Metabolomic-guided isolation of bioactive natural products from *Curvularia* sp., an endophytic fungus of *Terminalia laxiflora*

Ahmed F. Tawfike[1,2], Grainne Abbott[1], Louise Young[1], and RuAngelie Edrada-Ebel[1]

Affiliation

[1]Strathclyde Institute of Pharmacy and Biomedical Science, University of Strathclyde, United Kingdom

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

Correspondence

Dr. RuAngelie Edrada-Ebel, Strathclyde Institute of Pharmacy and Biomedical Science, University of Strathclyde, 161 Cathedral Street, G4 0RE, Glasgow, United Kingdom. E-mail: ruangelie.edrada-ebel@strath.ac.uk Tel: +44(0)141 548 5968 Fax:+44(0)141 552 2562

Ahmed F. Tawfike, Department of Pharmacognosy, Faculty of Pharmacy, Helwan University, 11795, Cairo, Egypt. E-mail: ahmedtawfike@gmail.com, Tel: +44(0)7875190840

Abstract

Endophytic fungi associated with medicinal plants are a potential source of novel chemistry and biology. Metabolomic tools were successfully employed to compare the metabolite fingerprints of solid and liquid culture extracts of endophyte *Curvularia* sp. isolated from the leaves of *Terminalia laxiflora*. Natural product databases were used to dereplicate metabolites in order to determine known compounds and the presence of new natural products. Multivariate analysis highlighted the putative metabolites responsible for the bioactivity of the fungal extract and its fractions on NF-kappaB and the myelogenous leukemia cell line K562. Metabolomic tools and dereplication studies using HRESIMS directed the fractionation and isolation of the bioactive components from the fungal extracts. This resulted in the isolation of *N*-acetylphenylalanine (**1**) and two linear peptide congeners of **1**: dipeptide *N*-acetylphenylalanyl-L-phenylalanine (**2**) and tripeptide *N*-acetylphenylalanyl-L-phenylalanyl-L-leucine (**3**).

Keywords:

Metabolomics; Dereplication; Endophytic fungi; *Curvularia; Terminalia laxiflora,* Combretaceae

No. of words excluding References: 4000

Introduction

In our search for new potential anticancer agents from endophytic fungi of Egyptian medicinal plants, *Terminalia laxiflora* (Combretaceae) was among the plants chosen for this study. In Egypt, *T. laxiflora* has been used to treat yaws, diarrhea, dysentery, and pulmonary troubles. Extracts and isolated compounds from plants of the genus *Terminalia* have been reported not only for their cytotoxicity [1-3], antioxidant and free radical scavenging effects [4-6] but as well as induction of cell cycle arrest and apoptosis [7-9]. Chebulagic acid from *T. chebula* was described to synergize the cytotoxicity of doxorubicin in human hepatocellular carcinoma [10]. *T. catappa* exhibited antimetastatic effects on hepatocellular carcinoma by modulating NF-kappaB and AP-1 activity [11]. Moreover, like the other members of the genus, *T. laxiflora* is rich with hydrolysable tannins and phenols, which were described as cancer cell growth inhibitors [12]. In addition, taxol-producing fungal endophytes *Pestalotiopsis terminaliae* [13] and *Chaetomella raphigera* [14] have also been isolated from *T. arjuna* (arjun tree). These earlier reports on the genus *Terminalia* prompted us to look into the endophytes of *T. laxiflora* from which we isolated three endophytic fungi, *Aspergillus aculeatus, Aspergillus oryzae* and *Curvularia* sp. Endophytes are microbes that inhabit living, internal tissues of plants without causing any immediate, apparent negative effects [15]. Endophytes associated with plants are an infinite undisclosed reservoir of chemically diverse natural products. Advanced methods in cultivation and dereplication procedures have provided access to a rich source of novel drug leads, having the advantage of vast-effective production through large-scale cultivation of the microorganisms. In many cases, endophytic fungi might be involved in the biosynthesis of plant natural products, but they might also be themselves the producers of new pharmacologically active and structurally diverse secondary metabolites [16].

In the preliminary screening, fungal extracts from both *A. aculeatus* and *oryzae* were found to be active against the prostate cancer cell line (PC-3) [17] while some of the prepared *Curvularia* extracts inhibited the growth of chronic myelogenous leukemia cell (K562). Between 2009 and 2013, leukemia was the fifth and the sixth most common cause of cancer deaths in men and women, respectively [18]. According to the International Agency for Research on Cancer (IARC), there were 351,965 cases of leukemia diagnosed worldwide in 2012. The number of deaths was 265,471 that is more than 75% on the incidence numbers [19]. To date, in 2017, 24,500 people are likely to die from leukemia (14,300 males and 10,200 females) [18]. According to the American Cancer Society, it is estimated that there are 363,794 people living with or in remission from leukemia [20]. In this study, we focused on the isolation of the bioactive metabolites of a *Curvularia* extract against the leukemia cell line K562. Metabolomic tools were employed to compare the metabolite fingerprints of solid and liquid culture extracts of a *Curvularia* sp. endophyte isolated from the leaves of *T. laxiflora*. *Curvularia* sp. was earlier found to be associated with both terrestrial and/or marine natural sources. Previous studies of endophytic *Curvularia* associated with marine alga [21,22] and the medicinal plant *Murraya koenigii* [23] yielded diverse types of curvularin macrolides, which exhibited a range of bioactivities like anticancer and antimicrobial activities [21,22].

Production of the fungal bioactive metabolites was targeted through a metabolomics approach [24-26]. Putative metabolites responsible for the activity of the fungal extracts and fractions against leukemia cell lines were then pinpointed by multivariate analysis. In parallel, a natural product database was utilized to dereplicate metabolites in order to identify the known compounds and the presence of new natural products.

Results and Dicussion

By repeated inoculation on Malt-Agar (MA) plates, purified colonies of *Curvularia* sp. (strain code: TL-F2A) was isolated from the surface-sterilized leaves of *Terminalia laxiflora* collected from Al-Zohria gardens, Giza, Egypt. The fungal extract of a MA plate of *Curvularia* exhibited neither anticancer nor antimicrobial activities. However, metabolomic profiling of extracts from two other different media, which were liquid broth (LC) and rice (RC), exhibited a change in chemical profiles as shown by the heat map analysis of their mass spectral data (Fig. 1). The *m/z* peaks are indicated by bands. The occurrence of more blue bands represented the diversity of the metabolites in terms of their occurrence and increase in intensity of respective metabolites prior to the normalized average yield. The red bands signified the absence or decrease in intensity of respective metabolites. On different culture media prepared in triplicates, optimal growth was observed in 30-days agar plate, 15-days liquid broth, and 30-days rice cultures. The 7- and 30-days rice cultures exhibited the highest chemical diversity. However, the displayed chemical diversity for the 7-days rice culture represents the unutilized media constituents originating from the malt agar inoculum. For the 15-days rice culture, low production of metabolites was observed. An increase in metabolite production was exhibited by the 30-days rice culture.

HR-LCMS metabolite profiling data for the rice culture extracts processed via MZmine 2.10 (modified version) [25] also detected more metabolites from both the 7- and 30-days than the 15-days incubation period as shown by the scatter plots (Fig. 2). From the scatter plot between the 7- and 30-days culture (Fig. 2C), it was also observed that a different set of metabolites were being synthesized and degraded over the respective incubation periods as unique metabolites were indicated below and above the diagonal. As exhibited by both the heat-map and scatter plots, an incubation period of 15 days was an intermediate stage when nutrients are being depleted. As a

survival mechanism for the fungus to beat environmental stress that also includes nutrient depletion, secondary metabolites are produced during the lag phase, which occurred at a period of 30 days on the rice culture media.

The change in chemical profile between culture media and duration of incubation raised a question whether these different fermentation conditions will demonstrate any change in the bioactivity. Investigation of the biological activity of the metabolite-rich 7- and 30-days rice culture extracts showed that only the 30-days culture extract of *Curvularia sp.* exhibited a good inhibition activity of 70% against NF-kappaB at 30 μg/mL. NF-kappaB is a protein complex that controls transcription of DNA and responsible for cytokine production as well as cell survival [27]. NF-kappaB activation contributes to proliferation and survival of a variety of cancer cells. NF-kappaB inhibition substantially suppresses tumor growth and angiogenesis. Compounds described to inhibit NF-kappaB may exhibit anti-cancer activity [28] while NF-kappaB activation also plays a major role in the regulation of proinflammatory and prothrombotic responses [29] making NF-kappaB inhibition important against inflammation-associated cancer development.

Multivariate analysis of the mass spectral data of the fungal and host plant extracts by Principal Component Analysis (PCA), an unsupervised classification method discriminated both the 7- and 30-days culture extracts (Fig 3A) indicating the presence of exclusive metabolites produced during these growth phases. The unique metabolites found on the bioactive 30-days culture extract were identified on the PCA loading plot (Fig 3B), and then dereplicated using the Dictionary of Natural products (DNP) database as shown in Table 1. Some metabolites obtained from the fungal extracts were also detected in the crude leaf extracts of the host plant *Terminalia laxiflora* but at lower levels of concentration, which indicated that the isolated *Curvularia* fungus is a "true" endophyte of *T. laxiflora*. "True" endophytic fungi are isolated from internal healthy tissues whose colonization is

not a result of visible symptoms of a disease [30]. Dereplicated metabolites with corresponding $[M+H]^+$ ion peaks at *m/z* 208.0969, 355.1651 and 468.2495 eluted at a retention time of 6.54, 9.17 and 11.55 minutes, respectively, were isolated from the scaled-up batch of the 30-days rice culture fungal extract by medium pressure liquid chromatography (MPLC). These metabolites were putatively identified as the drug afalanine or *N*-acetylphenylalanine (**1**), the lichen metabolite picroroccellin from *Roccella fuciformis* [31], and the fungal metabolite cyclopiamine B from *Aspergillus caespitosus* [32], respectively. However, only afalanine matched the identity of the isolated compound at *m/z* 208.0969. The purified compounds **1**, **2**, and **3** (Fig. 4) were chromatographically obtained from the active extracts and fractions that were tested against NF-kappaB's antiapoptotic transcription factors in human Philadelphia chromosome-positive chronic myelogenous leukemia cells (K562). The NMR data of the isolated compounds **2** and **3** (Fig. 4) at *m/z* 355.1651 and 468.2495, respectively, did not match those of the dereplicated hits from the DNP database indicating that the afforded compounds were suspected novel natural products requiring further structure elucidation.

Compound **1**, which was purified from fraction CV96-103, exhibited 95% inhibition of NF-kappaB in K562 at a concentration of 100 μg/mL (500 μM). Compound **3** was found to be inactive in the assays performed in this study. As depicted by the S-plot (Fig. 5B), compound **3** was situated almost near the 'zero' scale of the 'y' axis towards the quadrant of the inactive metabolites. The isolation of compound **3** was due to its co-occurrence with compound **2** in the same fractions. Compounds **2** and **3** (*m/z* 354.158 and 467.242, respectively) were both isolated from fraction CV4-32 that gave a weaker growth inhibition of the K562 cell line by 70% when compared to fraction CV96-103, which contained a higher concentration of compound **1** (*m/z* 207.090) as shown by its peak area (Fig. 5C). Compounds **2** and **3** were also found in the inactive fractions but at very low

concentrations as exhibited by their respective peak areas**.** The low concentration of **2** in the inactive fractions was below the potency level of the compound.

The mass spectral data of the MPLC fractions and extracts were subjected to OPLS-DA by grouping them according to their bioactivities on NF-kappaB and K562. The bioactive fractions CV4-32, CV60-86, and CV96-103 were grouped together against the inactive fractions. Assayed at a concentration of 30 μg/mL, fraction CV4-32 inhibited the growth of K562 cell line by 70%, while CV60-86 and CV96-103 both exhibited 95% inhibition. The score plot (Fig. 5A) showed that the two active fractions CV4-32 and CV96-103 were clustered together while CV60-86 was an outlier and considered to have a more diverse chemistry. The S-loading plot (Fig. 5B) classified the putative distinct metabolites belonging to the active fractions, which included $[M+H]^+$ ion peaks at *m/z* 208.0969 and 355.1651, representing compounds **1** and **2**, respectively, also found on the 30-day culture extract as shown in the PCA loading plot on Fig. 3B. Furthermore, ion peaks at *m/z* 172.0990 $[M-H]^-$ and 345.2050 $[2M-H]^-$ eluting at 5.41 min were also determined together with compounds **1** and **2**, as among the natural products predicted to be responsible for the extract's bioactivities against NF-kappaB and K562. The ion peak at *m/z* 172.0990 was dereplicated as *N*-acetyl-leucine, however, it was not possible to isolate and purify the compound present from CV60-86 due to a low fraction yield of 6.9 mg. Compound **1** was identified in fractions CV60-86 and CV96-103, while compound **2** was identified in both fractions CV4-32 and CV60-86 (Fig. 5C).

Compound **1** was isolated as a yellowish brown amorphous powder (8 mg), which exhibited a molecular formula of $C_{11}H_{13}O_3N$ established on the basis of ESI-HRMS at *m/z* 208.0969 $[M+H]^+$ (calcd. 208.0968), UV (MeOH) λ_{max} (log ε) 228(5.03), 308 (5.09) nm. The ^{1}H-NMR spectrum (DMSO-d_6, 400 MHz) (Supporting Information Fig. 1S) showed six proton signals at δ_H 8.15 (d, 1H, J = 7.8 Hz), 7.24 (m, 5H), 4.37 (m, 1H), 3.04 (dd, 1H, J = 13.6, 3.9 Hz), 2.82

(dd, 1H, J = 13.6, 9.5 Hz) and 1.77 (s, 3H). The structure of the amino acid spin system was detected *via* a ^{1}H-^{1}H TOCSY spectrum (Supporting Information Fig. 2S) which showed the correlation between the N*H* group at δ_H 8.15 and CH_{2A} (δ_H 3.04), CH_{2B} (δ_H 2.82) and α-C*H*NH at δ_H 4.37. Comparison of ^{1}H- and ^{13}C-NMR data of compound **1** (Table 2) with data from literature [33-35] along with the 2D spectra of the isolated compound, **1** was identified as the drug *N*-acetylphenylalanine (Fig. 4) also known as afalanine [36]. Afalanine is used as an antidepressant drug and in combination with antibiotics to prevent kidney damage [37].

The configuration of the isolated *N*-acetyl-phenylalanine (**1**) was established by Marfey's derivatisation implying an L configuration (Fig 3S). The NMR data and retention time of the compound **1** were identical to the commercially available D and L standards while the optical rotation of **1**, $[\alpha]_D^{25}$ = +12^o (*c* 0.05 in MeOH), followed a similar signage as the L standard with a value of $[\alpha]_D^{25}$ = +40^o (*c* 0.05 in MeOH). The difference in magnitude of the optical rotation between the isolated compound and the standard could be due to the lower racemic purity of compound **1**.

Compound **2**, obtained as a yellow powder (15 mg), which exhibited a molecular formula $C_{20}H_{22}N_2O_4$ established by ESI-HRMS at *m/z* 355.1645 $[M+H]^+$ (calcd 355.1652), UV (MeOH) λ_{max} (log ε) 223 (5.53), 257 (5.27) nm. The tandem MS/MS spectrum (Fig. 6A, Supporting Information Fig. 4S) showed fragment ions at *m/z* 190.0867 and 166.0868 for *N*-acetyl-L-phenylalanine $[C_{11}H_{12}O_2N - OH]$ and phenylalanine $[C_9H_{10}NO_2]$ moieties, respectively. The ^{1}H-NMR spectrum (DMSO-d_6, 400 MHz) (Table 2, Supporting Information Fig. 5S) showed proton resonances at δ_H 8.39 (d, J = 8.1 Hz, 1H), 8.35 (d, J = 8.5 Hz, 1H), 7.05-7.27 (m, 10H), 5.18 (m, 1H), 5.12 (m, 1H), 2.43-2.70 (m, 4H) and 1.80 (s, 3H). The ^{13}C-NMR spectrum (DMSO-d_6, 100 MHz) (Table 2 and Supporting Information Fig. 6S) showed 20 carbon signals including five quaternary resonances,

ten aromatic methines, two methylenes, and one methyl carbon. The TOCSY spectrum (Supporting Information Fig. 7S) showed two spin systems, a correlation between N*H* at δ_H 8.39 and α-C*H*NH-2''' (δ_H 5.12) and CH_2-3''' (δ_H 2.43-2.70) of a phenylalanine moiety, while the N*H* at δ_H 8.35 correlated with α-C*H*NH-2 (δ_H 5.18) and CH_2-3 (δ_H 2.43-2.70) of the *N*-acetyl-phenylalanine unit. The structure of the *N*-acetyl-phenylalanine was confirmed from the HMBC spectrum (Fig. 8S and Fig. 6A) through the correlations of α-C*H*NH-2 (δ_H 5.18) with C-3 (δ_C 43.0), C-2'' (δ_C 168.8), C-1' (δ_C 143.3) and C-2'/6' (δ_C 127.2). In addition, N*H* at δ_H 8.35 showed a cross peak with C-2'' carbonyl and C-2 (δ_C 50.5), the α carbon for *N*-acetyl-phenylalanine. The structure of a phenylalanine moiety was confirmed through the correlations of an α-C*H*NH-2''' at δ_H 5.12 with C-1''' (δ_C 172.3), C-3''' (δ_C 41.3), C-1 (δ_C 168.9), C-1'''' (δ_C 142.8), C-2''''/6'''' (δ_C 126.8); while the N*H* at δ_H 8.39 correlated with C-1 and C-2''' at δ_C 49.8. The correlation of α-C*H*NH and N*H* of phenylalanine moiety with the carbonyl C-1 at δ_C 168.9 of *N*-acetylphenylalanine confirmed the attachment of the two moieties through the amide linkage. Furthermore, the configuration of phenylalanine, whether it is D or L amino acid was assigned through derivatization with Marfey's reagent FDAA (1% *N*-(1-flouro-2,4-dinitrophenyl)-5-L-alanine amide) [38] and HPLC-MS data of the derived products of analysis (Supporting Information Fig. 9S). The hydrolysis product of compound **2** reacted with Marfey's reagent to afford $C_{18}H_{20}O_7N_5$ at *m/z* 418.1355 $[M+H]^+$ eluting at 13.70 min which corresponded with the standard L-phenylalanine as the hydrolysis product. On the other hand, the D-phenylalanine standard eluted at 14.87 min. Compound **2** (Fig. 4) was identified as *N*-acetyl-L-phenylalanyl-L-phenylalanine [39], a dipeptide isolated for the first time from a natural source. *N*-acetyl amino acids are very useful in the pharmaceutical and food industries as they play an important role in methylation, sulfuration, detoxication and antioxidation [39]. The related natural product *N*-acetyl-L-phenylalanyl-L-phenylalaninol has been previously isolated from

culture filtrates of the fungus *Emericellopsis salmosynnemata* [40], a producer of the peptide antibiotic zervamicin [41,42]. The hydroxylated methylene unit found in *N*-acetyl-L-phenylalanyl-L-phenylalaninol was instead substituted by a carboxylic acid moiety in compound **2**.

Compound **3**, obtained as a brownish white residue (16 mg), it has a molecular formula of $C_{26}H_{33}N_3O_5$ established by ESI-HRMS at *m/z* 468.2490 $[M+H]^+$ (calcd 468.2493), UV (MeOH) λ_{max} (log ε) 221 (5.41), 257 (4.84) nm. The MS/MS (Fig. 10S and Fig. 6B) showed three main fragments at *m/z* 337.1535, 279.1696 and 190.0858 for *N*-acetyl-phenylalanyl-phenylalanine [$C_{20}H_{21}O_3N_2$], phenylalanyl-leucine [$C_{15}H_{23}O_3N_2$], *N*-acetyl-L-phenylalanine [$C_{11}H_{12}O_2N - OH$], respectively. The fragment for *N*-acetyl-L-phenylalanine was also observed in compound **2**. The ^{1}H-NMR spectrum (DMSO-d_6, 400 MHz) (Table 2, Supporting Information Fig. 11S) was very similar to compound **2** particularly in the aromatic region. However, additional proton signals were observed at the aliphatic region at δ_H 4.11 (m), 1.37 (*m*, 2H), 1.27 (*m*, 1H), 0.78 (*d*, 5.9, 3H) and 0.68 (*d*, 5.8, 3H). The ^{13}C-NMR spectrum (DMSO-d_6, 100 MHz) (Table 2, Supporting Information Fig. 12S) showed 26 carbon signals instead of 20 when compared to compound **2** including six quaternary carbons, ten aromatic methines, three methylenes, three methyl units and four methines with three α-C*H*NHs and one aliphatic methine. The TOCSY spectrum (Supporting Information Fig. 13S) indicated the presence of three amino acids, that included leucine and two units of phenylalanine**.** The occurrence of leucine was established through the correlation of the N*H** group (8.09) with α-C*H*NH-2* (4.11), C*H*$_2$-3* (1.37) and the two methyl doublets Me-5* (δ_H 0.68) and Me-6* (δ_H 0.78). The connectivity of the three amino acid units were confirmed from the HMBC spectrum (Fig. 14S and Fig. 6B). As in compound **2**, *N*-acetyl-phenylalanine was linked to a phenylalanine moiety by a similar pattern of cross peaks. Leucine was linked to phenylalanine through correlations of α-C*H*NH-2* at δ_H 4.11 with C-1**'''** of phenylalanine, a carboxyl amide signal at δ_C 174.8 (C-1*) and a methylene carbon at

40.6 (C-3*). The leucine moiety was further confirmed with correlations of C*H*$_2$-3* (δ_H 1.37) with C-2* (δ_C 50.6) and the two methyl doublets Me-5* (δ_C 23.6) and Me-6* (δ_C 21.6) which in return correlated with C-3*, C-4* (δ_C 24.4). The configuration of the phenylalanine and leucine were assigned through a Marfey's reagent experiment (Supporting Information Fig. 15S). Marfey derivatives of the hydrolysed products of compound **3** afforded $C_{15}H_{22}N_5O_7$ at *m/z* 384.1548 $[M+H]^+$ eluting at 13.80 min which paralleled to that of standard L-leucine at 13.82 min while the standard D-leucine had a retention time of 15.24 min. For phenylalanine, a similar result was obtained as in compound **2**. Compound **3** (Fig. 4) was elucidated as *N*-acetyl-L-phenylalanyl-L-phenylalanyl-L-leucine and isolated for the first time from the endophytic fungus *Curvularia* sp. Compound **3** was first described as a product of α-chymotrypsin-catalysed syntheses of peptides using various *N*-acylated amino acid or peptide esters as donors with amino acid derivatives or their peptide derivatives as acceptors [43].

In this study, we were able to demonstrate a metabolomic-guided isolation protocol. The protocol afforded a logical systematic prediction of the biologically active target compounds. A combination of metabolomic- and a bioassay-guided protocol can efficiently predict the novelty of putative biologically active metabolites during the first stage of fractionation. This solves the problem on making false hypothesis on the dependence of bioactivity on the yield of respective metabolites in an active fraction or extract. The low yield of a very potent metabolite is challenging in determining the bioactive compound at the initial stage of the isolation work. In this study, a metabolomic-guided protocol provided focus on the potency of the metabolite(s) present in a complex extract or fraction at micro- or nano-gram levels.

Materials and Methods

Fungal Material

The endophytic fungus was isolated from fresh leaves of *Terminalia laxiflora* (Combretaceae) that were collected on the 15th of October 2010 from Al-Zohriya Gardens (30°02'45.0"N 31°13'32.5"E), Al-Zamalek in Giza, Egypt. Dr Therese L. Yousef from Orman Garden taxonomically identified the plant and a voucher specimen (No. 822) was deposited in her office in Orman Garden.

The collected 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 for 5 min under a laminar flow hood. With a sterile scalpel, outer tissues were removed from the plant samples and the inner tissues were carefully dissected under sterile conditions and laid over malt agar (MA) plates containing chloramphenicol. After four weeks of incubation at 30°C, hyphal tips of the fungi were removed and transferred to plates with fresh MA medium. Plates were prepared in duplicates to eliminate the possibility of contamination. Pure strains were isolated by repeated inoculation. The purified fungus was later transferred to rice media for scaling up.

Small-scale extraction for screening, metabolomics profiling and dereplication.

A plate of each fungal species was transferred into 250 ml flask and macerated overnight with ethyl acetate (EtOAc). The mycelia with the agar media was then homogenized with an ultra-turrax (IKA) and vacuum filtered. The filtered mycelia was again macerated sequentially 3× with 200 mL EtOAc and filtered. The pooled filtrate was then dried under vacuo, re-suspended in 200 mL H_2O and partitioned 3× with 200 mL EtOAc in a separating funnel. (*N.B*. Re-suspension of the dried crude extract in H_2O provided better separation of the aqueous and organic phase as well as a more efficient extraction of non- and semi-polar metabolites from the aqueous phase.) The ethyl acetate soluble portions were concentrated and dried under *vacuo*. The water-soluble portion was also

concentrated and then passed over DiAION HP-20 (Supelco) column (100 x 600 mm) using methanol as an eluant. The methanol eluate was concentrated and dried under *vacuo.* 1 mg of each of the dried extract was subjected to HRMS analysis for metabolomics profiling and dereplication studies. LC-MS spectra were processed using Thermo Xcalibur 2.1 (Thermo Scientific). The program msconvert from ProteoWizard [44] was used to convert the raw data into separate positive and negative ionisation mzXML files. The files were then imported to the data mining software MZmine 2.10 (modified version as described by MacIntyre et al, 2014 [24]) for peak picking, deconvolution, deisotoping, alignment and formula prediction [24,25]. An in-house macro file with built in databases written in Excel was employed for further clean-up of background peaks from the media components and to dereplicate positive and negative ion peaks [24]. The Dictionary of Natural Products (DNP) 2015 database was used for the dereplication study. MestReNova (MNova) 2.10 by Mestrelab Research S.L, was utilised to process all NMR data. SIMCA 14 analysis (Umetrics AB) was applied for multivariate data assessment.

Samples at concentrations of 8-10 mg in 600 μL DMSO-d_6 were sent to NMR analysis to monitor presence of lipids and steroids poorly detected by MS. Sample in duplicates of 1 mg/mL concentration of each fungal extract from the malt agar plate (MA), 15-days liquid broth culture (LC), 7-days rice culture (RC.7DAY), and 30-days rice culture (RC.30DAY) were prepared and sent to Strathclyde Institute for Drug Research SIDR for bioassay against NF-kappaB and chronic myelogenous leukaemia (K562) cancer cell line.

Fermentation, Extraction, and Isolation

The fermentation was carried out in two Erlenmeyer flasks (1L each) on rice medium, which was prepared with 100 g of basmati rice and approximately 100 mL of demineralized water just enough to cover the rice layer. The rice media was autoclaved prior to inoculation. A 15-days 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 concentrated under reduced pressure. The EtOAc extract was evaporated, dissolved in 10% aqueous MeOH and defatted by partitioning with n-hexane in a separating funnel. The MeOH soluble portion (2 g) was then dried and reconstituted with 30 mL of EtOAc to be loaded onto a 37 mm Biotage SNAP Ultra C18 Samplet 3 g frit and dried overnight under a fume hood. The dried loaded frit was placed over a Biotage C_{18} SNAP 60g silica gel cartridge (85 mm x 37 mm). Fractionation of the fungal extract was done using the Biotage IsoleraTM Spektra One Flash Purification System ISO-1SV. The flash system was equipped with photodiode array (PDA) detector with a wavelength range of 200 to 400 nm. The chromatographic run was accomplished using 0.1 % formic acid in water and 0.1 % formic acid in acetonitrile as solvents. A step-wise gradient elution was employed by commencing with 100 % water reaching to 50 % acetonitrile in 70 minutes followed by another 20 minutes to 100 % acetonitrile with flow rate 12 mL/min. The fractions were concentrated, pooled according to their PDA chromatogram peaks and their purity on the TLC; which yielded 8 mg of **1** eluting at 30% acetonitrile, 15 mg of **2** and 16 mg of **3** that eluted at 45% and 50% acetonitrile, respectively. The pooled fractions were also further monitored for their purity by HPLC-HRMS. The generated HPLC-HRMS data was also use for metabolomics profiling of the fractions.

Supporting information

1D and 2D NMR, MS spectra of compounds **1**, **2**, and **3** as well as part of the experimental methods that includes the taxonomical identification of the fungi by molecular biological methods, Marfey derivatisation and bioscreening of the isolated compounds are available as Supporting Information.

Acknowledgment

This work was financially supported by mission sector of the Egyptian Ministry of Higher Education awarded to Ahmed Tawfike. Chromatographic equipment used in this project was funded by the Royal Society Seed-Corn Research Grant awarded to RuAngelie Edrada-Ebel.

Conflict of interests

The authors declare that there are no conflicts of interests.

References

[1]Mbwambo ZH, Erasto P, Nondo RO, Innocent E, Kidukuli AW. Antibacterial and cytotoxic activities of terminalia stenostachya and *terminalia spinosa*. Tanzan J Health Res 2011; 13: 119-124

[2]Kim JH, Koo YC, Hong CO, Yang SY, Jun W, Lee KW. Mutagenicity and oral toxicity studies of *terminalia chebula*. Phytother Res 2012; 26: 39-47

[3]Ravi Shankara BE, Ramachandra YL, Rajan SS, Ganapathy PS, Yarla NS, Richard SA, Dhananjaya BL. Evaluating the anticancer potential of ethanolic gall extract of *terminalia chebula* (gaertn.) retz. (combretaceae). Pharmacognosy Res 2016; 8: 209-212

[4]Lin CC, Hsu YF, Lin TC. Antioxidant and free radical scavenging effects of the tannins of *terminalia catappa* l. Anticancer Res 2001; 21: 237-243

[5]Tan AC, Konczak I, Ramzan I, Zabaras D, Sze DM. Potential antioxidant, antiinflammatory, and proapoptotic anticancer activities of kakadu plum and illawarra plum polyphenolic fractions. Nutr Cancer 2011; 63: 1074-1084

[6]Basu T, Panja S, Ghate NB, Chaudhuri D, Mandal N. Antioxidant and antiproliferative effects of different solvent fractions from *terminalia belerica* roxb. Fruit on various cancer cells. Cytotechnology 2017; 69: 201-216

[7]Kuo PL, Hsu YL, Lin TC, Chang JK, Lin CC. Induction of cell cycle arrest and apoptosis in human non-small cell lung cancer a549 cells by casuarinin from the bark of *terminalia arjuna* linn. Anticancer Drugs 2005; 16: 409-415

[8]Ghate NB, Hazra B, Sarkar R, Chaudhuri D, Mandal N. Alteration of bax/bcl-2 ratio contributes to *terminalia belerica*-induced apoptosis in human lung and breast carcinoma. In Vitro Cell Dev Biol Anim 2014; 50: 527-537

[9]Messeha SS, Zarmouh NO, Taka E, Gendy SG, Shokry GR, Kolta MG, Soliman KF. The role of monocarboxylate transporters and their chaperone cd147 in lactate efflux inhibition and the anticancer effects of *terminalia chebula* in neuroblastoma cell line n2-a. Eur J Med Plants 2016; 12:

[10]Achari C, Reddy GV, Reddy TC, Reddanna P. Chebulagic acid synergizes the cytotoxicity of doxorubicin in human hepatocellular carcinoma through cox-2 dependant modulation of mdr-1. Med Chem 2011; 7: 432-442

[11]Yeh CB, Hsieh MJ, Hsieh YS, Chien MH, Lin PY, Chiou HL, Yang SF. *Terminalia catappa* exerts antimetastatic effects on hepatocellular carcinoma through transcriptional inhibition of matrix metalloproteinase-9 by modulating nf-kappab and ap-1 activity. Evid Based Complement Alternat Med 2012; 2012: 595292

[12]Saleem A, Husheem M, Harkonen P, Pihlaja K. Inhibition of cancer cell growth by crude extract and the phenolics of *terminalia chebula* retz. Fruit. J Ethnopharmacol 2002; 81: 327-336

[13]Gangadevi V, Muthumary J. Taxol production by *pestalotiopsis terminaliae*, an endophytic fungus of *terminalia arjuna* (arjun tree). Biotechnol Appl Biochem 2009; 52: 9-15

[14]Gangadevi V, Muthumary J. A novel endophytic taxol-producing fungus *chaetomella raphigera* isolated from a medicinal plant, *terminalia arjuna*. Appl Biochem Biotechnol 2009; 158: 675-684

[15]Newman DJ, Cragg GM, Snader KM. Natural products as sources of new drugs over the period 1981-2002. J Nat Prod 2003; 66: 1022

[16]Arnold EA. Understanding the diversity of foliar endophytic fungi: Progress, challenges, and frontiers. Fungal Biology Rev 2007; 21: 51-66

[17]Tawfike A, Tate R, Abbott G, Young L, Viegelmann C, Schumacher M, Diederich M, Edrada-Ebel R. Metabolomic tools to assess the chemistry and bioactivity of endophytic aspergillus strain. Chem Biodivers 2017, DOI: 10.1002/cbdv.201700040:

[18]International agency for research on cancer. In. http://www.lls.org/http%3A/llsorg.prod.acquia-sites.com/facts-and-statistics/facts-and-statistics-overview/facts-and-statistics: Leukemia & Lymphoma Society; 2017

[19]Cancer research uk. In. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/leukaemia#heading-One; 2017

[20]Cancer treatment & survivorship facts & figures 2016-2017. In. https://www.cancer.org/research/cancer-facts-and-statistics/cancer-treatment-and-survivorship-facts-and-figures/: American Cancer Society; 2017

[21]Dai J, Krohn K, Floerke U, Pescitelli G, Kerti G, Papp T, Koever KE, Benyei AC, Draeger S, Schulz B, Kurtan T. Curvularin-type metabolites from the fungus *curvularia* sp. Isolated from a marine alga. Eur J Org Chem 2010; 2010: 6928-6937

[22]Greve H, Schupp PJ, Eguereva E, Kehraus S, Kelter G, Maier A, Fiebig H-H, Konig GM. Apralactone a and a new stereochemical class of curvularins from the marine fungus *curvularia* sp. Eur J Org Chem 2008; 2008: 5085-5092

[23]Mondol MA, Farthouse J, Islam MT, Schuffler A, Laatsch H. Metabolites from the endophytic fungus curvularia sp. M12 act as motility inhibitors against phytophthora capsici zoospores. J Nat Prod 2017; 80: 347-355

[24]MacIntyre L, Zhang T, Viegelmann C, Juarez Martinez I, Cheng C, Dowdells C, Abdelmohsen UR, Gernert C, Hentschel U, Edrada-Ebel R. Metabolomic tools for secondary metabolite discovery from marine microbial symbionts. Mar Drugs 2014; 12: 3416-3448

[25]Tawfike AF, Viegelmann C, Edrada-Ebel R. Metabolomics and dereplication strategies in natural products. In: Roessner U, Dias DA eds, Metabolomics tools for natural product discovery: Humana Press; 2013: 227-244

[26]Kamal N, Viegelmann CV, Clements CJ, Edrada-Ebel R. Metabolomics-guided isolation of anti-trypanosomal metabolites from the endophytic fungus *lasiodiplodia theobromae*. Planta Medica 2016, DOI: 10.1055/s-0042-118601:

[27]Gilmore TD. Introduction to nf-kappab: Players, pathways, perspectives. Oncogene 2006; 25: 6680-6684

[28]Escarcega RO, Fuentes-Alexandro S, Garcia-Carrasco M, Gatica A, Zamora A. The transcription factor nuclear factor-kappa b and cancer. Clin Oncol (R Coll Radiol) 2007; 19: 154-161

[29]Monaco C, Andreakos E, Kiriakidis S, Mauri C, Bicknell C, Foxwell B, Cheshire N, Paleolog E, Feldmann M. Canonical pathway of nuclear factor κb activation selectively regulates proinflammatory and prothrombotic responses in human atherosclerosis. Proc Natl Acad Sci USA 2004; 101: 5634-5639

[30]Hyde KD, Soytong K. The fungal endophyte dilemma. Fungal Diversity 2008; 33: 163-173

[31]Jones MP, Keane J, Nolan TJ. Lichen substances containing nitrogen. Nature 1944; 154: 580-580

[32]Bond R, Boeyens J, Holzapfel C, Steyn P. Cyclopiamines a and b, novel oxindole metabolites of penicillium cyclopium westling. J Chem Soc, Perkin Trans 1979; 1: 1751–1761

[33]Dressen MHCL, van de Kruijs BHP, Meuldijk J, Vekemans JAJM, Hulshof LA. From batch to flow processing: Racemization of n-acetylamino acids under microwave heating. Org Process Res Dev 2009; 13: 888-895

[34]Naturale G, Lamblin M, Commandeur C, Felpin F-X, Dessolin J. Direct c–h alkylation of naphthoquinones with amino acids through a revisited kochi–anderson radical decarboxylation: Trends in reactivity and applications. Eur J Org Chem 2012; 2012: 5774-5788

[35]Koshti N, Naik S, Parab B. Polymer-bound cationic rh(i) phosphine catalyst for homogeneous asymmetric hydrogenation. Indian J Chem, Sect B: Org Chem including Med Chem 2005; 44B: 2555-2559

[36]Andrew W. Afalanine. In, Pharmaceutical Manufacturing Encyclopedia. 3rd ed. Norwich, NY., U.S.A.: William Andrew Publishing; 2007

[37]Stout KL, Hallock KJ, Kampf JW, Ramamoorthy A. N-acetyl-l-phenylalanine. Acta Crystallogr C 2000; 56: E100

[38]Ibrahim SRM, Min CC, Teuscher F, Ebel R, Kakoschke C, Lin W, Wray V, Edrada-Ebel R, Proksch P. Callyaerins a-f and h, new cytotoxic cyclic peptides from the indonesian marine sponge *callyspongia aerizusa*. Bioorg Med Chem 2010; 18: 4947-4956

[39]Li G-N, Li J, Liang Z-P. (s)-2-[(s)-2-acetamido-3-phenylpropanamido]-3-phenylpropanoic acid. Acta Crystallogr E 2007; 63: o372-o373

[40]Argoudelis AD, Mizsak SA, Baczynskyj L. N-acetyl-l-phenylalanyl-l-phenylalaninol a metabolite of *emericellopsis salmosynnemata*. J Antibiot (Tokyo) 1975; 28: 733-736

[41]Argoudelis AD, Dietz A, Johnson LE. Zervamicins i and ii, polypeptide antibiotics produced by emericellopsis salmosynnemata. J Antibiot (Tokyo) 1974; 27: 321-328

[42]Ovchinnikova TV, Levitskaya NG, Voskresenskaya OG, Yakimenko ZA, Tagaev AA, Ovchinnikova AY, Murashev AN, Kamenskii AA. Neuroleptic properties of the ion-channel-forming peptaibol zervamicin: Locomotor activity and behavioral effects. Chem Biodivers 2007; 4: 1374-1387

[43]Morihara K, Oka T. A-chymotrypsin as the catalyst for peptide synthesis. Biochem J 1977; 163: 531-542

[44]Chambers MC, Maclean B, Burke R, Amodei D, Ruderman DL, Neumann S, Gatto L, Fischer B, Pratt B, Egertson J, Hoff K, Kessner D, Tasman N, Shulman N, Frewen B, Baker TA, Brusniak MY, Paulse C, Creasy D, Flashner L, Kani K, Moulding C, Seymour SL, Nuwaysir LM, Lefebvre B, Kuhlmann F, Roark J, Rainer P, Detlev S, Hemenway T, Huhmer A, Langridge J, Connolly B, Chadick T, Holly K, Eckels J, Deutsch EW, Moritz RL, Katz JE, Agus DB, MacCoss M, Tabb DL, Mallick P. A cross-platform toolkit for mass spectrometry and proteomics. Nat Biotechnol 2012; 30: 918-920

Legends for Figures

Fig. 1 - Heat map analysis of the mass spectral data of rice culture (RC) extracts at 7, 15, and 30 days in comparison with the 15-days liquid broth culture extract and 30-days malt-agar (MA) culture extract. Legend: RC.7DAY = 7-days rice culture; RC.15DAY = 15-days rice culture; RC.30DAY = 30-days rice culture; LC.15DAY=15-days liquid broth culture; MA.plate = Malt Agar plate

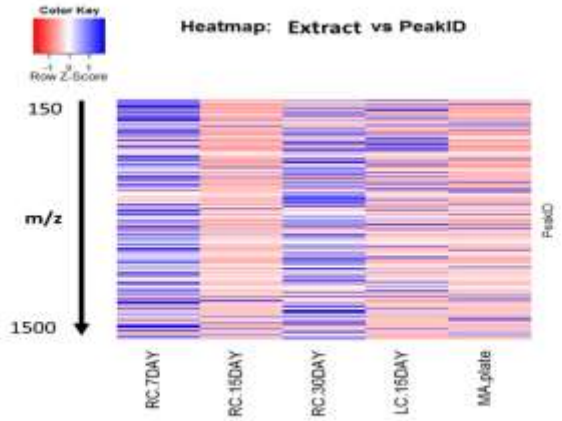


Fig. 2 - Scatter plot of the mass spectral data between (A) 15- and 7-days, (B) 15- and 30-days, and (C) 7- and 30-days rice culture extracts.

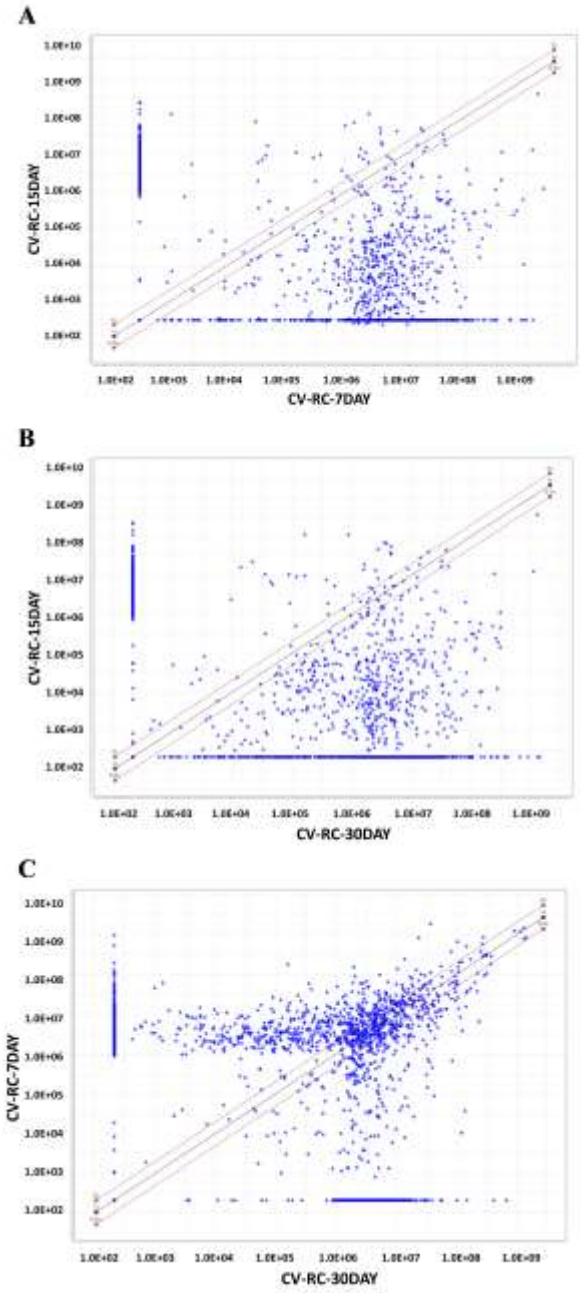


Fig. 3 - (A) PCA score plot of the fractions where the blue dots are the inactive fractions and the red dots are the active fractions and (B) loading plot of the mass spectral data of crude extracts obtained from the rice and liquid culture of *Curvularia* along with the leaf extracts of its host plant *Terminalia laxiflora*. Shown data are the molecular weights. $R_2 = 0.648$, $Q_2 = 0.309$

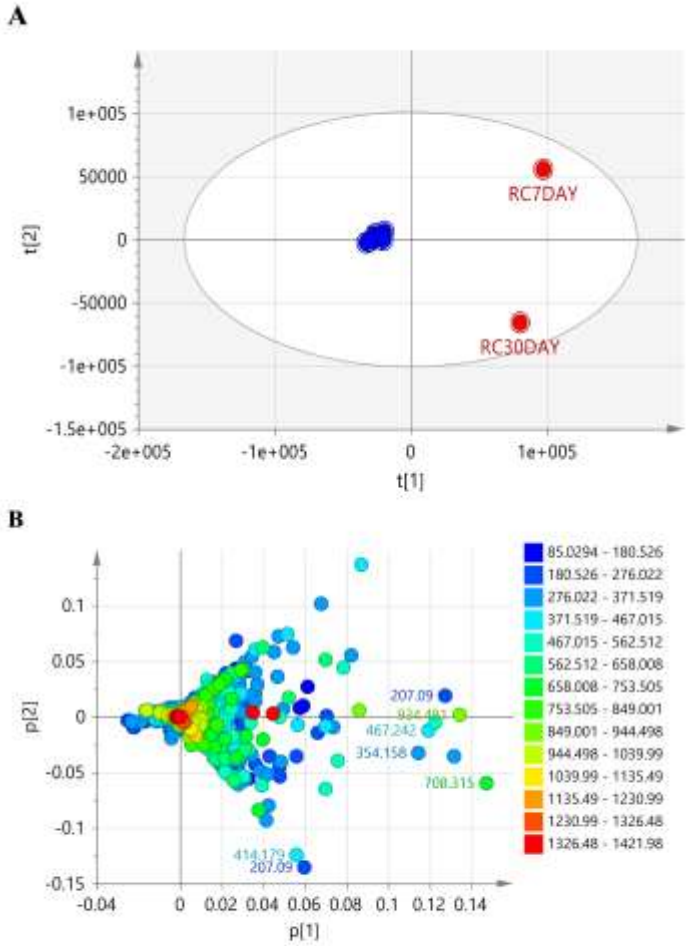
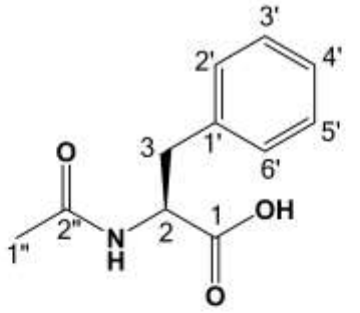
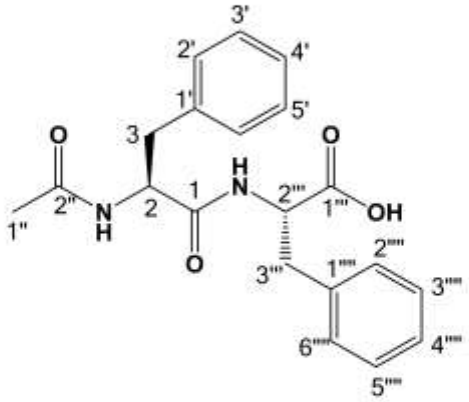


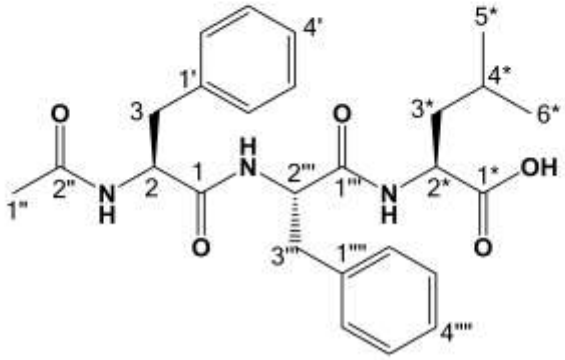
Fig. 4 - Isolated compounds *N*-acetylphenylalanine (**1**), *N*-acetylphenylalanyl-L-phenylalanine (**2**) and *N*-acetylphenylalanyl-L-phenylalanyl-L-leucine (**3**) from *Curvularia* sp.



1



2



3

Fig. 5 - (A) OPLS-DA score and (B) S-loading plots of the mass spectral data of active versus inactive fractions on NF-kappaB in K562 obtained from the 30-day rice culture of the endophytic fungus *Curvularia*. Shown data are the molecular weights. $R_2 = 1$, $Q_2 = 1$. (C) Relative abundance and peak area of predicted and isolated biologically active metabolites in the fractions.

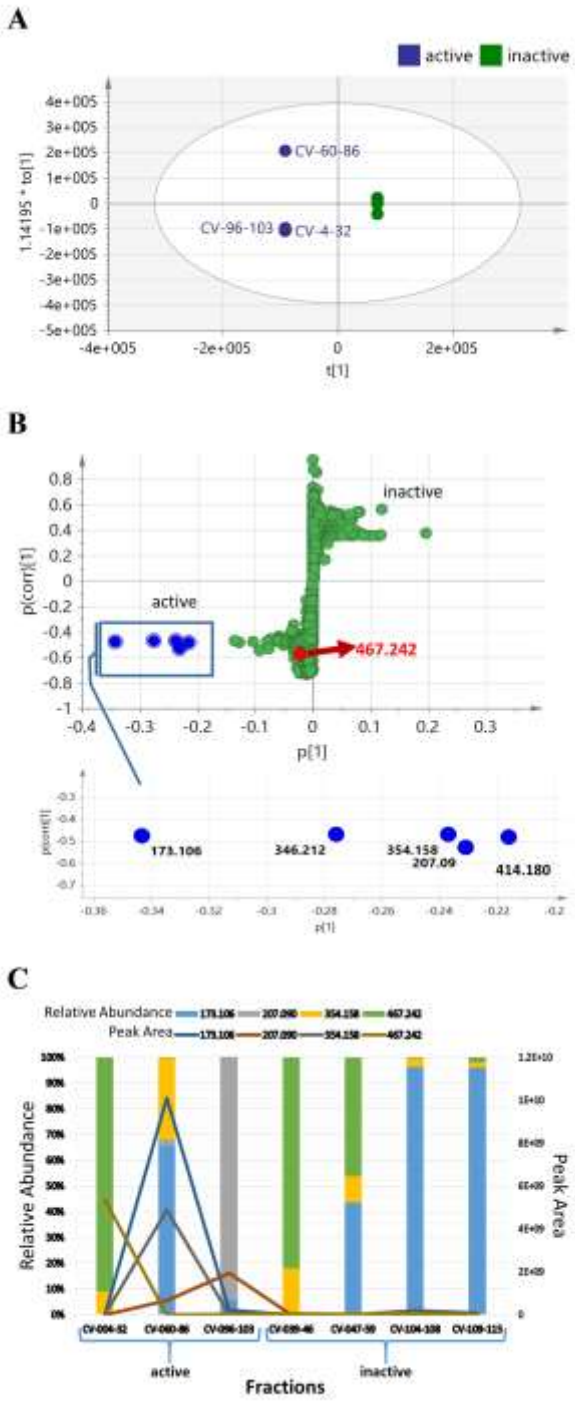


Fig. 6 - HMBC correlations and MS/MS fragments of compounds **2** (A) and **3** (B). Red arrows (●) represent *N*-acetylphenylalanine moiety, blue arrows (●) for L-phenylalanine, and green (●) for L-leucine.

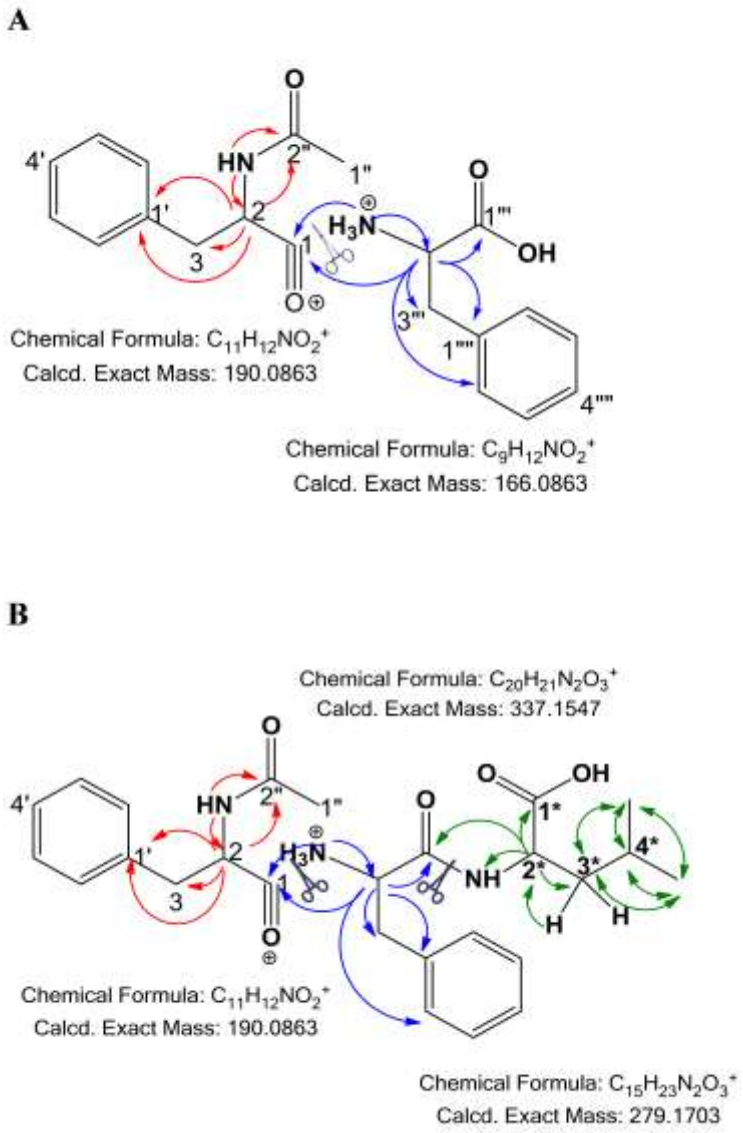


Table 1: Dereplication of the mass spectral data for the major total ion chromatogram peaks for the 30-days rice culture extract in comparison with the 15-days liquid broth culture extracts of *Curvularia* along with the butanol (BuOH) and ethyl acetate (EtOAc) leaf extracts of the host plant *Terminalia laxiflora*. The highlighted rows represent the isolated fungal metabolites with identical molecular formula as predicted by MZmine 2.10 (modified version). RT = retention time; MW = molecular weight; [P] = plant; [L] = lichen; [F] = fungus

RT (min)	*m/z* [ionization]	MW	Putative compound identified / Molecular formula/ Known source	Peak Area			
				Rice culture	Liquid broth culture	Plant leaf extracts	
				30-days	15-days	BuOH	EtOAc
5.41	172.0990 $[M-H]^-$	173.106	*N*-acetyl-leucine $C_8H_{15}NO_3$	1.79E+07	3.42E+05	5.19E+03	1.43E+04
5.41	174.1126 $[M+H]^+$	173.106	*N*-acetyl-leucine $C_8H_{15}NO_3$	1.91E+08	3.05E+06	2.73E+04	4.56E+04
5.41	345.2050 $[2M-H]^-$	346.212	Complex of 173.06	2.07E+06	0	0	0
6.54	206.0820 $[M-H]^-$	207.096	afalanine, INN (**1**) $C_{11}H_{13}NO_3$	8.00E+07	5.45E+05	1.77E+04	6.32E+04
6.54	208.0969 $[M+H]^+$	207.096	afalanine, INN $C_{11}H_{13}NO_3$	3.01E+08	1.48E+06	7.58E+03	2.96E+04
6.54	413.1720 $[2M-H]^-$	414.180	Complex of 207.096	2.27E+03	5.20E+02	2.70E+05	1.07E+04
7.99	237.0760 $[M+H]^+$	236.064	herbarin A $C_{12}H_{12}O_5$ *Cladosporium herbarum* [F]	2.74E+07	9.46E+08	1.57E+03	5.85E+03
9.17	355.1651 $[M+H]^+$	354.158	picroroccellin $C_{20}H_{22}N_2O_4$ *Roccella fuciformis* [L]	1.46E+09	1.30E+06	1.66E+04	7.79E+02
9.17	709.3240 $[2M+H]^+$	708.315	Complex of 354.158	2.90E+08	0	0	0
9.34	426.2389 $[M+H]^+$	425.234	dihydroxyisoechinulin A $C_{24}H_{31}N_3O_4$ *Aspergillus* strain MFA 212 KACC [F]	2.08E+07	4.40E+06	0	0
11.55	468.2495 $[M+H]^+$	467.242	cyclopiamine B $C_{26}H_{33}N_3O_5$ *Aspergillus caespitosus* [F]	9.20E+07	5.48E+06	0	0
11.55	935.4923 $[2M+H]^+$	934.481	Complex of 467.242	3.31E+08	2.95E+05	7.58E+02	0
23.14	535.2910 $[M-H]^-$	536.298	sengosterone $C_{29}H_{44}O_9$ *Cyathula capitata* [P]	2.25E+07	4.49E+05	0	0
29.27	297.2430 $[M-H]^-$	298.251	*E*-11-hydroxy-octadeca-12-enoic acid $C_{18}H_{34}O_3$	2.27E+07	5.88E+05	2.88E+02	2.05E+03

Table 2. ^{1}H- (400 MHz) and ^{13}C-NMR (100 MHz) data of compounds **1**, **2**, and **3** measured in DMSO-d_6.

Atom No.	**1**		**2**		**3**	
	δ_C (m)	δ_H (m, *J* in Hz)	δ_C (m)	δ_H (m, *J* in Hz)	δ_C (m)	δ_H (m, *J* in Hz)
1	173.9 (C)		168.9 (C)		168.7 (C)	
2	54.4 (CH)	4.37 (m)	50.5 (CH)	5.18 (m, 1H)	50.5 (CH)	5.19 (m)
3A	37.5 (CH_2)	3.04 (dd, 3.9, 13.6)	43.0 (CH_2)	2.43-2.70 (m)	43.0 (CH_2)	2.40-2.61 (m)
3B		2.82 (dd, 9.5, 13.6)		2.43-2.70 (m)		2.40-2.61 (m)
1'	138.6 (C)		143.3 (C)		143.4 (C)	
2'/6'	128.7 (CH)	7.24 (m)	127.2 (CH)	7.05-7.27 (m)	127.1 (CH)	7.05-7.25 (m)
3'/5'	129.7 (CH)	7.24 (m)	128.6 (CH)	7.05-7.27 (m)	128.6 (CH)	7.05-7.25 (m)
4'	126.9 (CH)	7.24 (m)	127.3 (CH)	7.05-7.27 (m)	127.2 (CH)	7.05-7.25 (m)
1"	23.0 (CH_3)	1.77 (s)	23.3 (CH_3)	1.80 (s)	23.3 (CH_3)	1.80 (s)
2"	169.6 (C)		168.8 (C)		168.8 (C)	
1"'			172.3 (C)		169.7 (C)	
2"'			49.8 (CH)	5.12 (m)	50.5 (CH)	5.09 (m)
3"'			41.3 (CH_2)	2.43-2.70 (m)	42.6 (CH_2)	2.40-2.61 (m)
1""			142.8 (C)		142.8 (C)	
2""/6""			126.8 (CH)	7.05-7.27 (m)	126.9 (CH)	7.05-7.25 (m)
3""/5""			128.7 (CH)	7.05-7.27 (m)	128.4 (CH)	7.05-7.25 (m)
4""			127.3 (CH)	7.05-7.27 (m)	127.2 (CH)	7.05-7.25 (m)
1*					174.8 (C)	
2*					50.6 (CH)	4.11 (m)
3*					40.6 (CH_2)	1.37 (m)
4*					24.4 (CH)	1.27 (m)
5*					23.6 (CH_3)	0.68 (d, 5.8)
6*					21.6 (CH_3)	0.78 (d, 5.9)
NH		8.15 (d, 7.8)		8.35 (d, 8.1)		8.38 (d, 8.5)
NH"'				8.39 (d, 8.5)		8.29 (d, 7.9)
NH*						8.08 (d, 8.2)