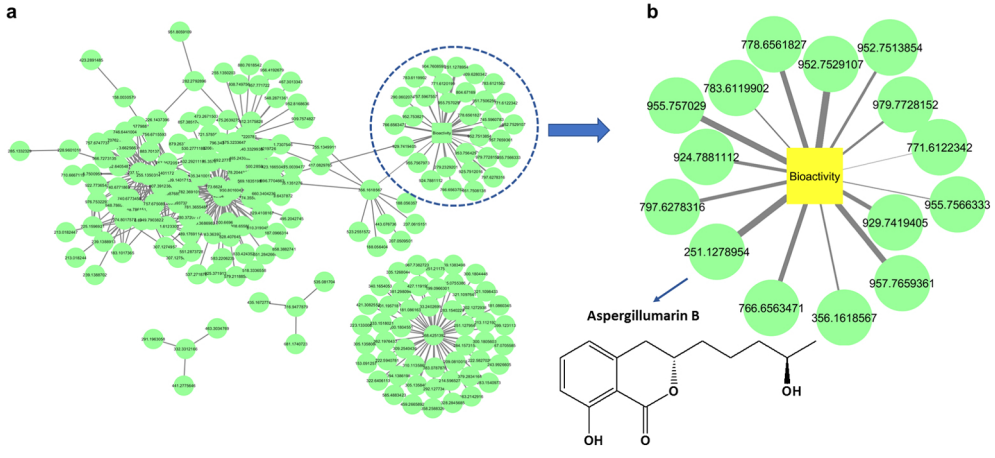

a
b
Bioactivity
778.6561827
952.7513854
952.7529107
783.6119902
955.757029
979.7728152
924.7881112
771.6122342
797.6278316
955.7566333
929.7419405
251.1278954
957.7659361
766.6563471
356.1618567
Aspergillumarin B
OH
O
OH
O

Figure 3

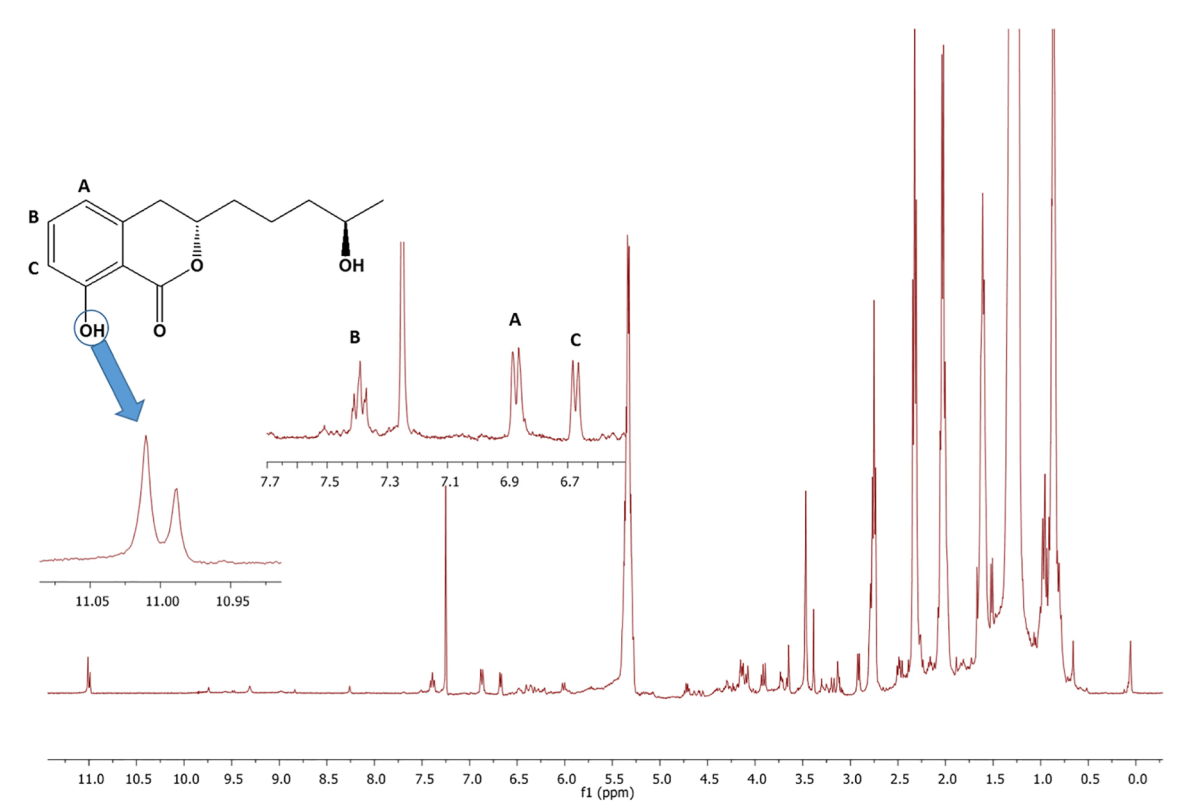
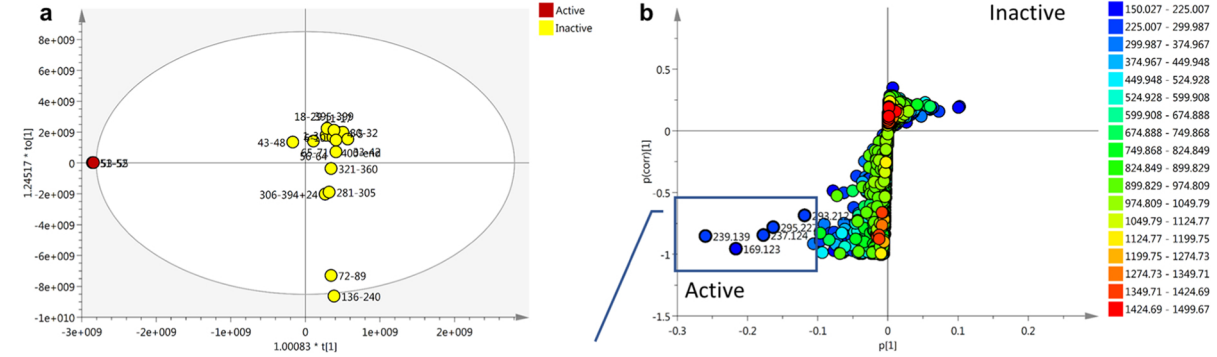
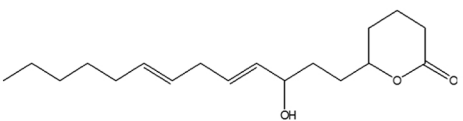
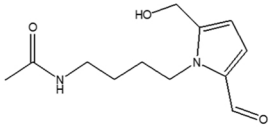

A
B
C
O
OH
O
OH
B
A
C
7.7
7.5
7.3
7.1
6.9
6.7
11.05
11.00
10.95
11.0
10.5
10.0
9.5
9.0
8.5
8.0
7.5
7.0
6.5
6.0
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
f1 (ppm)

Figure 4

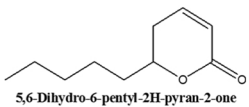




Tetrahydro-6-(3-hydroxy-4,7-tridecadienyl)-2H-pyran-2-one
Chemical Formula: $C_{18}H_{30}O_3$
Exact Mass: 294.2195
Source: Metab. of *Aspergillus nidulans*



N-[4-(2-formyl-5-hydroxymethyl-pyrrol-1-yl)-butyl]-acetamide
Chemical Formula: $C_{12}H_{18}N_2O_3$
Exact Mass: 238.1317
Source: The mangrove-derived *Fusarium incarnatum* HKI00504



5,6-Dihydro-6-pentyl-2H-pyran-2-one
Chemical Formula: $C_{10}H_{16}O_2$
Exact Mass: 168.1150
Source: *Simplicillium lamellicola*

Figure 5

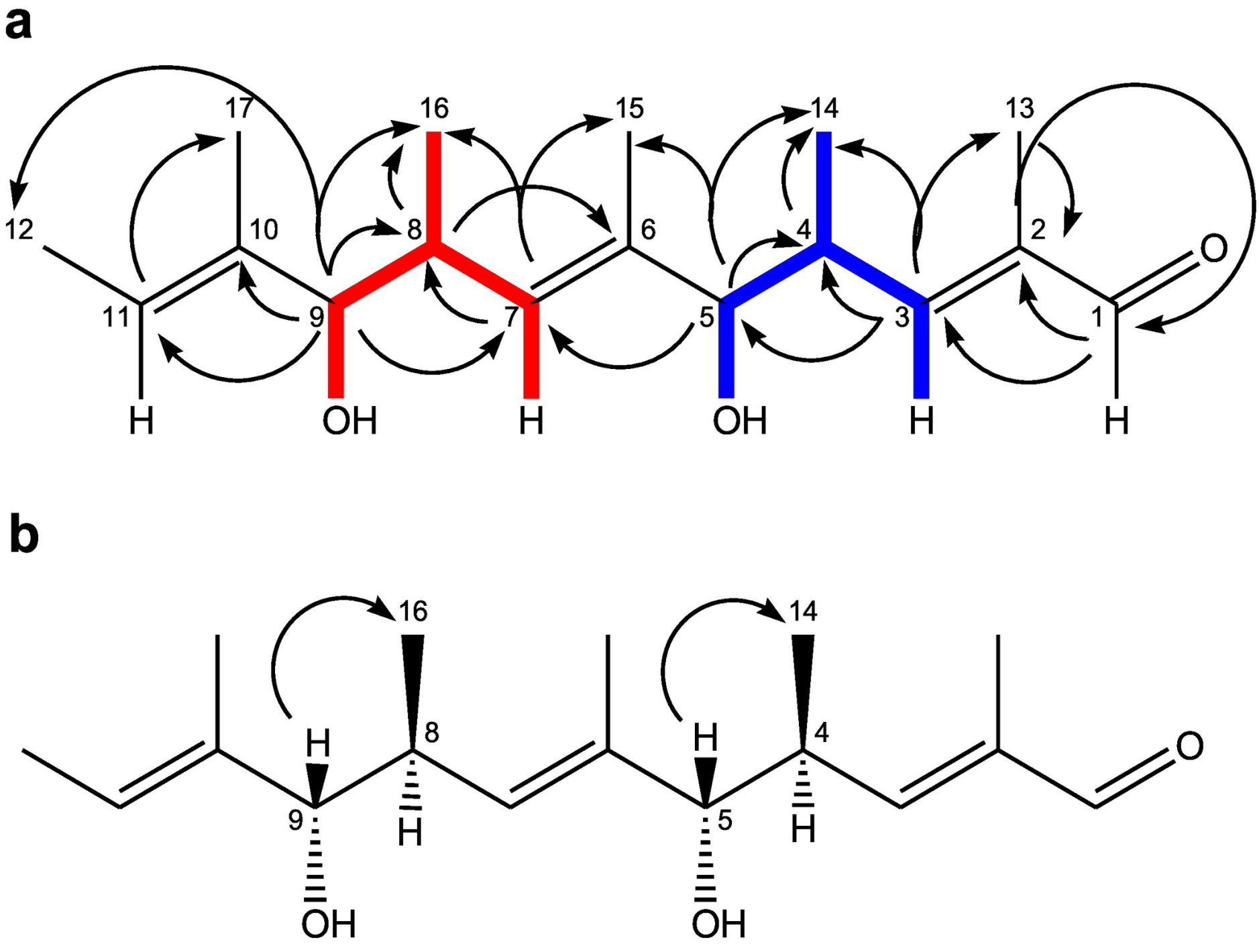
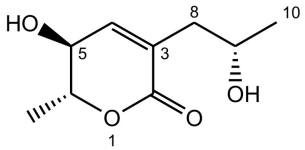
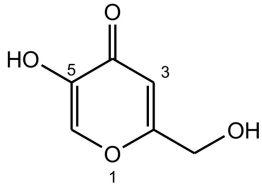

a
17
16
15
14
13
12
10
8
6
4
2
O
11
9
7
5
3
1
H
OH
H
OH
H
H
b
16
14
H
8
H
4
O
9
H
5
H
OH
OH

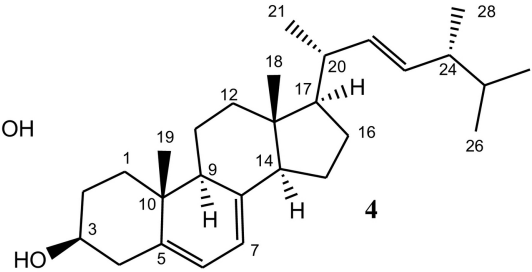
Figure 6



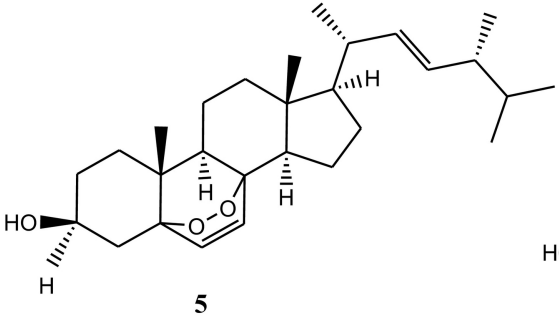
2



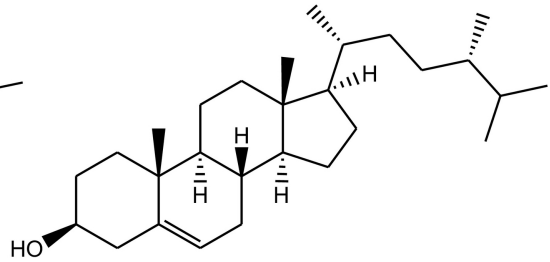
3



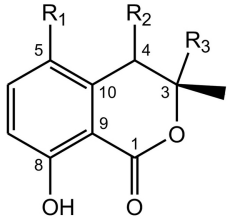
4



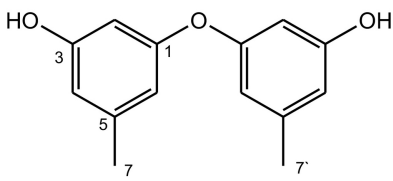
5



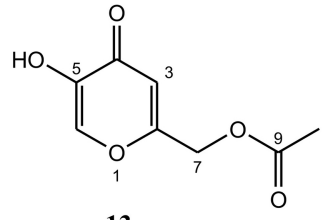
6



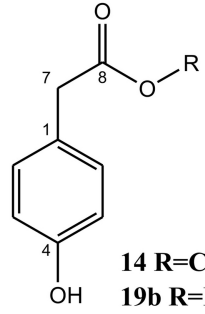
7a R_1=R_2=H, R_3=OCH_3
7b R_1=R_2=R_3=H
8 R_1=R_3=H, R_2=OH(R)
9 R_1=R_3=H, R_2=OH(S)
10 R_1=R_2=H, R_3=OH
12 R_1= OH, R_2=R_3=H
20 R_1=R_2=OH, R_3=H



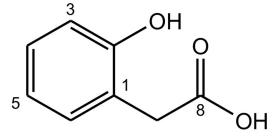
11



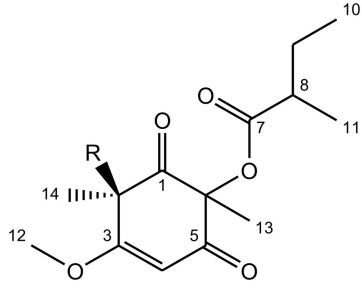
13



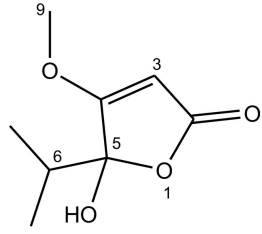
14 R=CH_3
19b R=H



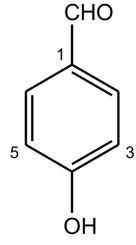
19a



15 R=OH(R)
16 R=OH(S)



17



18

Figure 7