(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
20 September 2018 (20.09.2018)

WIPO | PCT

(10) International Publication Number
WO 2018/167270 A1

(51) International Patent Classification:
C07D 407/12 (2006.01) *A61P 35/00* (2006.01)
A61K 31/351 (2006.01)

(21) International Application Number:
PCT/EP2018/056665

(22) International Filing Date:
16 March 2018 (16.03.2018)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
17382140.6 17 March 2017 (17.03.2017) EP

(71) Applicant: **PHARMA MAR, S.A.** [ES/ES]; Avenida de los Reyes, 1, E-28770 Colmenar Viejo-Madrid (ES).

(72) Inventors: **CAÑEDO HERNÁNDEZ, Librada María**; PHARMA MAR, S.A., Avenida de los Reyes, 1, Polígono Industrial La Mina-Norte, E-28770 Colmenar Viejo-Madrid (ES). **DE LA CALLE VERDÚ, Fernando**; PHARMA MAR, S.A., Avenida de los Reyes, 1, Polígono Industrial La Mina-Norte, E-28770 Colmenar Viejo-Madrid (ES). **RODRÍGUEZ RAMOS, María Pilar**; PHARMA MAR, S.A., Avenida de los Reyes, 1, Polígono Industrial La Mina-Norte, E-28770 Colmenar Viejo-Madrid (ES). **SCHLEISSNER SÁNCHEZ, María del Carmen**; PHARMA MAR, S.A., Avenida de los Reyes, 1, Polígono Industrial La Mina-Norte, E-28770 Colmenar Viejo-Madrid (ES). **ZÚÑIGA GIRÓN, Paz**; PHARMA MAR, S.A., Avenida de los Reyes, 1 Polígono Industrial La Mina-Norte, E-28770 Colmenar Viejo-Madrid (ES).

(74) Agent: **BERNARDO NORIEGA, Francisco**; ABG PATENTES, S.L., Avda. de Burgos, 16D , Edificio EUROMOR, E-28036 Madrid (ES).

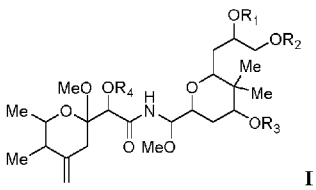
(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *with international search report (Art. 21(3))*
— *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
— *with (an) indication(s) in relation to deposited biological material furnished under Rule 13bis separately from the description (Rules 13bis.4(d)(i) and 48.2(a)(viii))*
— *with sequence listing part of description (Rule 5.2(a))*

WO 2018/167270 A1

(54) Title: ANTICANCER COMPOUNDS



I

(57) Abstract: Anticancer compounds of general formula (I), wherein R_1 to R_4 take various meanings, for use in the treatment of cancer. A novel *Labrenziasp.* strain named PHM005 with Accession Deposit Number CECT-9225, a method of producing compounds of the invention and analogues thereof by using the PHM005 strain and the Lab gene cluster codifying the biosynthesis of pederin-like and onnamide-like compounds are also provided.

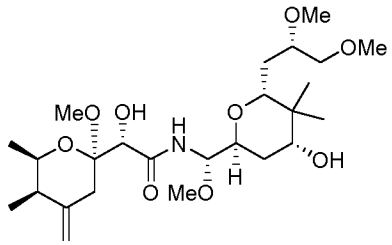
ANTICANCER COMPOUNDS

FIELD OF THE INVENTION

The present invention relates to direct or indirect production of anticancer compounds from bacteria and to new anticancer compounds, pharmaceutical compositions comprising them and their use as anticancer agents.

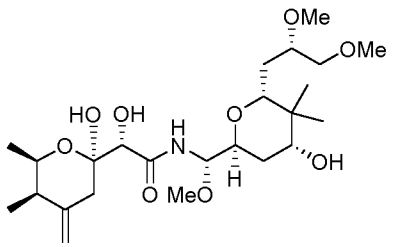
BACKGROUND OF THE INVENTION

In 1949, Ueta reported the isolation of the toxic principle from the beetle *Paederus fuscipes* (*Kyushu Igaku Zasshi,* 1949, 249). Four years later, a substance with identical physical properties from the same beetle species was also described by Pavan and Bo (*Physiol. Comp. Oecol.* 1953, 3, 307). The structure of this toxic compound, named pederin, was first proposed in 1965 by Cardani and co-workers (*Tetrahedron Lett.* 1965, 2537) but it was corrected in 1968 by Furusaki and co-workers based upon the crystal structure of a derivative. (*Tetrahedron Lett.* 1968, 6301). The structure of pederin is:

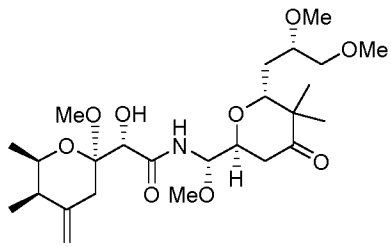


Pederin

Additionally, Cardani's group has reported the isolation of two additional compounds from *Paederus fuscipes* that were named pseudopederin and pederone. Pederone was described two years later (*Tetrahedron Lett.* 1967, *41*, 4023).



Pseudopederin



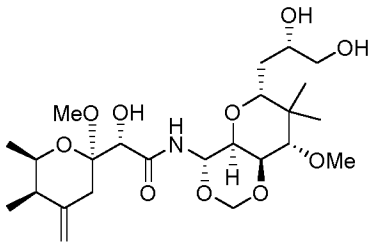
Pederone

Pederin is a potent cytotoxic and vesicant agent. Brega and co-workers (*J. Cell Biol.* 1968, 485-496) have tested pederin against a number of cell lines such as EUE, E6D, HeLa, KB, Hep, AS, MEF, CE, BHK, Z1 and M1 and have reported that concentrations of the order of 3 nM are

sufficient to cause cellular death within four days in all the cell lines analyzed. In addition pederin causes an immediate impairment of protein and DNA synthesis.

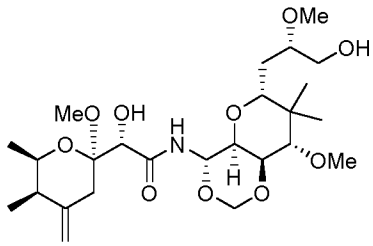
The cytotoxicity of pseudopederin has also been described by Soldati and co-workers (*Experientia* 1966, *3*, 176-178). Pseudopederin has toxicity lower than pederin, being active at doses 10 times higher.

European patent **EP0289203** discloses the isolation and antitumoral and antiviral activity of Mycalamide A, a compound isolated from *Mycale sp.* sponges collected in New Zealand.



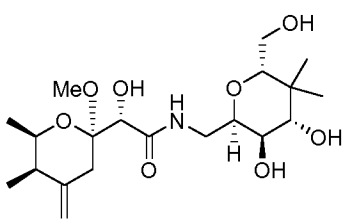
Mycalamide A

Its inventors, the Munro's group, further reported the isolation of Mycalamide B, a closely related compound with antitumoral and antiviral activity, from the same source (*J. Org. Chem.* 1990, *55*, 223).

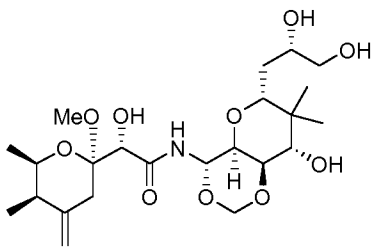


Mycalamide B

They further isolated two additional mycalamides, Mycalamides C and D, from *Stylinos* sponges (*J. Nat. Prod.* 2000, *63*, 704). Mycalamides A, B, C and D have IC_{50} values against the P-388 murine leukemia cell line of 3.0, 0.7, 95.0 and 35 ng/mL, respectively.



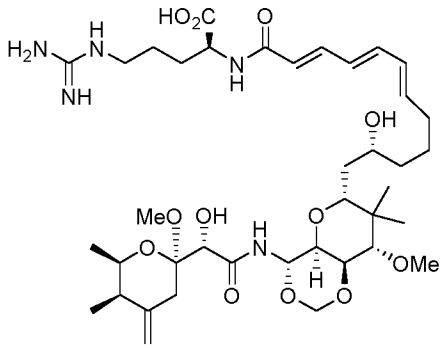
Mycalamide C



Mycalamide D

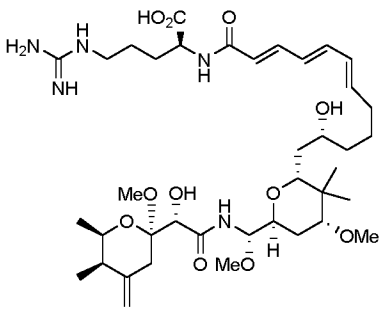
Mycalamides have also been shown to be powerful immunosuppressive agents with comparable in vitro efficacy to the clinical agent cyclosporine A.

US4801606 describes the isolation of Onnamide A from *Theonella sp.* samples collected off the coast of Japan. Onnamide A is an antitumoral compound with an IC_{50} value against the murine P388 cell line of 1 ng/mL. It also has antiviral activity.

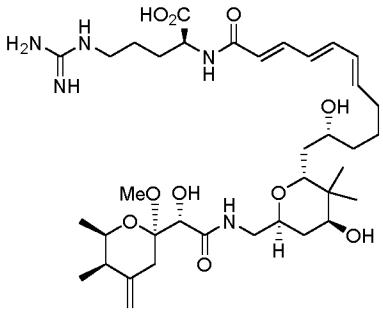


Onnamide A

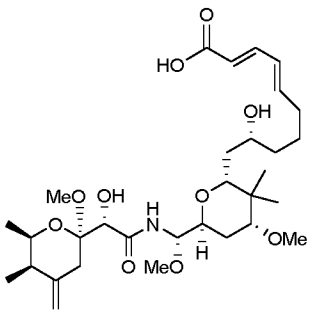
The onnamide family contains several members. Three of them, Onnamides D-F, lack of the dioxolane ring of onnamide A. Onnamides D and E were isolated from *Theonella* sponges by Matsunaga and co-workers (*Tetrahedron*, 1992, *48*, 8369) and Onnamide F was collected by the Capon group from the sponge *Trachycladus laevispirulifer* (*J. Nat. Prod.* 2001, *64*, 640).



Onnamide D



Onnamide E



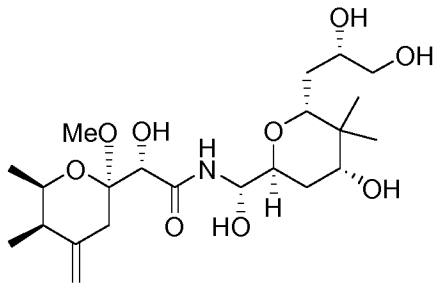
Onnamide F

Onnamide E did not show cytotoxic activity against the P388 cell line at a concentration of 0.4 µg/mL and Onnamide F has been described as a potent nematocide.

Experimental evidence for a bacterial biosynthesis of pederin was first provided by Kellner, who reported that the pederin-producing trait could be transferred to nonproducing *Paederus spp.* lines by feeding eggs of pederin-positive females (*Chemoecology,* 2001, *11*, 127). In contrast, eggs treated with antibiotics did not cause this effect. This result indicated the existence of a pederin-producing bacterium that is able to colonize the nonproducers (*J. Insect. Physiol.,* 2001, *47*, 475).

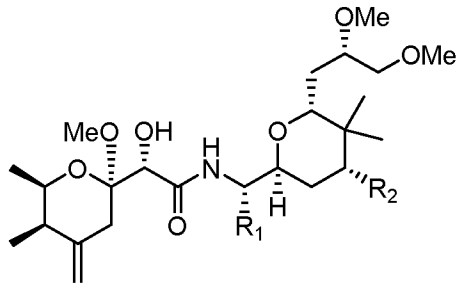
Piel and co-workers isolated the gene cluster for the polyketide synthase (PKS) of pederin (*Proc. Natl. Acad. Sci. U.S.A.*, 2002, *99*, 14002 and **WO2003044186**), and onnamides (*Proc. Natl. Acad. Sci. U.S.A.*, 2004, *101*, 16222). This work strongly implicated bacterial symbionts as the true sources of these compounds, which provides an explanation for the isolation of structurally similar compounds from disparate organisms. For a review about the symbiont proposal see Piel, J., *Curr. Med. Chem.* 2006, *13*, 39.

Another closely related compound, diaphorin, was isolated from the insect *Diaphorina citri* by Nakabachi and co-workers (*Current Biology* 2013, *23(15)*, 1478-1484). This compound is also cytotoxic with an IC_{50} value of *ca.* 1 μM and *ca.* 2 μM against B104 and HeLa cells, respectively. Its presence in extracts of *Diaphorina citri* was predicted in the same publication by the analysis of the polyketide synthase (PKS) system of Candidatus *Profftella armatura*, a defensive bacterial symbiont associated with *Diaphorina citri*.



Diaphorin

On the other hand, patent application **WO2013016120** describes a total synthesis of pederin and analogues thereof of formula:



wherein at least one of R_1 or R_2 includes a linker that includes a reactive functional group that can bind to a targeting moiety. This total synthesis is based on a multicomponent acyl aminal construction.

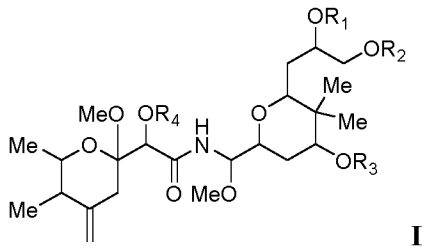
Detailed studies on the pharmacological properties of pederins, mycalamides and onnamides have been hampered by the scarcity of these compounds from natural sources. For example, approximately 100 kg of *Paederus fuscipes* were required to isolate sufficient material to elucidate

the structure of pederin. Although the PKS systems of pederins and onnamides have been described, it has not yet been possible to obtain these compounds by biotechnological methods. Therefore, the only practical way to obtain these interesting compounds at the moment is total synthesis. A number of total syntheses of pederin and mycalamides have been reported. They have been recently reviewed by Witezak and co-workers (*Mini Rev. Med. Chem.* 2012, *12(14)*, 1520-1532) and by Floreancig and Mosey (*Nat. Prod. Rep.* 2012, *29*, 980).

These syntheses have led to routes that are sufficiently brief to deliver sufficient material for biological testing and have provided analogues that have been useful in developing structure-activity relationships for these compounds. However, the need remains to provide a more concise route to these compounds and new antitumoral analogues thereof.

SUMMARY OF THE INVENTION

In a first aspect, the present invention is directed to a compound of general formula **I** or a pharmaceutically acceptable salt, tautomer, or stereoisomer thereof



I

wherein:

R_1, R_2, and R_3 are each independently selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, –C(=O)R_a, –C(=O)OR_b and –(C=O)N$R_c$$R_d$;

R_4 is selected from hydrogen, -C(=O)R_a, -C(=O)OR_b, and –C(=O)N$R_c$$R_d$;

R_a is selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl, and heterocyclyl;

R_b is selected from substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl, and heterocyclyl;

R_c and R_d are independently selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl and heterocyclyl;

with the proviso that R_1 and R_2 are not simultaneously methyl.

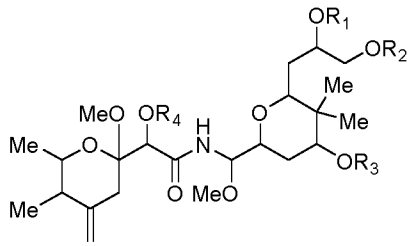
In a second aspect, the present invention is directed to pharmaceutical compositions comprising a compound of formula **I**, or a pharmaceutically acceptable salt, tautomer, or stereoisomer thereof, together with a pharmaceutically acceptable carrier or diluent.

In a third aspect, the present invention is directed to compounds of formula **I**, or pharmaceutically acceptable salts, tautomers, or stereoisomers thereof, for use as a medicament, in particular as a medicament for treating cancer.

In a fourth aspect, the present invention is directed to pharmaceutical compositions comprising a compound of formula **I**, for use as a medicament, in particular as a medicament for treating cancer.

In a fifth aspect, the present invention is also directed to the use of a compound of formula **I**, or a pharmaceutically acceptable salt, tautomer, or stereoisomer thereof, in the treatment of cancer, or in the preparation of a medicament, preferably for the treatment of cancer. Other aspects of the invention are methods of treatment, and compounds for use in these methods. Therefore, the present invention further provides a method of treating a patient, notably a human affected by cancer which comprises administering to said affected individual in need thereof a therapeutically effective amount of a compound as defined above.

In a sixth aspect, the present invention is directed to a process for obtaining a compound of formula **II** or a pharmaceutically acceptable salt, tautomer, or stereoisomer thereof



II

wherein

R_1, R_2, and R_3 are each independently selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, $-C(=O)R_a$, $-C(=O)OR_b$ and $-(C=O)NR_cR_d$;

R_4 is selected from hydrogen, $-C(=O)R_a$, $-C(=O)OR_b$, and $-C(=O)NR_cR_d$;

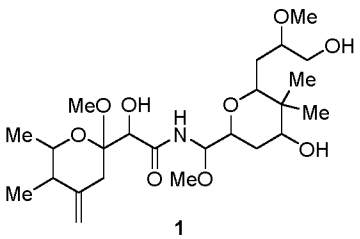
R_a is selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl, and heterocyclyl;

R_b is selected from substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl, and heterocyclyl;

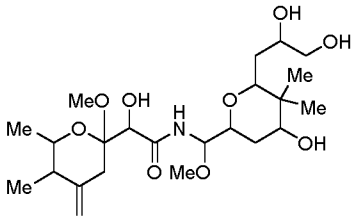
R_c and R_d are independently selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl and heterocyclyl;

the process comprising the steps of:

- culturing the wild type marine bacteria strain PHM005 or their mutants under suitable conditions to produce compounds **1** and/or **2** of formula:



1



2

;

- isolating compound **1** or **2**; and, if needed,
- derivatizing compound **1** or **2**.

In a seventh aspect, the present invention relates to strain PHM005. The free-living marine alphaproteobacteria producer of compounds **1** and **2** has been deposited for patent purposes in the CECT collection with the code CECT-9225.

In an eighth aspect, the present invention provides an isolated nucleic acid sequence comprising the **Lab biosynthetic gene cluster** or being complementary to a sequence comprising the Lab biosynthetic gene cluster. This gene cluster represents the first example of genes from a cultivable bacterium encoding the biosynthesis of pederin-like and onnamide-like compounds.

In a nineth aspect, the present invention provides nucleic acid fragments selected from the group consisting of genes *lab706*, *lab707*, *lab708*, *lab709*, *lab710*, *lab711*, *lab712*, *lab713*, *lab714*, *lab715*, *lab716*, *lab717*, *lab718*, *lab719*, *lab720*, *lab721*, *lab722*, *lab723*, *lab724*, *lab725* and/or *lab726* as shown in Figure 3.

In a tenth aspect, the invention is directed to a modular enzymatic system encoded by a nucleic acid sequence as described above. The modular enzymatic system preferably has functional

activity in the biosynthesis of pederin-like and onnamide-like compounds and/or a polyketide moiety and/or a nonribosomal peptide moiety.

In an eleventh aspect, the present invention is directed to a vector comprising a nucleic acid consisting essentially of the Lab biosynthetic gene cluster derived from *Labrenzia* sp. and in particular from strain PHM005 or a vector comprising a nucleic acid sequence as described above.

In a twelfth aspect the present invention is directed to a recombinant host cell or a transgenic organism comprising said nucleic acid or containing said vector.

In a thirteenth aspect the present invention is directed to a method for producing pederin-like or onnamide-like compounds using a mutant of PHM005 or a recombinant host cell or a transgenic organism as described above, comprising the step of:

- Culturing the mutant of PHM005 or the recombinant host cell or the transgenic organism under conditions to express the Lab biosynthetic gene cluster; and
- Isolating the produced pederin-like and/or onnamide-like compounds.

Other aspects of the present invention are directed to the use of a nucleic acid as defined above in the preparation of a modified Lab biosynthetic gene cluster, to the use of a nucleic acid as defined above in the preparation of a pederin-like or onnamide-like compound and to processes for improving production of pederin-like and onnamide-like compounds in bacteria comprising the steps of a) culturing strain PHM005 in the presence of a mutagenic agent for a period of time sufficient to allow mutagenesis, and b) selecting said mutants by a change of the phenotype that results in an increased production of pederin-like or onnamide-like compounds. The mutagenic agent may be a chemical agent, such as daunorubicin and nitrosoguanidine; a physical agent, such as gamma radiation or ultraviolet radiation; or a biological agent, such as a transposon, for example. Exemplary modifications include knock-out of tailoring genes to avoid methylations and hidroxylations.

BRIEF DESCRIPTION OF THE FIGURES AND THE SEQUENCES

Figure 1. Electron microscopy (negative staining) of *Labrenzia* sp. PHM005. Cells in the mid-exponential growth phase were adsorbed on 400-mesh carbon-collodion-coated grids for 2 min, negatively stained with 2% uranyl acetate, imaged with a Jeol JEM 1011 transmission electron microscope operated at 100 kV and photographed with a CCD Gatan Erlangshen ES1000W camera.

Figure 2. Neighbour-joining tree based on 16S rRNA gene sequences that show the relationship between PHM005 and the type strains of closely related species of the genera *Labrenzia* and *Stappia*. The phylogenetic tree was generated by the Pairwise alignment based similarity coefficient and UPGMA for Cluster analysis using BioNumerics V7.5 (Applied Maths). The phylogenetic neighbors were identified and pairwise 16S rDNA gene sequence similarities calculated by comparison with the SILVA LTPs123 database.

Figure 3. Map of **Lab** biosynthetic gene cluster. Total Lab gene cluster island: **69** Kb.

Figure 4. ^{1}H NMR spectrum of compound **1** in $CDCl_3$.

Figure 5. ^{13}C NMR spectrum of compound **1** in $CDCl_3$.

Figure 6. gCOSY spectrum of compound **1** in $CDCl_3$.

Figure 7. TOCSY spectrum of compound **1** in $CDCl_3$.

Figure 8. gHSQC spectrum of compound **1** in $CDCl_3$.

Figure 9. LR-HSQMBC spectrum of compound **1** in $CDCl_3$.

Figure 10. ROESY spectrum of compound **1** in $CDCl_3$.

The sequences mentioned in this application are listed in the attached sequence listing. These sequences are shortly summarized in the following:

SEQ ID NO: 1 Sequence (1355 bp) of 16S rRNA gene of *Labrenzia* sp. PHM005

SEQ ID NO: 2 nucleic acid sequence of the Lab biosynthetic gene cluster.

SEQ ID NO: 3 protein sequence of Lab706 putative acyl carrier protein.

SEQ ID NO: 4 protein sequence of Lab707 putative HMGS.

SEQ ID NO: 5 protein sequence of Lab708 PKS.

SEQ ID NO: 6 protein sequence of Lab709 TransAT PKS.

SEQ ID NO: 7 protein sequence of Lab710 putative acyl carrier protein.

SEQ ID NO: 8 protein sequence of Lab711 putative FAD oxigenase.

SEQ ID NO: 9 protein sequence of Lab712 putative oMethyltransferase.

SEQ ID NO: 10 protein sequence of Lab713 putative cytochrome P450.

SEQ ID NO: 11 protein sequence of Lab714 putative Malonyl CoA-ACP transacylase or FMT oxidoreductase.

SEQ ID NO: 12 protein sequence of Lab715 putative Malonyl CoA-ACP transacylase or acyltransferase.

SEQ ID NO: 13 protein sequence of Lab716 Malonyl CoA-ACP transacylase.

SEQ ID NO: 14 protein sequence of Lab717 Enoyl-CoA Hydratase.

SEQ ID NO: 15 protein sequence of Lab718 Beta-ketoacyl synthetase.

SEQ ID NO: 16 protein sequence of Lab719 TransAT PKS/NRPS.

SEQ ID NO: 17 protein sequence of Lab720 putative FAD monooxigenase.

SEQ ID NO: 18 protein sequence of Lab721, part of TransAT PKS.

SEQ ID NO: 19 protein sequence of Lab722, part of TransAT PKS.

SEQ ID NO: 20 protein sequence of Lab723, part of PKS.

SEQ ID NO: 21 protein sequence of Lab724, part of TransAT PKS/NRPS.

SEQ ID NO: 22 protein sequence of Lab725, part of PKS.

SEQ ID NO: 23 protein sequence of Lab726 putative oMethyltransferase.

DETAILED DESCRIPTION OF PREFERRED EMBODIMENTS

The present invention relates to compounds of general formula **I** as defined above.

In the compounds defined by Markush formulae in this specification, the groups can be selected in accordance with the following guidance:

Alkyl groups may be branched or unbranched, and preferably have from 1 to about 12 carbon atoms. One more preferred class of alkyl groups has from 1 to about 6 carbon atoms. Even more preferred are alkyl groups having 1, 2, 3 or 4 carbon atoms. Methyl, ethyl, *n*-propyl, isopropyl, and butyl, including *n*-butyl, *tert*-butyl, *sec*-butyl and isobutyl are particularly preferred alkyl groups in the compounds of the present invention. As used herein, the term alkyl, unless otherwise stated, refers to both cyclic and non-cyclic groups, although cyclic groups will comprise at least three carbon ring members.

Alkenyl and alkynyl groups in the compounds of the present invention may be branched or unbranched, have one or more unsaturated linkages and from 2 to about 12 carbon atoms. One more preferred class of alkenyl and alkynyl groups has from 2 to about 6 carbon atoms. Even more preferred are alkenyl and alkynyl groups having 2, 3 or 4 carbon atoms. The terms alkenyl and alkynyl as used herein refer to both cyclic and noncyclic groups, although cyclic groups will comprise at least three carbon ring atoms.

Suitable aryl groups in the compounds of the present invention include single and multiple ring compounds, including multiple ring compounds that contain separate and/or fused aryl groups. Typical aryl groups contain from 1 to 3 separated or fused rings and from 6 to about 18 carbon ring atoms. Preferably aryl groups contain from 6 to about 14 carbon ring atoms. Specially preferred aryl groups include substituted or unsubstituted phenyl, substituted or unsubstituted naphthyl,

substituted or unsubstituted biphenyl, substituted or unsubstituted phenanthryl and substituted or unsubstituted anthryl. The most preferred aryl group is substituted or unsubstituted phenyl.

Suitable heterocyclic groups include heteroaromatic and heteroalicyclic groups containing from 1 to 3 separated and/or fused rings and from 5 to about 18 ring atoms. Preferably heteroaromatic and heteroalicyclic groups contain from 5 to about 10 ring atoms, more preferably 5, 6 or 7 ring atoms. Suitable heteroaromatic groups in the compounds of the present invention contain one, two or three heteroatoms selected from N, O or S atoms and include, e.g., coumarinyl including 8-coumarinyl, quinolyl, including 8-quinolyl, isoquinolyl, pyridyl, pyrazinyl, pyrazolyl, pyrimidinyl, furyl, pyrrolyl, thienyl, thiazolyl, isothiazolyl, triazolyl, tetrazolyl, isoxazolyl, oxazolyl, imidazolyl, indolyl, isoindolyl, indazolyl, indolizinyl, phthalazinyl, pteridinyl, purinyl, oxadiazolyl, thiadiazolyl, furazanyl, pyridazinyl, triazinyl, cinnolinyl, benzimidazolyl, benzofuranyl, benzofurazanyl, benzothiophenyl, benzothiazolyl, benzoxazolyl, quinazolinyl, quinoxalinyl, naphthyridinyl, and furopyridyl. Suitable heteroalicyclic groups in the compounds of the present invention contain one, two or three heteroatoms selected form N, O or S atoms and include, e.g., pyrrolidinyl, tetrahydrofuranyl, tetrahydrothienyl, tetrahydrothiopyranyl, piperidyl, morpholinyl, thiomorpholinyl, thioxanyl, piperazinyl, azetidinyl, oxetanyl, thietanyl, homopiperidyl, oxepanyl, thiepanyl, oxazepinyl, diazepinyl, thiazepinyl, 1,2,3,6-tetrahydropyridil, 2-pyrrolinyl, 3-pyrrolinyl, indolinyl, 2H-pyranyl, 4H-pyranyl, dioxanyl, 1,3-dioxolanyl, pyrazolinyl, dithianyl, dithiolanyl, dihydropyranyl, dihydrothienyl, dihydrofuranyl, pyrazolidinyl, imidazolinyl, imidazolidinyl, 3-azabicyclo[3.1.0]hexyl, 3-azabicyclo[4.1.0]heptyl, 3H-indolyl, and quinolizinyl.

The groups above mentioned may be substituted at one or more available positions by one or more suitable groups such as OR', =O, SR', SOR', SO_2R', OSO_2R', NO_2, NHR', NR'R', =N-R', N(R')COR', $N(COR')_2$, $N(R')SO_2R$, N(R')C(=NR')N(R')R', CN, halogen, COR' COOR', OCOR', OCOOR', OCONHR', OCON(R')R', CON(R')R', CON(R')OR', $CON(R')SO_2R'$, $PO(OR')_2$, PO(OR')R', PO(OR')(N(R')R'), protected OH, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, substituted or unsubstituted aryl, and substituted or unsubstituted heterocyclic group, wherein each of the R' groups is independently selected from the group consisting of hydrogen, OH, NO_2, NH_2, SH, CN, halogen, COH, COalkyl, COOH, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, substituted or unsubstituted aryl, and substituted or unsubstituted heterocyclic group. Where such groups are themselves substituted, the substituents may be chose from the foregoing list.

Suitable halogen groups or substituents in the compounds of the present invention include F, Cl, Br, and I.

Suitable protecting groups for OH, including protecting groups for 1,2-diols, are well known for the person skilled in the art. A general review of protecting groups in organic chemistry is provided by Wuts, PGM and Greene TW in Protecting Groups in Organic Synthesis 4th Ed. Wiley-Interscience, and by Kocienski PJ in Protecting Groups, 3rd Ed. Georg Thieme Verlag. These references provide sections on protecting groups for OH. All these references are incorporated by reference in their entirely.

Within the scope of the present invention an OH protecting group is defined to be the *O*-bonded moiety resulting from the protection of the OH group through the formation of a suitable protected OH group. Examples of such protected OH include ethers, silyl ethers, esters, sulfonates, sulfenates and sulfinates, carbonates and carbamates. In the case of ethers the protecting group for the OH can be selected from methyl, methoxymethyl, methylthiomethyl, (phenyldimethylsilyl)methoxymethyl, benzyloxymethyl, *p*-methoxybenzyloxymethyl, [(3,4-dimethoxybenzyl)oxy]methyl, *p*-nitrobenzyloxymethyl, *o*-nitrobenzyloxymethyl, [(*R*)-1-(2-nitrophenyl)ethoxy]methyl, (4-methoxyphenoxy)methyl, guaiacolmethyl, [(*p*-phenylphenyl)-oxy]methyl, *t*-butoxymethyl, 4-pentenyloxymethyl, siloxymethyl, 2-methoxyethoxymethyl, 2-cyanoethoxymethyl, bis(2-chloroethoxy)methyl, 2,2,2-trichoroethoxymethyl, 2-(trimethylsilyl)-ethoxymethyl, menthoxymethyl, *O*-bis(2-acetoxyethoxy)methyl, tetrahydropyranyl, fluorous tetrahydropyranyl, 3-bromotetrahydropyranyl, tetrahydrothiopyranyl, 1-methoxycyclohexyl, 4-methoxytetrahydropyranyl, 4-methoxytetrahydrothiopyranyl, 4-methoxytetrahydrothiopyranyl *S,S*-dioxide, 1-[(2-chloro-4-methyl)phenyl]-4-methoxypiperidin-4-yl, 1-(2-fluorophenyl)-4-methoxypiperidin-4-yl, 1-(4-chlorophenyl)-4-methoxypiperidin-4-yl, 1,4-dioxan-2-yl, tetrahydrofuranyl, tetrahydrothiofuranyl, 2,3,3*a*,4,5,6,7,7*a*-octahydro-7,8,8-trimethyl-4,7-methanobenzofuran-2-yl, 1-ethoxyethyl, 1-(2-chloroethoxy)ethyl, 2-hydroxyethyl, 2-bromoethyl, 1-[2-(trimethylsilyl)ethoxy]ethyl, 1-methyl-1-methoxyethyl, 1-methyl-1-benzyloxyethyl, 1-methyl-1-benzyloxy-2-fluoroethyl, 1-methyl-1-phenoxyethyl, 2,2,2-trichloroethyl, 1,1-dianisyl-2,2,2-trichloroethyl, 1,1,1,3,3,3-hexafluoro-2-phenylisopropyl, 1-(2-cyanoethoxy)ethyl, 2-trimethylsilylethyl, 2-(benzylthio)ethyl, 2-(phenylselenyl)ethyl, *t*-butyl, cyclohexyl, 1-methyl-1'-cyclopropylmethyl, allyl, prenyl, cinnamyl, 2-phenallyl, propargyl, *p*-chlorophenyl, *p*-methoxyphenyl, *p*-nitrophenyl, 2,4-dinitrophenyl, 2,3,5,6-tetrafluoro-4-(trifluoromethyl)phenyl, benzyl, *p*-methoxybenzyl, 3,4-dimethoxybenzyl, 2,6-dimethoxybenzyl, *o*-nitrobenzyl, *p*-nitrobenzyl, pentadienylnitrobenzyl, pentadienylnitropiperonyl, halobenzyl, 2,6-dichlorobenzyl, 2,4-dichlorobenzyl, 2,6-difluorobenzyl, *p*-cyanobenzyl, fluorous benzyl, 4-fluorousalkoxybenzyl, trimethylsilylxylyl, *p*-phenylbenzyl, 2-phenyl-2-propyl, *p*-acylaminobenzyl, *p*-azidobenzyl,

4.azido-3-chlorobenzyl, 2-trifluoromethylbenzyl, 4-trifluoromethylbenzyl, *p*-(methylsulfinyl)benzyl, *p*-siletanylbenzyl, 4-acetoxybenzyl, 4-(2-trimethylsilyl)ethoxymethoxybenzyl, 2-naphthylmethyl, 2-picolyl, 4-picolyl, 3-methyl-2-picolyl *N*-oxide, 2-quinolinylmethyl, 6-methoxy-2-(4-methylphenyl-4-quinolinemethyl, 1-pyrenylmethyl, diphenylmethyl, 4-methoxydiphenylmethyl, 4-phenyldiphenylmethyl, *p,p'*-dinitrobenzhydryl, 5-dibenzosuberyl, triphenylmethyl, tris(4-*t*-butylphenyl)methyl, α-naphthyldiphenylmethyl, *p*-methoxyphenyldiphenylmethyl, di(*p*-methoxyphenyl)phenylmethyl, tri(*p*-methoxyphenyl)methyl, 4-(4'-bromophenacyloxy)phenyldiphenylmethyl, 4,4',4''-tris(4,5-dichlorophthalimidophenyl)methyl, 4,4'4''-tris(levulinoyloxyphenyl)methyl, 4,4',4''-tris(benzoyloxyphenyl)methyl, 4,4'-dimethoxy-3''-[*N*-(imidazolylmethyl)]trityl, 4,4'-dimethoxy-3''-[*N*-(imidazolylethyl)carbamoyl]trityl, bis(4-methoxyphenyl)-1'-pyrenylmethyl, 4-(17-tetrabenzo[*a,c,g,i*]fluorenylmethyl)-4,4''-dimethoxytrityl, 9-anthryl, 9-(9-phenyl)xan-thenyl, 9-phenylthioxanthyl, 9-(9-phenyl-10-oxo)anthryl, 1,3-benzodithiolan-2-yl, and 4,5-bis(ethoxycarbonyl)-[1,3]-dioxolan-2-yl, benzisothiazolyl *S,S*-dioxide. In the case of silyl ethers the protecting group for the OH can be selected from trimethylsilyl, triethylsilyl, triisopropylsilyl, dimethylisopropylsilyl, diethylisopropylsilyl, dimethylhexylsilyl, 2-norbornyldimethylsilyl, *t*-butyldimethylsilyl, *t*-butyldiphenylsilyl, tribenzylsilyl, tri-*p*-xylylsilyl, triphenylsilyl, diphenylmethylsilyl, di-*t*-butylmethylsilyl, bis-(*t*-butyl)-1-pyrenylmethoxysilyl, tris(trimethylsilyl)silyl, (2-hydroxystyryl)dimethylsilyl, (2-hydroxystyryl)diisopropylsilyl, *t*-butylmethoxyphenylsilyl, t-butoxydiphenylsilyl, 1,1,3,3-tetraisopropyl-3-[2-(triphenylmethoxy)ethoxy]disiloxane-1-yl, and fluorous silyl. In the case of esters the protecting group for the OH together with the oxygen atom of the unprotected OH to which it is attached form an ester that can be selected from formate, benzoylformate, acetate, chloroacetate, dichloroacetate, trichloroacetate, trichloroacetamidate, trifluoroacetate, methoxyacetate, triphenylmethoxyace-tate, phenoxyacetate, *p*-chlorophenoxyacetate, phenylacetate, diphenylacetate, 3-phenylpropionate, bisfluorous chain type propanoyl, 4-pentenoate, 4-oxopentanoate, 4,4-(ethylenedithio)pentanoate, 5-[3-bis(4-methoxyphenyl)hydroxymethylphenoxy]levulinate, pivaloate, 1-adamantoate, crotonate, 4-methoxycrotonate, benzoate, *p*-phenylbenzoate, 2,4,6-trimethylbenzoate, 4-bromobenzoate, 2,5-difluorobenzoate, *p*-nitrobenzoate, picolinate, nicotinate, 2-(azidomethyl)benzoate, 4-azidobutyrate, (2-azidomethyl)phenylacetate, 2-{[tritylthio)oxy]methyl}benzoate, 2-{[(4-methoxytritylthio)oxy]methyl}benzoate, 2-{[methyl(tritylthio)amino]methyl}benzoate, 2-{{[(4-methoxytrityl)thio]methylamino]-methyl}benzoate, 2-(allyloxy)phenylacetate, 2-(prenyloxymethyl)benzoate, 6-(levulinyloxymethyl)-3-methoxy-2-nitrobenzoate, 6-(levulinyloxymethyl)-3-methoxy-4-nitrobenzoate, 4-benzyloxybutyrate, 4-trialkylsilyloxybutyrate, 4-acetoxy-2,2-dimethylbutyrate, 2,2-dimethyl-4-pentenoate, 2-iodobenzoate, 4-nitro-4-methylpentanoate, *o*-(dibromomethyl)benzoate, 2-formylbenzenesulfonate, 4-

(methylthiomethoxy)butyrate, 2-(methylthiomethoxymethyl)benzoate, 2-(chloroacetoxymethyl)benzoate, 2-[(2-choroacetoxy)ethyl]benzoate, 2-[2-benzyloxy)ethyl]benzoate, 2-[2-(4-methoxybenzyloxy)ethyl]benzoate, 2,6-dichloro-4-methylphenoxyacetate, 2,6-dichloro-4-(1,1,3,3-tetramethylbutyl)phenoxyacetate, 2,4-bis(1,1-dimethylpropyl)phenoxyacetate, chlorodiphenylacetate, isobutyrate, monosuccinoate, (*E*)-2-methyl-2-butenoate, *o*-(methoxycarbonyl)benzoate, α-naphthoate, nitrate, alkyl *N,N,N',N'*-tetramethylphosphorodiamidate, and 2-chlorobenzoate. In the case of sulfonates, sulfenates and sulfinates the protecting group for the OH together with the oxygen atom of the unprotected OH to which it is attached form a sulfonate, sulfenate or sulfinate that can be selected from sulfate, allylsulfonate, methanesulfonate, benzylsulfonate, tosylate, 2-[(4-nitrophenyl)ethyl]sulfonate, 2-trifluoromethylbenzenesulfonate, 4-monomethoxytritylsulfenate, alkyl 2,4-dinitrophenylsul-fenate, 2,2,5,5-tetramethylpyrrolidin-3-one-1-sulfinate, and dimethylphosphinothiolyl. In the case of carbonates the protecting group for the OH together with the oxygen atom of the unprotected OH to which it is attached form a carbonate that can be selected from methyl carbonate, methoxymethyl carbonate, 9-fluorenylmethyl carbonate, ethyl carbonate, bromoethyl carbonate, 2-(methylthiomethoxy)ethyl carbonate, 2,2,2-trichloroethyl carbonate, 1,1-dimethyl-2,2,2-trichloroethyl carbonate, 2-(trimethylsilyl)ethyl carbonate, 2-[dimethyl(2-naphthylmethyl)silyl]ethyl carbonate, 2-(phenylsulfonyl)ethyl carbonate, 2-(triphenylphos-phonio)ethyl carbonate, *cis*-[4-[[(methoxytrityl)sulfenyl]oxy]tetrahydrofuran-3-yl]oxy carbonate, isobutyl carbonate, *t*-butyl carbonate, vinyl carbonate, allyl carbonate, cinnamyl carbonate, propargyl carbonate, *p*-chlorophenyl carbonate, *p*-nitrophenyl carbonate, 4-ethoxy-1-naphthyl carbonate, 6-bromo-7-hydroxycoumarin-4-ylmethyl carbonate, benzyl carbonate, *o*-nitrobenzyl carbonate, *p*-nitrobenzyl carbonate, *p*-methoxybenzyl carbonate, 3,4-dimethoxybenzyl carbonate, anthraquinon-2-ylmethyl carbonate, 2-dansylethyl carbonate, 2-(4-nitrophenyl)ethyl carbonate, 2-(2,4-dinitrophenyl)ethyl carbonate, 2-(2-nitrophenyl)propyl carbonate, alkyl 2-(3,4-methylenedioxy-6-nitrophenyl)propyl carbonate, 2-cyano-1-phenylethyl carbonate, 2-(2-pyridylamino-1-phenylethyl carbonate, 2-[*N*-methyl-*N*-(2-pyridyl)]amino-1-phenylethyl carbonate, phenacyl carbonate, 3',5'-dimethoxybenzoin carbonate, methyl dithiocarbonate, and *S*-benzyl thiocarbonate. And in the case of carbamates the protecting group for the OH together with the oxygen atom of the unprotected OH to which it is attached forms a carbamate that can be selected from dimethylthiocarbamate, *N*-phenylcarbamate, and *N*-methyl-*N*-(*o*-nitrophenyl)carbamate.

Within the scope of the present invention an 1,2-diol protecting group is defined to be the O-bonded moiety resulting from the simultaneous protection of the 1,2-diol through the formation of a protected 1,2-diol. Examples of such protected 1,2-diols include cyclic acetals and ketals, cyclic ortho esters, silyl derivatives, dialkylsilylene derivatives, cyclic carbonates, cyclic boronates.

Examples of cyclic acetals and ketals include methylene acetal, ethylidene acetal, *t*-butylmethylidene acetal, 1-*t*-buylethylidene ketal, 1-phenylethylidene ketal, 2-(methoxycarbonyl)ethylidene (Mocdene) acetal, or 2-(*t*-butylcarbonyl)ethylidene (Bocdene) acetal, phenylsulfonylethylidene acetal, 2,2,2-trichloroethylidene acetal, 3-(benzyloxy)propyl acetal, acrolein acetal, acetonide (isopropylidene ketal), cyclopentylidene ketal, cyclohexylidene ketal, cycloheptylidene ketal, benzylidene acetal, *p*-methoxybenzylidene acetal, 1-(4-methoxyphenyl)ethylidene ketal, 2,4-dimethoxybenzylidene acetal, 3,4-dimethoxybenzylidene acetal, *p*-acetoxybenzylidene acetal, 4-(*t*-butyldimethylsilyloxy)benzylidene acetal, 2-nitrobenzylide acetal, 4-nitrobenzylidene acetal, mesitylene acetal, 6-bromo-7-hydroxycoumarin-2-ylmethylidene acetal, 1-naphthaladehyde acetal, 2-naphthaldehyde acetal, 9-anthracene acetal, benzophenone ketal, di-(*p*-anisyl)methylidene acetal, xanthen-9-ylidene ketal, 2,7-dimethylxanthen-9-ylidene ketal, diphenylmethylene ketal, camphor ketal, and menthone ketal. Examples of cyclic ortho esters include methoxymethylene acetal, ethoxymethylene acetal, 2-oxacyclopentylidene ortho ester, dimethoxymethylene ortho ester, 1-methoxyethylidene ortho ester, 1-ethoxyethylidene ortho ester, phthalide ortho ester, 1,2-dimethoxyethylidene ortho ester, α-methoxybenzylidene ortho ester, 1-(*N*,*N*-dimethylamino)ethylidene derivative, α-(*N*,*N*-dimethylamino)benzylidene derivative, butane 2-3-bisacetal (BBA), cyclohexane-1,2-diacetal (CDA), and dispiroketals. Examples of silyl derivatives include di-*t*-butylsilylene group (DTBS$(OR)_2$), 1-(cyclohexyl)-1-(methyl)silylene (Cy)(Me)Si$(OR)_2$, di-isopropylsilylene (*i*-propyl)$_2$Si$(OR)_2$, dicyclohexylsilylene (Cy)$_2$Si$(OR)_2$, 1,3-(1,1,3,3-tetraisopropyldisiloxanylidene) derivative (TIPDS$(OR)_2$), 1,1,3,3-tetra-*t*-butoxydisiloxanylidene derivative (TBDS$(OR)_2$), methylene-bis-(diisopropylsilanoxanylidene) (MDPS$(OR)_2$), and 1,1,4,4-tetraphenyl-1,4-disilanylidene (SIBA$(OR)_2$). Examples of cyclic boronates include methyl boronate, ethyl boronate, phenyl boronate, and *o*-acetamidophenyl boranate.

The mention of these groups should not be interpreted as a limitation of the scope of the invention, since they have been mentioned as a mere illustration of protecting groups for OH, but further groups having said function may be known by the skilled person in the art, and they are to be understood to be also encompassed by the present invention.

The term “pharmaceutically acceptable salt” refers to any pharmaceutically acceptable salt which, upon administration to the patient is capable of providing (directly or indirectly) a compound as described herein. However, it will be appreciated that non-pharmaceutically acceptable salts also fall within the scope of the invention since those may be useful in the preparation of pharmaceutically acceptable salts. The preparation of salts can be carried out by methods known in the art.

For instance, pharmaceutically acceptable salts of compounds provided herein are synthesized from the parent compound, which contains a basic or acidic moiety, by conventional chemical methods. Generally such salts are, for example, prepared by reacting the free acid or base forms of these compounds with a stoichiometric amount of the appropriate base or acid in water or in an organic solvent or in a mixture of both. Generally, nonaqueous media like ether, ethyl acetate, ethanol, 2-propanol or acetonitrile are preferred. Examples of the acid addition salts include mineral acid addition salts such as, for example, hydrochloride, hydrobromide, hydroiodide, sulfate, nitrate, phosphate, and organic acid addition salts such as, for example, acetate, trifluoroacetate, maleate, fumarate, citrate, oxalate, succinate, tartrate, malate, mandelate, methanesulfonate and *p*-toluenesulfonate. Examples of the alkali addition salts include inorganic salts such as, for example, sodium, potassium, calcium and ammonium salts, and organic alkali salts such as, for example, ethylenediamine, ethanolamine, *N,N*-dialkylenethanolamine, triethanolamine and basic aminoacids salts.

The compounds of the invention may be in crystalline or amorphous form either as free compounds or as solvates (e.g. hydrates, alcoholates, particularly methanolates) and it is intended that any of these forms are within the scope of the present invention. Methods of solvation are generally known within the art. The compounds of the invention may present different polymorphic forms, and it is intended that the invention encompasses all such forms.

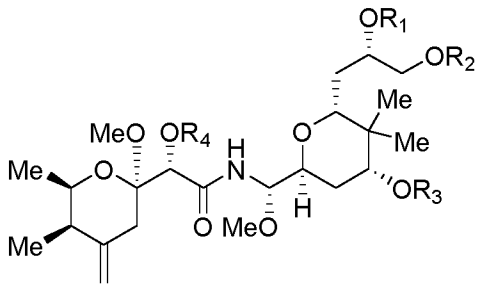
Any compound referred to herein is intended to represent such specific compound as well as certain variations or forms. In particular, compounds referred to herein may have asymmetric centers and therefore exist in different enantiomeric or diastereomeric forms. Thus, any given compound referred to herein is intended to represent any one of a racemate, one or more enantiomeric forms, one or more diastereomeric forms, and mixtures thereof. Likewise, stereoisomerism or geometric isomerism about the double bond is also possible, therefore in some cases the molecule could exist as (*E*)-isomer or (*Z*)-isomer (*trans* and *cis* isomers). If the molecule contains several double bonds, each double bond will have its own stereoisomerism, that could be the same, or different to, the stereoisomerism of the other double bonds of the molecule. Furthermore, compounds referred to herein may exist as atropisomers. All the stereoisomers including enantiomers, diastereoisomers, geometric isomers and atropisomers of the compounds referred to herein, and mixtures thereof, are considered within the scope of the present invention.

Furthermore, any compound referred to herein may exist as tautomers. Specifically, the term tautomers refer to one of two or more structural isomers of a compound that exist in equilibrium and are readily converted from one isomeric form to another. Common tautomeric pairs are amine-imine, amide-imidic acid, keto-enol, lactam-lactim, etc.

Unless otherwise stated, the compounds of the invention are also meant to include isotopically-labelled forms i.e. compounds which differ only in the presence of one or more isotopically-enriched atoms. For example, compounds having the present structures except for the replacement of at least one hydrogen atom by a deuterium or tritium, or the replacement of at least one carbon atom by ^{13}C- or ^{14}C-enriched carbon, or the replacement of at least one nitrogen atom by ^{15}N-enriched nitrogen are within the scope of this invention.

To provide a more concise description, some of the quantitative expressions given herein are not qualified with the term "about". It is understood that, whether the term "about" is used explicitly or nor, every quantity given herein is meant to refer to the actual given value, and it is also meant to refer to the approximation to such given value that would reasonably be inferred based on the ordinary skill in the art, including equivalents and approximations due to the experimental and/or measurement conditions for such given value.

More particularly, preferred compounds of formula **I** are those also having general formula **III** or a pharmaceutically acceptable salt, tautomer, and stereoisomer thereof



III

wherein R_1, R_2, R_3 and R_4 are as defined above in general formula **I**.

In compounds of general formula **I** and **III**, particularly preferred R_1 is selected from hydrogen and substituted or unsubstituted C_1-C_{12} alkyl. More preferably R_1 is selected from hydrogen and substituted or unsubstituted C_1-C_6 alkyl. Even more preferably, R_1 is selected from hydrogen, methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *tert*-butyl, *sec*-butyl, and isobutyl. Most preferred R_1 are hydrogen and methyl.

In compounds of general formula **I** and **III**, particularly preferred R_2 is selected from hydrogen and $-C(=O)R_a$, wherein R_a is substituted or unsubstituted C_1-C_{12} alkyl. More preferred R_a is substituted or unsubstituted C_1-C_6 alkyl. Even more preferably R_a is selected from methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *tert*-butyl, *sec*-butyl and isobutyl. Most preferred R_2 are hydrogen and acetyl.

In compounds of general formula **I** and **III**, particularly preferred R_3 and R_4 are independently selected from hydrogen and $-C(=O)R_a$, wherein R_a at each occurrence is independently selected from substituted or unsubstituted C_1-C_{12} alkyl. More preferably R_a at each occurrence is independently selected from substituted or unsubstituted C_1-C_6 alkyl. Even more preferably, R_a at each occurrence is independently selected from methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *tert*-butyl, *sec*-butyl and isobutyl. Most preferably R_3 and R_4 are independently selected from hydrogen and acetyl.

In additional preferred embodiments, the preferences described above for the different substituents are combined. The present invention is also directed to such combinations of preferred substitutions in the general formula **I** and **III** above.

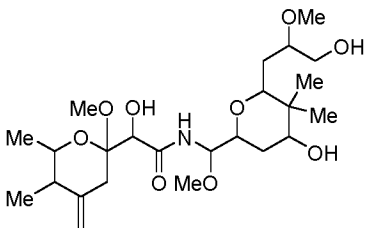
In one embodiment, R_1 is selected from substituted or unsubstituted C_1-C_6 alkyl and R_2 is hydrogen.

In another embodiment, R_1 is selected from substituted or unsubstituted C_1-C_6 alkyl and R_2 is $-C(=O)R_a$, wherein R_a is substituted or unsubstituted C_1-C_{12} alkyl.

In a further embodiment, both R_1 and R_2 are hydrogen.

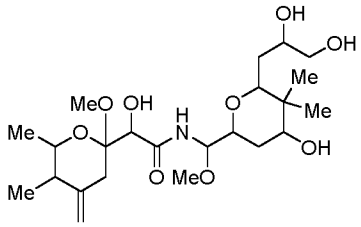
In the present description and definitions, when there are several groups R_a, R_b, R_c, R_d or R' present in the compounds of the invention, and unless it is stated explicitly so, it should be understood that they can be each independently different within the given definition, i.e. R_a does not represent necessarily the same group simultaneously in a given compound of the invention.

Particularly preferred compounds of the invention are the following



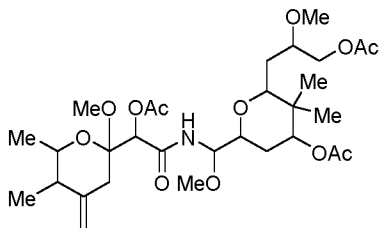
1

,



2

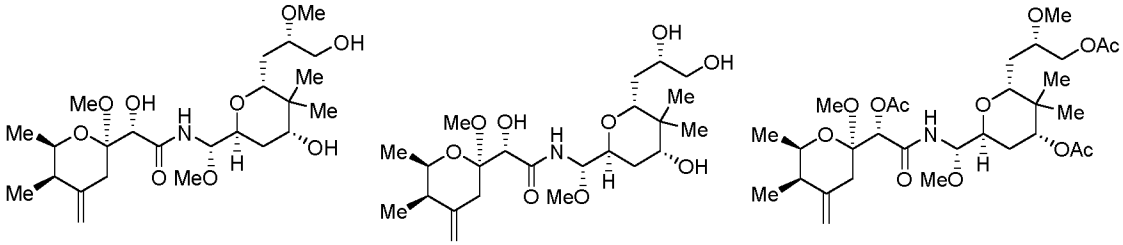
,



3

or pharmaceutically acceptable salts, tautomers or stereoisomers thereof.

Most preferred compounds of the invention are the following:



1' , 2' , 3'

or pharmaceutically acceptable salts, tautomers or stereoisomers thereof.

Compounds **1** and **2** were isolated from *Labrenzia sp.*, named strain PHM005. This alphaproteobacteria was isolated from a marine sediment sample collected in the Indian Ocean. Observation of the cells by transmission electron microscopy (Figure 1) allowed the identification of motile rods (0.6-0.8 μm wide and 1.6-2.1 μm long) with single, subpolar inserted flagella. A culture of this strain has been deposited in the CECT ("Colección Española de Cultivos Tipo") at the University of Valencia, Spain, under the accession number CECT-9225. This deposit has been made under the provisions of the Budapest Treaty.

The bacteria is clearly marine salt dependent since it needs more than 2.5% NaCl to grow, with the optimal concentration of marine salt for production of **1** being 36 g/L, similar to ocean conditions. Colonies on Marine Agar 2216 (DIFCO) are beige, almost transparent, smooth and with entire margin. After three weeks the colonies become darker-brown, maybe due to the effect of bacteriochlorophyll a and carotenoid production, as described for *Labrenzia alexandrii* DFL-11^T (Biebl and co-workers, *Evol, Microbiol*, **2007,** *57*, 1095-1107).

For the isolation of the producer microorganism, all the manipulations were carried out in aseptic conditions. PHM005 was isolated from a sediment frozen sample spread directly on Petri dishes with a sea salt medium of the following composition (g/L): marine salts (Tropic Marin® PRO-REEF, 27; agar, 16; supplemented with cycloheximide 0.2 mg/mL. The plates were incubated at 28 °C for three weeks under atmospheric pressure. After this period a slightly brown colony was picket and transferred to the same sea salt medium to confirm the purity and start taxonomy and fermentation studies.

A taxonomic evaluation of PHM005 was conducted by partial sequence of 16S rRNA following standard procedures. PHM005 was grown in marine broth (DIFCO 1196) for 72 hours. Cells were recovered and lysed by boiling with 4% NP40 for 10 minutes. Cell debris was discarded by centrifugation. The 16S rRNA was amplified by the polymerase chain reaction using the bacterial primers F1 and R5 described by Cook and Myers (*International Journal of Systematics and*

Evolutionary Microbiology, 2003, *53*, 1907-1915. The nearly full-length 16S rRNA gene sequence obtained is shown in SEQ NO: 1.

The phylogenetic tree was generated by the Pairwise alignment based similarity coefficient and UPGMA for Cluster analysis using BioNumerics V7.5 for Cluster Analysis. The phylogenetic neighbors were identified and pairwise 16S rRNA gene sequence similarities calculated by comparison with the SILVA LTPs123 database. The phylogenetic tree is shown in Figure 2.

PHM005 produces compounds **1** and **2** when it is cultured under controlled conditions in a suitable medium. This strain clearly needs marine salt to grow. This strain is preferably grown in a conventional aqueous nutrient medium. The culture must be driven in aerobic conditions and the production of compounds **1** and **2** should start after 3 days of growth controlling temperature between 26-28ºC. Conventional fermentation tanks have been found to be well suited for carrying out the cultivation of this organism. The addition of nutrients and pH control as well as antifoaming agents during the different stages of fermentation may be needed for increasing production and avoid foaming.

Compounds of the present invention can be produced starting from a colony or a frozen pure culture of strain PHM005 for developing enough biomass. This step may be repeated several times as needed and the material collected will be used as an inoculum to seed one or several fermentation flasks or tanks with the appropriate culture medium. These flasks or tanks can be used for developing the inoculum or for the production stage, depending on the broth volume needed. Sometimes the production medium may be different from the ones used for inoculum development.

Compounds of the present invention can be isolated from the fermentation broth mainly from cells and from the supernatant of strain PHM005 by extraction with a suitable mixture of solvents or absorbing in adequate resins.

Separation and purification of the present invention from the crude active extract can be performed using the proper combination of conventional chromatographic techniques.

Additionally, compounds of the invention can be obtained by modifying those already obtained from the natural source or by further modifying those already modified by using a variety of chemical reactions. Thus, hydroxyl groups can be acylated by standard coupling or acylation procedures, for instance by using acetyl chloride or acetic anhydride in pyridine or the like. Formate groups can be obtained by reacting the corresponding alkoxyde with acetic formic anhydride. Carbamates can be obtained by heating hydroxyl precursors with isocyanates. Carbonates can be obtained by using the corresponding anhydride and an activator such as

$Mg(ClO_4)_2$ or $Zn(OAc)_2$, Hydroxyl groups can also be converted into alkoxy groups by alkylation using an alkyl bromide iodide or sulfonate or into amino lower alkoxy groups by using, for instance, a protected 2-bromoethylamine. When necessary, appropriate protecting groups can be used on the substituents to ensure that reactive groups are not affected and to all selective functionalization of the hydroxyl groups. The procedures and reagents needed to prepare these derivatives are known to the skilled person and can be found in general textbooks such as March's Advanced Organic Chemistry 7th Edition 2013, Wiley Interscience.

An important feature of the above described compounds of formula **I** and **III** is their bioactivity and in particular their cytotoxic activity against tumor cells. Thus, with this invention we provide pharmaceutical compositions of compounds of general formula **I** and **III**, or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof that possess cytotoxicity activities and their use as anticancer agents. The present invention further provides pharmaceutical compositions comprising a compound of general formula **I** and **III**, or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof, with a pharmaceutically acceptable carrier or diluent.

Examples of pharmaceutical compositions include any solid (tablet, pills, capsules, granules, powder for vials, etc.) or liquid (solutions, suspensions or emulsions) composition for oral, topical or parenteral administration.

Administration of the compounds or compositions of the present invention may be by any suitable method, such as intravenous infusion, oral preparations, and intraperitoneal and intravenous administration. We prefer that infusion times of up to 24 hours are used, more preferably 1-12 hours, with 1-6 hours most preferred. Short infusion times which allow treatment to be carried out without an overnight stay in hospital are especially desirable. However, infusion may be 12 to 24 hours or even longer if required. Infusion may be carried out at suitable intervals of say 1 to 4 weeks. Pharmaceutical compositions containing compounds of the invention may be delivered by liposome or nanosphere encapsulation, in sustained release formulations or by other standard delivery means.

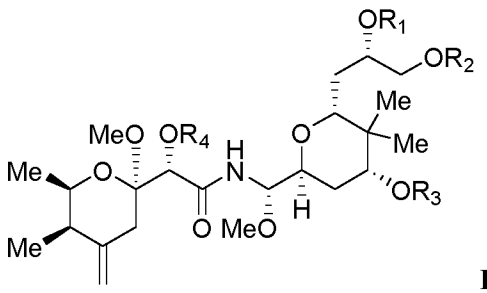
The correct dosage of the compounds will vary according to the particular formulation, the mode of application, ant the particular status, host and tumor being treated. Other factors like age, body weight, sex, diet, time of administration, rate of excretion, condition of the host, drug combinations, reaction sensitivities and severity of the disease shall be taken into account. Administration can be carried out continuously or periodically within the maximum tolerated dose.

As used herein, the terms "treat", "treating" and "treatment" include the eradication, removal, modification, or control of a tumor or primary, regional, or metastatic cancer cells or tissue and the minimization of delay of the spread of cancer.

The compounds of the invention have anticancer activity against several cancer types which include, but are not limited to, lung cancer, colon cancer, breast cancer and pancreas cancer.

Thus in alternative embodiments of the present invention, the pharmaceutical composition comprising a compound of formula **I** and **III** as defined above is for the treatment of lung cancer, colon cancer, breast cancer or pancreas cancer.

In a sixth aspect, the present invention is directed to a process for the production of compounds of formula **II**. Preferred processes according to this aspect of the invention are those that produce a compound also having formula **IV**



IV

or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof;

wherein R_1, R_2, R_3 and R_4 are as defined above in general formula **II**.

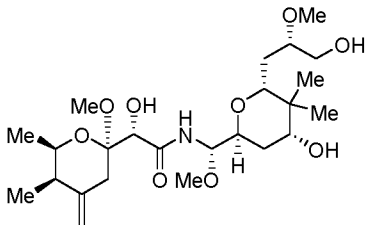
In processes for the synthesis of compounds of formula **II** and **IV**, particularly preferred R_1 is selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, and –C(=O)R_a where R_a is substituted or unsubstituted C_1-C_{12} alkyl. More preferably R_1 is selected from hydrogen, substituted or unsubstituted C_1-C_6 alkyl and –C(=O)R_a where R_a is substituted or unsubstituted C_1-C_6 alkyl. Even more preferably, R_1 is selected from hydrogen, methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *tert*-butyl, *sec*-butyl, isobutyl and –C(=O)R_a wherein R_a is selected from methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *tert*-butyl, *sec*-butyl and isobutyl. Most preferred R_1 is selected from hydrogen and methyl.

In processes for the synthesis of compounds of formula **II** and **IV**, particularly preferred R_2 is selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, and –C(=O)R_a where R_a is substituted or unsubstituted C_1-C_{12} alkyl. More preferably R_2 is selected from hydrogen, substituted or unsubstituted C_1-C_6 alkyl and –(C=O)R_a where R_a is substituted or unsubstituted C_1-C_6 alkyl.

Even more preferably, R_2 is selected from hydrogen, methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *tert*-butyl, *sec*-butyl, isobutyl, and $-C(=O)R_a$ where R_a is selected from methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *tert*-butyl, *sec*-butyl and isobutyl. Most preferred R_2 are hydrogen, methyl and acetyl.

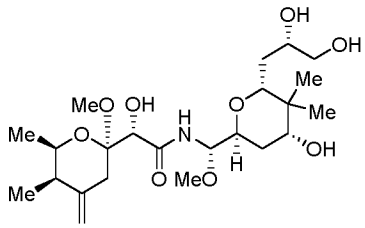
In processes for the synthesis of compounds of formula **II** and **IV**, particularly preferred R_3 and R_4 are independently selected from hydrogen and $-C(=O)R_a$, wherein R_a at each occurrence is independently selected from substituted or unsubstituted C_1-C_{12} alkyl. More preferably R_a at each occurrence is independently selected form substituted or unsubstituted C_1-C_6 alkyl. Even more preferably, R_a is selected from methyl, ethyl, *n*-propyl, isopropyl, *n*-butyl, *tert*-butyl, *sec*-butyl and isobutyl. Most preferred R_3 and R_4 are independently selected from hydrogen and acetyl.

In processes for the synthesis of compounds of formula **II** and **IV**, particularly preferred compounds **1** and **2** have, respectively, the following relative stereochemistry:



1'

,



2'

.

In additional preferred embodiments, the preferences described above for the different substituents are combined. The present invention is also directed to such combinations of preferred substitutions in the processes for the synthesis of compounds of formula **II** and **IV** above.

In a more preferred embodiment of this aspect of the invention the compound of formula **II** or **IV** is pederin.

In an even more preferred embodiment, pederin is obtained from compound **1'** by:

- Protecting all the hydroxy groups in compound **1'** with a protecting group for –OH suitable to be selectively removed from a protected primary OH in presence of protected secondary OH. Examples of such protecting group include trimethylsilyl, triethylsilyl, triisopropylsilyl, and *tert*-butyldimethylsilyl. Most preferred protecting group for this step is *tert*-butyldimethylsilyl;
- Selectively removing the primary OH protecting group;
- Methylating the resulting primary hydroxy group with a suitable methylation reagent; and

- Removing the other protecting groups for OH.

In another more preferred embodiment, pederin is obtained from compound **2'** by:

- Protecting the 1,2-diol group with a suitable protecting group for 1,2-diols. Examples of suitable protecting groups for 1,2-diols include, but are not limited to, those groups that after reaction with the corresponding 1,2-diol generate Mocdene acetal, Bocdene acetal, acrolein acetal, benzylidene acetal, (*t*-butyldimethylsilyloxy)benzylidene acetal, mesitylene acetal, methoxymethylene acetal, ethoxymethylene acetal, cyclic carbonates, methyl boronate and ethyl boronate. More preferred protecting groups for this step are those that generate a Mocdene acetal, Bocdene acetal, benzylidene acetal, and cyclic carbonates being the protecting group that generates a benzylidene acetal the most preferred;
- Protecting the other hydroxy groups with a protecting group for –OH that is orthogonal with the 1,2-diol protecting group of previous step. Examples of protecting groups for OH suitable for this step are trimethylsilyl, triethylsilyl, triisopropylsilyl *tert*-butyldimethylsilyl, and acetyl. Most preferred protecting group for this step are *tert*-butyldimethylsilyl and acetyl;
- Removing the 1,2-diol protecting group;
- Methylating the resulting 1,2-diol with a suitable methylation reagent; and
- Removing the other protecting groups for OH.

Examples of suitable methylation reagents include methyl iodide, methyl bromide, dimethylsulfate, and methyl triflate.

The isolated nucleic acid according to the eighth aspect of the invention is preferably derived from *Labrenzia sp*, and in particular from strain PHM005.

The complete genome sequence of this bacterium revealed the biosynthetic gene cluster responsible for the pederin and onnamide synthesis. Bioinformatic analysis was used to predict the function of the genes in the cluster.

This gene cluster, named **Lab** gene cluster, is a Trans-AT hybrid polyketide synthase / non ribosomal synthetase (PKS/NRPS) gene cluster with 69 Kb. It was deduced from genome mining from the whole sequenciation of the genome of strain PHM005 composed by 20 ORF homologous to the described for pederin gene cluster. It contains genes encoding enzymes for the biosynthesis of pederin-like and onnamide-like compounds.

In a preferred embodiment, the isolated nucleic acid preferably comprises nucleic acid fragments forming individual units and/or modules of the Lab biosynthetic gene cluster as it is

shown in more detail in Figure 3. As depicted in Figure 3, the Lab gene cluster contains the units *lab706* to *lab726*.

In a particularly preferred embodiment, the isolated nucleic acid according to the eighth aspect of the present invention comprises:

a nucleotide sequence as shown in SEQ ID NO: 2; or

a nucleotide sequence which is the complement of SEQ ID NO: 2; or

a nucleotide sequence hybridising under highly stringent conditions to SEQ ID NO: 2; or to the complement thereof; or

a nucleotide sequence having at least 80% sequence identity with SEQ ID NO: 2 or with the complement thereof.

Particularly preferred nucleic acid fragments according to the nineth aspect of the present invention are nucleic acid fragments essentially comprising at least one of the genes *lab708*, *lab709*, *lab710*, *lab721*, *lab722*, *lab723*, *lab724* and *lab725*. Further preferred are the nucleic acid fragments comprising one or more nucleotide sequences encoding the protein sequences as shown in SEQ ID NOs: 3 to 23. Also preferred parts are the corresponding parts of the nucleotide sequence SEQ ID NO: 2.

In another preferred embodiment particularly preferred fragments are those essentially consisting of *lab719* and/or *lab720*. Further preferred is the nucleic acid fragment comprising the nucleotide sequence encoding the protein sequence as shown in SEQ ID NO: 16 and/or SEQ ID NO: 17. Also preferred are the corresponding parts of the nucleotide sequence SEQ ID NO: 2.

The annotation of the whole genome of PHM005 reveals a circular chromosome with a length of 6167 bp, 5651 coding sequences (CDS), 53 tRNA and 10 rRNA. 55% G+C.

Exploring the entire genome into a unique contig using software for prediction/identification of secondary metabolisms as antiSMASH V 3.0 (Weber and co-workers, Nucleic Acid Research, 2015 *doi: 10.1093/nar/gkv437)* a 102 Kb of a large hybrid PKS/NRPS gene cluster is detected. Among the 317 ORF analyzed, 20 genes (69 Kb) shown homologies to pederin (ped) and onnamide (onn) sequences based on BLASTp against symbiont bacterium of *Paedeus fascipens* (GenBank AH013687.2) and bacterium symbiont of *Theonella swinhoei* (GenBank AY688304.1) as shown in more detail in **Table 1**.

Table 1.Homologies of the ***lab*** genes respect to ***ped*** (pederin) and ***onn*** (onnamide) genes.

lab gene	Prot. Size (aas)	Putative function in *Labrenzia* sp. PHM005	Bacterium Symbiont *Paedeus fuscipens (AH013687.2)*		Bacterium Symbiont *Theonella swinhoei (AY688304.1)*	
			Gene	% H/Q*	Gene	% H/Q
706	80	Polyketide Biosynthesis Acyl carrier protein (ACP)	***ped*N**	47/87	-	No homology
707	425	Polyketide Biosynthesis 3-hydroxy-3 methylglutaryl ACP synthase (HMGS)	***ped*P**	61/99	***onn*A**	60/99
708	1165	Polyketide synthase (GNAT-ACP-KS-DHt)	***ped*I**	42/93	***onn*B**	39/98
709	3219	TransAT PKS (KR-cMT-ACP-KS-TransAT-ECH-ACPb-ACPb-KS—KR)	***ped*I**	49/94	***onn*B** ***onn*I**	41/73 45/73
710	97	Phosphopantetheine attached site (ACP)	***ped*I**	46/90	***onn*I**	34/73
711	373	Monooxygenase (OX)	***ped*J**	60/98	***onn*C**	58/98
712	318	Methyltransferase (oMT)	***ped*A**	47/97	***onn*G** ***onn*D**	51/99 46/97
713	414	Cytochrome P450		No homology		No homology
714	447	Malonyl CoA-ACP transacylase (or oxidorectase)	***ped*B**	56/98		No homology
715	337	Malonyl CoA-ACP transcylase	***ped*C**	38/94		No homology
716	375	Malonyl CoA-ACP transcylase	***ped*D**	51/95		No homology
717	253	Enoyl transferase	***ped*L**	43/91		No homology
718	411	Beta-ketocacyl-synthase	***ped*M**	30/81		No homology
719	2254	Mixed TransAT PKS/NRPS (ACP-KS-TransAT-DH-KR-ACP-KS-DH-DH-ACP-KS-TransAT-KR-ACP-KS-TransAT-C-A-PCP-TE)	***ped*H**	42/99	***onn*I**	35/84
720	437	Oxidoreductase (Ox)	***ped*G**	73/94		No homology
721	1986	TransAT-PKS (PS-KR-ACP-KS-TransAT-KR-KS-TransAT)	***ped*F**	40/99	***onn*I**	30/82
722	1949	TransAT Polyketide synthase (TransAT-KR-cMT-ACPb-KS-TransAT-DH)	***ped*F**	44/97	***onn*I** ***onn*B**	36/86 34/85
723	875	Polyketide synthase (KR-ACP-KS)	***ped*F**	49/93	***onn*B** ***onn*I**	52/96 45/95
724	1986	Mixed PKS/NRPS (DHt-ACP-C-A(gly)-PCP-KS-TransAT)	***ped*F**	42/99	***onn*I**	38/96
725	377	Polyketide synthase (KS)	***ped*F**	48/99	***onn*I** ***onn*B**	46/92 41/88
726	278	Methyltransferase (MT)	***ped*E**	51/98	***onn*H**	43/99

(*) H: Homology in %. Q: Query covered in %

The putative Lab gene cluster comprises a 69 Kb nucleic acid fragments forming individual units and/or modules similar to those described for pederin biosynthetic gene cluster as it is shown in more detail in Figure 3.

The TransAT hybrid PKS/NRPS Lab gene cluster is mainly composed by one PKS (Composed by ORF *lab708*, *lab709* and *lab710*) and two mixed PKS/NRPS systems (*lab721*, *lab722*, *lab723*, *lab724*, *lab725* and *lab719*) flanked by oxygenases, oxidoreductases and methylases in closed similar architecture to the described by J. Piel for *ped* genes. The predicted functions and the composition of the aminoacids of each ORF is detailed in Table 1.

The TransAT-PKS *lab708*, *lab709*, *lab710* (4.481 amino acids) is composed by the modules GNAT-ACP-KS-DHt-KR-cMT-ACP-KS-TransAT-ECH-ACP-ACP-KS-KR-ACP) similar to the described for *ped*I with homologies 42-49%. This biosynthetic gene cluster may be the responsible of the biosynthesis of the six membered ring bearing the exomethylene group of the pederin structure. Where the domains are GNAT: Gcn5-related-N-acetyltransferase; ACP: Acyl Carrier Protein; KS: Ketosynthase; DHt Dehydratase; KR: Ketoreductase; cMT: Methyltransferase; ECH Enoyl-CoA-hydratase o crotonase; TransAT: Trans Acyl Transferase).

The hybrid Trans-AT PKS/NRPS formed by *lab721*, *lab722*, *lab723*, *lab724*, *lab725* (5.385 aa) is composed by 6 Kethosyntases and 1 NRPS with a clear adenylation for glycine. (PS-KR-ACP-KS-TransAT-KR-KS-TransAT-transAT-KR-cMT-ACP-KS-TransAT-DH-KR-ACP-KS-DHt-ACP-C-A (gly)-PCP-KS-TransAT-KS). With 40-49% homology to *ped*F but essentially the same functions and architecture of modules. Where the domains are C: nonribosomal peptide Condensation; A: nonribosomal peptide Adenylation; PCP: Thiolation and Peptide Carrier Protein.

According to a preferred embodiment of the nineth aspect, we have identified the *lab*719 PKS/NRPS system related to the biosynthesis of any onnamide-like compound from the Lab gene cluster. This putative new compound has not been identified in the fermentation broth of PHM005. It is possible that the product of the gen *lab*720, an oxidoreductase, will prevent the formation of the onnamide-like compound by cleaving the pederin structure before to add to the first domain ACP in *lab*719 or a final oxidative breakout is produced after its biosynthesis. The same doubt has been discussed by J. Piel in the WO 03/044186 A2. The genetic modification of the gene *lab*719 (homology to *ped*G) will solve this uncertainty.

This “silent” hybrid transAT PKS/NRPS gene, represented by *lab719* (2.254 aa) is composed by 4 KS and 1 NRPS with uncertain adenylation domain, maybe for the incorporation of

arg (as the case of onnamide), but asp, asn, glu and gln could be other possible alternatives as propodes by NRPSPredictor2 SVM algorithm. The composition of this ORF is (ACP-KS-TransAT-DH-KR-ACP-KS-DH-DH-ACP-KS-TransAT-KR-ACP-KS-TransAT-C-A -PCP-TE). Where TE: Thioesterase domain.

The single ORF in the lab region without sequence-homology to ped, onn or nsp (nosperin) islands is the *lab713*, putative for a cytochrome P450, maybe playing a role in oxygenation of polyketides, as described by J. Piel in the case of the ped islands. (J. Bacteriol. 2004. 186(5), 1280-1286) with similar function-assigned genes.

Particularly preferred modular enzymatic system according to the tenth aspect of the present invention comprises a protein sequence according to any of the sequences SEQ ID NO: 3 to SEQ ID NO: 23 or a protein sequence having at least 80% sequence identity with these sequences.

Particularly preferred host cells according to the twelfth aspect of the present invention are bacterial cells. More particularly preferred host cells are *Pseudomonas*, *Acinetobacter*, *Bacillus*, *Streptomyces* and *E. coli*.

The inventive modification on Lab biosynthetic gene cluster can be used in the preparation of a modified Lab biosynthetic gene cluster or in the preparation of pederin-like or onnamide-like compounds.

In a preferred embodiment according to the thirteenth aspect of the present invention the product of the *lab719* is expressed.

EXAMPLES

General Structure Elucidation Procedure. Optical rotations were determined using a Jasco P-1020 polarimeter. NMR spectra were obtained on a Varian "Unity 500" spectrometer at 500/125 MHz ($^{1}H/^{13}C$) and on a Varian "Unity 400" spectrometer at 400/100 MHz ($^{1}H/^{13}C$). Chemical shifts were reported in ppm using residual solvent peak for $CDCl_3$ (δ 7.26 ppm for ^{1}H and 77.0 ppm for ^{13}C) as an internal reference. (+)ESIMS were recorded using an Agilent 1100 Series LC/MSD spectrometer. High Resolution Mass Spectroscopy (HRMS) was performed using an Agilent 6230 TOF LC/MS system and the ESI-MS technique.

Example 1: Bacteria isolation

The pederin-type producing bacteria, *Labrenzia* sp., PHM005 was isolated from a sediment sample collected at a depth of 18 m from a highly epiphytic and unidentified coral-sponge habitat off the coast of Kenya in 2005. Approximately 5 grams of sea gravel material was collected in a 50 ml Falcon tube containing sterile artificial sea water (ASW) and was maintained at 5 ºC for 5 days before being processed. Once in the laboratory, the sample was homogenized and 100 µl of a 1:100 dilution with ASW spread directly on Petri dishes with a sea salt medium consisting of 27 g/L marine salts (Tropic Marin® PRO-REEF), 16 g/L agar and 0.2 mg/mL of cycloheximide. After incubation for three weeks at 28ºC, a slightly brown colony was picked and transferred to the same sea salt medium to confirm the purity and generate biomass for molecular characterization with one colony being inoculated on liquid marine broth for further conservation on 20% glycerol at -80ºC as a cell bank.

Example 2: Electron Microscopy.

Cells in the mid-exponential growth phase were adsorbed on 400-mesh carbon-collodion-coated grids for 2 min, negatively stained with 2% uranyl acetate, imaged with a Jeol JEM 1011 transmission electron microscope operated at 100 kV and photographed with a CCD Gatan Erlangshen ES1000W camera.

Example 3: 16S rRNA characterization.

For DNA extraction the strain was grown in marine broth (DIFCO 1196) for 72 hours. Cells were recovered and lysed by boiling with 4% NP40 for 10 minutes. Cell debris was discarded by centrifugation. The 16S rDNA gene was amplified by the polymerase chain reaction using the bacterial primers F1 and R5. The phylogenetic tree (Figure 2) was generated by the Pairwise alignment based similarity coefficient and UPGMA for Cluster analysis using BioNumerics V7.5 (Applied Maths). The phylogenetic neighbors were identified and pairwise 16S rDNA gene sequence similarities calculated by comparison with the SILVA LTPs123 database.

Example 4: Cultivation and extraction.

The strain clearly needs marine salt to grow. After culture, the whole broths were lyophilized and extracted with a mixture of organic solvents and a 0.5 mL sample of the crude extract dried and screened for cytotoxic activity. The best cytotoxic activity was achieved in the 16B/d medium at 120h. This medium consisted of 17.5 g/L of brewer's yeast (Sensient, G2025), 76 g/L mannitol, 7 g/L $(NH_4)_2SO_4$, 13 g/L $CaCO_3$, 0.09 g/L $FeCl_3$ and 36 g/L marine salts (Tropic Marin® PRO-REEF). A 50 L scale-up of this bacterium in 16B/d medium was prepared in 200 x 2L Erlenmeyer flasks each with a working volume of 250 mL. The production flasks were inoculated with 2% of

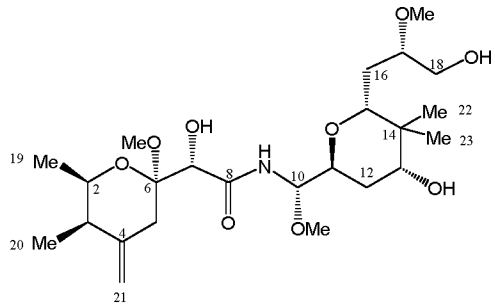
the bacteria grown during 72 h in marine broth (DIFCO 1196) from another highly grown pre-inoculum. The scale-up was incubated during 120h at 28ºC in a rotatory shaker at 220 rpm with 5cm eccentricity. The culture was then centrifuged at 6.000 rpm during 20 minutes to give 45 L of aqueous supernatant which was extracted twice with EtOAc and the organic phase dried to give a crude extract (1.8 g).

Example 5: Isolation of Compound **1**.

The extract was applied to a silica gel VFC (vacuum flash chromatography) system, using a stepwise gradient elution with *n*-hexane-EtOAc and EtOAc-MeOH mixtures to give eleven fractions. The active fractions were eluted with EtOAc and EtOAc-MeOH 9:1 (550.0 mg) and were subjected to preparative reversed-phase HPLC using a symmetry C_{18} column (19x150mm, 7µm) and a linear gradient of H_2O/CH_3CN from 5% to 35% CH_3CN over 30 min at a flow rate of 13.5 mL/min, to afford a very active peak-fraction (77.0 mg) with a retention time of 24.5 min containing **1** based on the HPLC-MS chromatogram. This fraction was further purified by semipreparative HPLC on a XBridge C_{18} column (10x250mm, 5µm) and isocratic elution with H_2O/CH_3CN (78:22) at a flow of 4 mL/min, to yield 24.5 mg of pure compound **1** with a retention time of 25.0 min at these HPLC conditions.

(1): colorless oil; $[\alpha]_D^{20}$ + 82.4 (c=0.49; $CHCl_3$) and $[\alpha]_D^{20}$ + 81.3 (c=0.36; MeOH); ^{1}H NMR ($CDCl_3$) δ 3.99 (1H, dq, *J*=6.6, 2.7 Hz, H-2), 2.25 (1H, dq, *J*=7.1, 2.7 Hz, H-3), 2.43 (1H, d, *J*=14.1 Hz, H-5a), 2.36 (1H, dt, *J* = 14.1, 2.3 Hz, H-5b), 4.31 (1H, s, H-7), 7.18 (1H, d, *J*=9.8 Hz, NH), 5.37 (1H, dd, *J*=9.8, 7.8 Hz, H-10), 3.83 (1H, dt, *J*=7.8, 2.7 Hz, H-11), 2.04 (1H, dt, *J*=13.5, 3.6 Hz, H-12a), 1.75 (1H, m, H-12b), 3.64 (1H, m, H-13), 3.31 (1H, m, H-15), 1.75 (1H, m, H-16a), 1.57 (1H, dd, *J*=14.3, 9.7 Hz, H-16b), 3.36 (1H, m, H-17), 3.65 (1H, m, H-18a), 3.48 (1H, m, H-18b), 1.19 (3H, d, *J*=6.6 Hz, H-19), 1.01 (3H, d, *J*=7.1 Hz, H-20), 4.85 (1H, t, *J* = 2.3 Hz, H-21a), 4.73 (1H, t, *J* = 2.3 Hz, H-21b), 0.95 (3H, s, C-22), 0.88 (3H, s, C-23), 3.32 (3H, s, MeO-6), 3.38 (3H, s, MeO-10), 3.32 (3H, s, MeO-17); ^{13}C NMR ($CDCl_3$) δ 69.6 (d, C-2), 41.3 (d, C-3), 145.7 (s, C-4), 34.1 (t, C-5), 99.7 (s, C-6), 73.1 (d, C-7), 171.9 (s, C-8), 79.4 (d, C-10), 72.6 (d, C-11), 29.6 (t, C-12), 71.8 (d, C-13), 38.4 (s, C-14), 75.4 (d, C-15), 29.2 (t, C-16), 79.0 (d, C-17), 63.8 (t, C-18), 17.9 (q, C-19), 12.0 (q, C-20), 110.5 (t, C-21), 23.1 (s, C-22), 13.5 (s, C-23), 49.1 (q, MeO-6), 56.4 (q, MeO-10), 56.6 (q, MeO-17); (+)-ESIMS *m/z* 512.3 $[M + Na]^+$; (+)-HRES-TOFMS *m/z* 512.2873 $[M + Na]^+$ (calcd. for $C_{24}H_{43}NO_9Na$, 512.2830).

The relative stereochemistry of compound **1** was established as



on the basis of ROESY data and analysis of coupling constants. The optical rotation of compound **1** ($[\alpha]_D^{20}$ + 82.4, c = 0.49; $CHCl_3$ and $[\alpha]_D^{20}$ +81.3, c = 0.36; MeOH) showed the same sign as pederin ($[\alpha]_D^{20}$ + 86.8, c = 1.00; $CHCl_3$). The absolute stereochemistry of pederin has been established by X-ray crystallographic study (Simpson, J. S. *et. al.* J. Nat. Prod. 2000, 63, 704-706) and stereoselective synthesis (Matsuda, F., *et. al.* Tetrahedron 1988, 44, 7063-7080). Therefore, we tentatively propose the absolute configuration of compound **1** to be the same as pederin and other reported analogous compounds (Wan, S. *et. al.* J. Am. Chem. Soc. 2011, 133, 16668-16679).

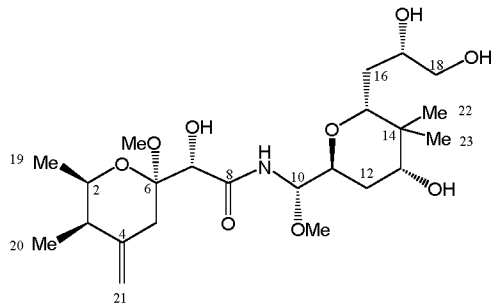
Example 6. Isolation of compound **2**.

Compound **2** was isolated from the whole broth crude extract (9.5 g) of the fermentation broth (15 L) of the marine derived strain PHM005. The extract was applied to a silica gel VFC (vacuum flash chromatography) system, using a stepwise gradient elution with *n*-hexane-EtOAc and EtOAc-MeOH mixtures to give seven fractions. The active fraction containing compound **2** was eluted with EtOAc-MeOH 4:1 (659.0 mg) and were subjected to semipreparative reversed-phase HPLC equipped with a Symmetry C_{18} column (7.8 x 150 mm, 5µm) using a linear gradient of H_2O/CH_3CN from 5% to 60% of CH_3CN in 25 min at a flow rate of 3.0 mL/min, to afford a very active time-fraction between 25-30 min (28.0 mg) containing compound **2** based on HPLC-MS chromatogram. This fraction was again purified by semipreparative HPLC on a Symmetry C18 column (7.8 x 150 mm, 5µm), using a linear gradient of H_2O/CH_3CN from 20% to 30% of CH_3CN in 20 min at a flow rate of 2.5 mL/min, to yield 2.6 mg of pure compound **2** with a retention time of 11.5 min at these HPLC conditions.

2: colorless oil; $[\alpha]_D^{20}$ + 64.5 (c=0.16; $CHCl_3$); ^1H NMR ($CDCl_3$) δ 3.97 (1H, dq, *J*=6.6, 2.6 Hz, H-2), 2.25 (1H, dq, *J*=7.1, 2.6 Hz, H-3),), 2.50 (1H, dt, *J*=14.2, 1.45 Hz, H-5a), 2.45 (1H, d, *J*=14.1 Hz, H-5b), 4.32 (1H, s, H-7), 7.17 (1H, d, *J*=9.9 Hz, NH), 5.44 (1H, dd, *J*=9.9, 7.5 Hz, H-10), 3.95 (1H, m, H-11), 2.05 (1H, dt, *J*=13.5, 4.0 Hz, H-12a), 1.75 (1H, m, H-12b), 3.66 (1H, m, H-13), 3.58 (1H, m, H-15), 1.80 (1H, m, H-16a), 1.55 (1H, m, H-16b), 3.80 (1H, m, H-17), 3.57 (1H, m,

H-18), 3.44 (1H, dd, J=11.5, 6.5 Hz, H-18), 1.19 (3H, d, *J*=6.6 Hz, H-19), 1.01 (3H, d, *J*=7.1 Hz, H-20), 4.85 (1H, t, J=1.45 Hz, H-21a), 4.75 (1H, t, J=1.45 Hz, H-21b), 0.96 (3H, s, C22), 0.89 (3H, s, C-23), 3.34 (3H, s, MeO-6), 3.41 (3H, s, MeO-10); ^{13}C NMR ($CDCl_3$) δ 69.6 (d, C-2), 41.3 (d, C-3), 146.1 (s, C-4), 34.2 (t, C-5), 99.6 (s, C-6), 74.5 (d, C-7), 171.9 (s, C-8), 79.3 (d, C-10), 72.2 (d, C-11), 29.8 (t, C-12), 71.6 (d, C-13), 38.4 (s, C-14), 80.9 (d, C-15), 31.4 (t, C-16), 72.8 (d, C-17), 66.6 (t, C-18), 17.8 (q, C-19), 11.9 (q, C-20), 110.2 (t, C-21), 23.4 (s, C-22), 14.3 (s, C-23), 49.6 (q, MeO-6), 56.3 (q, MeO-10); (+)-ESIMS *m/z* 498.4 $[M + Na]^+$; (+)-HRES-TOFMS *m/z* 498.2713 $[M + Na]^+$ (calcd. for $C_{23}H_{41}NO_9Na$, 498.2674).

The relative stereochemistry of compound **2** was assigned as



on the basis of an analysis of coupling constants. The optical rotation of compound **2** ($[\alpha]_D^{20}$ + 64.5, c=0.16; $CHCl_3$) showed the same sign as pederin ($[\alpha]_D^{20}$ + 86.8, c = 1.00; $CHCl_3$). Therefore, we tentatively propose the absolute configuration of compound **2** to be the same as pederin and other reported analogous compounds (Wan, S. *et. al.* J. Am. Chem. Soc. 2011, 133, 16668-16679).

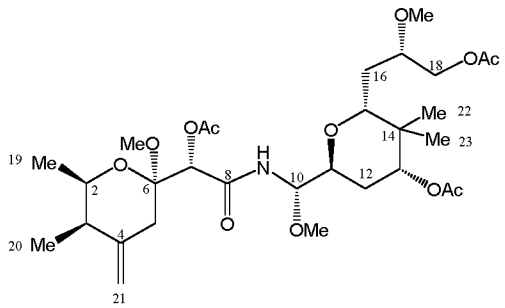
Example 7. Synthesis of compound **3**

To a solution of **1** (2.5 mg, 5.1 µmol) in dry DCM (2 mL) under a nitrogen atmosphere, were added pyridine (10 µL, 124 µmol), DMAP (catalytic amount) and Ac_2O (2.9 µL, 31 mmol). The reaction was allowed to stand at room temperature overnight. The mixture was concentrated under vacuum and purified via flash column chromatography on silica gel (*n*-hexane/EtOAc 1:1) to afford **3** (3 mg, 95%) as a white solid.

3: ^{1}H NMR ($CDCl_3$) δ 3.96 (1H, dq, *J*=6.6, 2.6 Hz, H-2), 2.24 (1H, dq, *J*=7.0, 2.6 Hz, H-3), 2.62 (1H, dt, *J*=14.5, 2.2 Hz, H-5a), 2.37 (1H, d, *J*=14.5 Hz, H-5b), 5.25 (1H, s, H-7), 6.62 (1H, d, *J*=9.6 Hz, NH), 5.27 (1H, dd, *J*=9.6, 4.1Hz, H-10), 3.91(1H, dt, *J*=6.3, 4.6, Hz, H-11), 2.02 (1H, m, H-12a), 1.66 (1H, m, H-12b), 4.91 (1H, dd, J=4.7, 4.1Hz, H-13), 3.55 (1H, m, H-15), 2.02 (1H,

m, H-16a), 1.67 (1H, m, H-16b), 3.60 (1H, dd, J=11.3, 2.2 Hz, H-17), 4.32 (1H, dd, J=12.1, 2.6 Hz, H-18a), 4.12 (1H, m, H-18b), 1.15 (3H, d, *J*=6.6 Hz, H-19), 0.97 (3H, d, *J*=7.0 Hz, H-20), 4.86 (1H, t, J=2.0 Hz, H-2a), 4.76 (1H, t, J=2.0 Hz, H-21b), 0.97 (3H, s, C22), 0.89 (3H, s, C-23), 3.21 (3H, s, MeO-6), 3.39 (3H, s, MeO-10), 3.38 (3H, s, MeO-17), 2.20 (3H, s, OCOMe-7), 2.08 (3H, s, OCOMe-13), 2.10 (3H, s, OCOMe-18) ; 13C NMR ($CDCl_3$) δ 69.6 (d, C-2), 41.3 (d, C-3), 145.5 (s, C-4), 33.8 (t, C-5), 99.1 (s, C-6), 72.1 (d, C-7), 167.4 (s, C-8), 81.8 (d, C-10), 70.0 (d, C-11), 26.7 (t, C-12), 74.2 (d, C-13), 36.7 (s, C-14), 76.5 (d, C-15), 29.3 (t, C-16), 76.4 (d, C-17), 64.0 (t, C-18), 17.9 (q, C-19), 12.0 (q, C-20), 110.4 (t, C-21), 24.7 (s, C-22), 17.2 (s, C-23), 48.4 (q, MeO-6), 56.3 (q, MeO-10), 57.0 (q, MeO-17), 20.7 (q, OCOMe-7), 169.8 (s, OCOMe-7), 21.2 (q, OCOMe-13), 170.3 (s, OCOMe-13), 20.9 (q, OCOMe-18), 170.0 (s, OCOMe-18), ; (+)-ESIMS *m/z* 638.3 $[M + Na]^+$.

The relative stereochemistry of compound **3** was established as



by analogy with its precursor, compound **1**.

Example 8. In vitro bioassays for the detection of antitumor activity

The aim of this assay is to evaluate the in vitro cytostatic (ability to delay or arrest tumor cell growth) or cytotoxic (ability to kill tumor cells) activity of the samples being tested.

CELL LINES

Name	Nº ATCC	Species	Tissue	Characteristics
A549	CCL-185	human	lung	lung carcinoma (NSCLC)
HT29	HTB-38	human	colon	colorectal adenocarcinoma
MDA-MB-231	HTB-26	human	breast	breast adenocarcinoma
PSN1	CRM-CRL-3211	human	pancreas	pancreas adenocarcinoma

EVALUATION OF CYTOTOXIC ACTIVITY USING THE SBR COLORIMETRIC ASSAY

A colorimetric assay, using sulforhodamine B (SRB) reaction has been adapted to provide a quantitative measurement of cell growth and viability (following the technique described by Skehan *et al.* J. Natl. Cancer Inst. 1990, 82, 1107-1112).

This form of assay employs 96-well cell culture microplates following the standards of the American National Standards Institute and the Society for Laboratory Automation and Screening (ANSI SLAS 1-2004 (R2012) 10/12/2011). All the cell lines used in this study were obtained from the American Type Culture Collection (ATCC) and derive from different types of human cancer.

Cells were maintained in Dulbecco's Modified Eagle Medium (DMEM) supplemented with 10% Fetal Bovine Serum (FBS), 2mM L-glutamine, 100 U/mL penicillin and 100 U/mL streptomycin at 37 °C, 5% CO_2 and 98% humidity. For the experiments, cells were harvested from subconfluent cultures using trypsinization and resuspended in fresh medium before counting and plating.

Cells were seeded in 96 well microtiter plates, at 5000 cells per well in aliquots of 150 µL, and allowed to attach to the plate surface for 18 hours (overnight) in drug free medium. After that, one control (untreated) plate of each cell line was fixed (as described below) and used for time zero reference value. Culture plates were then treated with test compounds (50 µL aliquots of 4X stock solutions in complete culture medium plus 4% DMSO) using ten 2/5 serial dilutions (concentrations ranging from 10 to 0.003 µg/mL) and triplicate cultures (1% final concentration in DMSO). After 72 hours treatment, the antitumor effect was measured by using the SRB methodology: Briefly, cells were washed twice with PBS, fixed for 15 min in 1% glutaraldehyde solution at room temperature, rinsed twice in PBS, and stained in 0.4% SRB solution for 30 min at room temperature. Cells were then rinsed several times with 1% acetic acid solution and air-dried at room temperature. SRB was then extracted in 10 mM trizma base solution and the absorbance measured in an automated spectrophotometric plate reader at 490 nm. Effects on cell growth and survival were estimated by applying the NCI algorithm (Boyd MR and Paull KD. Drug Dev. Res. 1995, 34, 91-104).

The values obtained in triplicate cultures were fitted to a four-parameter logistic curve by nonlinear regression analysis. Three reference parameters were calculated (according to the NCI algorithm) by automatic interpolation of the curves obtained from such fitting: GI_{50} = compound concentration that produces 50% cell growth inhibition, as compared to control cultures; TGI =

total cell growth inhibition (cytostatic effect), as compared to control cultures, and LC_{50} = compound concentration that produces 50% net cell killing cytotoxic effect).

Table 2 illustrates data on the biological activity of compounds of the present invention.

	Biological activity (M)				
Compound		Cell Line			
		A549	HT29	MDA-MB-231	PSN-1
1	GI_{50}	2.04E-09	2.86E-09	2.66E-09	2.66E-09
	TGI	7.97E-09	8.99E-09	5.31E-09	5.72E-09
	LC_{50}	3.68E-08	>2.04E-07	1.08E-08	1.94E-08
2	GI_{50}	7.15E-09	8.83E-09	8.20E-09	8.62E-09
	TGI	2.52E-08	4.42E-08	1.56E-08	1.91E-08
	LC_{50}	1.22E-07	>2.10E-06	3.15E-08	7.78E-08
3	GI_{50}	1.15E-07	1.62E-07	3.09E-07	1.62E-07
	TGI	8.77E-07	9.26E-07	2.44E-06	6.66E-07
	LC_{50}	8.61E-06	>1.62E-05	>1.62E-05	3.90E-06

PCT

Print Out (Original in Electronic Form)
(This sheet is not part of and does not count as a sheet of the international application)

0-1	**Form PCT/RO/134 Indications Relating to Deposited Microorganism(s) or Other Biological Material (PCT Rule 13bis)**	
0-1-1	Prepared Using	PCT Online Filing Version 3.5.000.251e MT/FOP 20141031/0.20.5.20
0-2	**International Application No.**	
0-3	**Applicant's or agent's file reference**	P14452PC00

1	**The indications made below relate to the deposited microorganism(s) or other biological material referred to in the description on:**	
1-1	**page**	18-19
1-2	**line**	17 and 18
1-3	**Identification of deposit**	
1-3-1	Name of depositary institution	CECT Colección Española de Cultivos Tipo (CECT)
1-3-2	Address of depositary institution	Edificio 3 CUE, Parc Cientific Universitat de Valencia, Catedrático Agustín Escardino, 9, 46980 Paterna (Valencia), Spain
1-3-3	Date of deposit	22 November 2016 (22.11.2016)
1-3-4	Accession Number	CECT 9225
1-5	**Designated States for Which Indications are Made**	All designations

FOR RECEIVING OFFICE USE ONLY

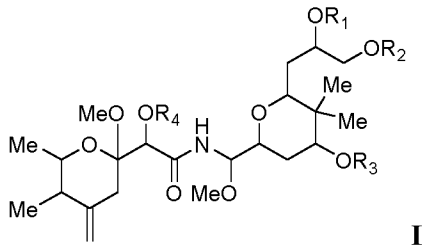
0-4	**This form was received with the international application:** (yes or no)	Yes
0-4-1	Authorized officer	Olsson, Johanna

FOR INTERNATIONAL BUREAU USE ONLY

0-5	**This form was received by the international Bureau on:**	
0-5-1	Authorized officer	

CLAIMS

1. A compound of general formula **I** or a pharmaceutically acceptable salt, tautomer, or stereoisomer thereof



I

wherein:

R_1, R_2, and R_3 are each independently selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, $-C(=O)R_a$, $-C(=O)OR_b$ and $-(C=O)NR_cR_d$;

R_4 is selected from hydrogen, $-C(=O)R_a$, $-C(=O)OR_b$, and $-C(=O)NR_cR_d$;

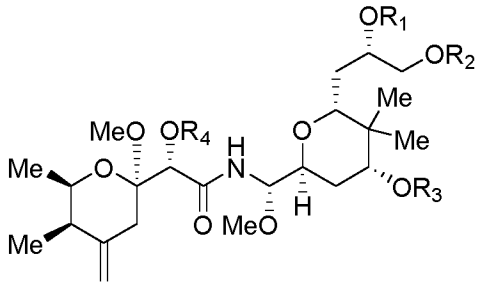
R_a is selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl, and heterocyclyl;

R_b is selected from substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl, and heterocyclyl;

R_c and R_d are independently selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl and heterocyclyl;

with the proviso that R_1 and R_2 are not simultaneously methyl.

2. The compound according to claim 1, also having general formula **III** or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof



III

wherein R_1, R_2, R_3 and R_4 are as defined for formula **I** in claim 1.

3. The compound according to claim 1 or 2, wherein R_1 is selected from hydrogen and substituted or unsubstituted C_1-C_6 alkyl.

4. The compound according to claim 3, wherein R_1 is selected from hydrogen and methyl.

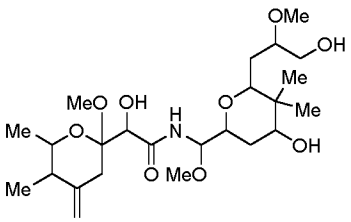
5. The compound according to any preceding claim, wherein R_2 is selected from hydrogen and $-C(=O)R_a$ where R_a is selected from substituted or unsubstituted C_1-C_6 alkyl.

6. The compound according to claim 5, wherein R_2 is selected from hydrogen and acetyl.

7. The compound according to any preceding claim, wherein R_3 and R_4 are independently selected from hydrogen and $-C(=O)R_a$, wherein R_a at each occurrence is independently selected from substituted or unsubstituted C_1-C_6 alkyl.

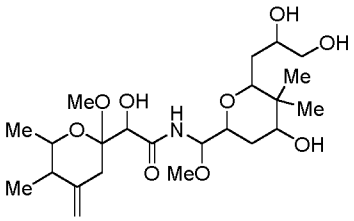
8. The compound according to claim 7 wherein R_3 and R_4 are independently selected from hydrogen and acetyl.

9. The compound according to claim 1 of formula:



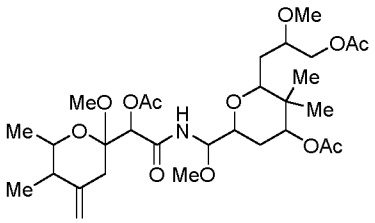
1

,



2

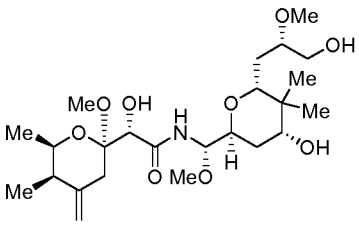
or



3

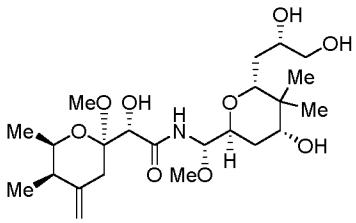
; or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof.

10. The compound according to claim 9 of formula:



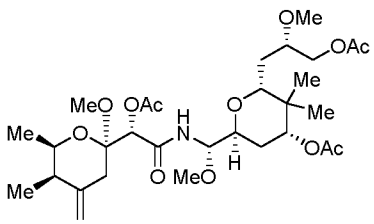
1'

,



2'

or



3'

; or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof.

11. A pharmaceutical composition comprising a compound as defined in any preceding claim, or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof, and a pharmaceutically acceptable carrier or diluent.

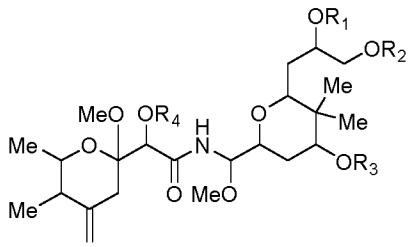
12. A compound as defined in any of claims 1 to 10, or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof, or a composition as defined in claim 11, for use as a medicament.

13. The compound or composition according to claim 12, for use as a medicament for the treatment of cancer.

14. Use of a compound as defined in any of claims 1 to 10, or pharmaceutically acceptable salts, tautomers, or stereoisomers thereof, in the preparation of a medicament for the treatment of cancer.

15. A method of treating a patient, notably a human, affected by cancer, which comprises administering to the affected individual in need thereof a therapeutically effective amount of a compound as defined in any of claims 1 to 10, or a pharmaceutically acceptable salt, tautomer, or stereoisomer thereof.

16. A process for obtaining a compound of formula **II** or a pharmaceutically acceptable salt, tautomer, or stereoisomer thereof



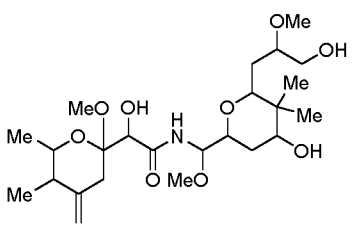
II

wherein

- R_1, R_2, and R_3 are each independently selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, $–C(=O)R_a$, $–C(=O)OR_b$ and $–(C=O)NR_cR_d$;
- R_4 is selected from hydrogen, $-C(=O)R_a$, $-C(=O)OR_b$, and $–C(=O)NR_cR_d$;
- R_a is selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl, and heterocyclyl;
- R_b is selected from substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl, and heterocyclyl;
- R_c and R_d are independently selected from hydrogen, substituted or unsubstituted C_1-C_{12} alkyl, substituted or unsubstituted C_2-C_{12} alkenyl, substituted or unsubstituted C_2-C_{12} alkynyl, aryl and heterocyclyl;

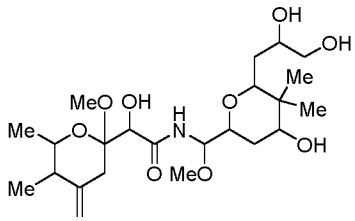
the process comprising the steps of:

- culturing the wild type marine bacterial strain PHM005 or their mutants under suitable conditions to produce compounds **1** and/or **2** of formula:



1

,



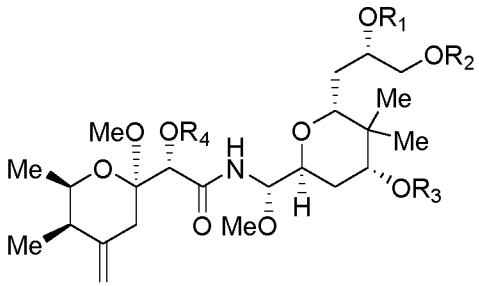
2

;

- isolating compounds **1** or **2**; and, if needed,

-derivatizing compounds **1** or **2**.

17. The process according to claim 16, wherein the compound of formula **II** has also formula **IV**



IV

or a pharmaceutically acceptable salt, tautomer or stereoisomer thereof;

wherein R_1, R_2, R_3, and R_4 are as defined for formula **II** in claim 16.

18. Biologically pure strain PHM005, deposited under the accession number CECT-9225 in the Colección Española de Cultivos Tipo at the University of Valencia, Spain.

19. An isolated nucleic acid comprising the Lab biosynthetic gene cluster or being complementary to a sequence comprising the Lab biosynthetic gene cluster which is derived from *Labrenzia* sp. and in particular from strain PHM005.

20. The isolated nucleotide sequence according to claim 19 comprising:
a nucleotide sequence as shown in SEQ ID NO: 2; or
a nucleotide sequence which is the complement of SEQ ID NO: 2; or
a nucleotide sequence hybridising under highly stringent conditions to SEQ ID NO: 2 or to the complement thereof; or
a nucleotide sequence having at least 80% sequence identity with SEQ ID NO: 2 or with the complement thereof.

21. An isolated nucleic acid comprising nucleic acid fragments forming individual units and/or modules of the Lab biosynthetic gene cluster as shown in Figure 3.

22. A modular enzymatic system encoded by a nucleic acid sequence as defined in any of claims 19 to 21.

23. A modular enzymatic system according to claim 22 comprising one or more of a protein sequence selected from the group formed by the sequences SEQ ID NO: 3 to SEQ ID NO: 23 or a protein sequence having at least 80% sequence identity with these sequences.

24. A modular enzymatic system according to claim 22 or 23 having functional activity in the synthesis of pederin-like or onnamide-like compounds and/or a polyketide moiety and/or a nonribosomal peptide moiety.

25. A vector comprising a nucleic acid consisting essentially of the Lab biosynthetic gene cluster derived from *Labrenzia* sp. and in particular from strain PHM005.

26. A vector comprising a nucleic acid sequence according to any of claims 19 to 21.

27. A recombinant host cell or a transgenic organism comprising a nucleic acid according to any of claims 19 to 21 or containing a vector according to claim 25 or 26.

28. A recombinant host cell according to claim 27 which is a bacterial cell and in particular is a *Pseudomonas*, *Acinetobacter*, *Bacillus*, *Streptomyces, or E. coli* cell.

29. A method for producing pederin-like or onnamide-like compounds using a mutant of PHM005 or a recombinant host cell according to claim 27 or a transgenic organism according to claim 28 comprising the steps of:

- culturing the mutant of PHM005 or the recombinant host cell or the transgenic organism under conditions to express the Lab biosynthetic gene cluster; and

- isolating the produced pederin-like or onnamide-like compounds.

30. The method according to claim 29 wherein the product of the *lab*719 is expressed to provide an onnamide-like compound.

31. Use of a nucleic acid according to any of claims 19 to 21 in the preparation of a modified Lab biosynthetic gene cluster.

32. Use of a nucleic acid according to any of claims 19 to 21 in the preparation of a pederin-like compound.

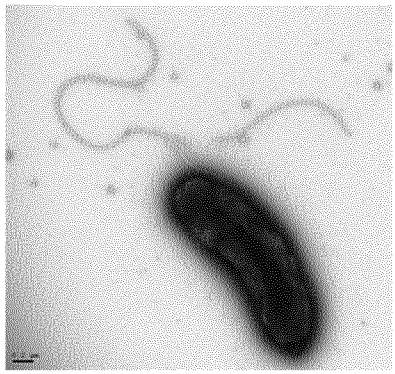


Figure 1

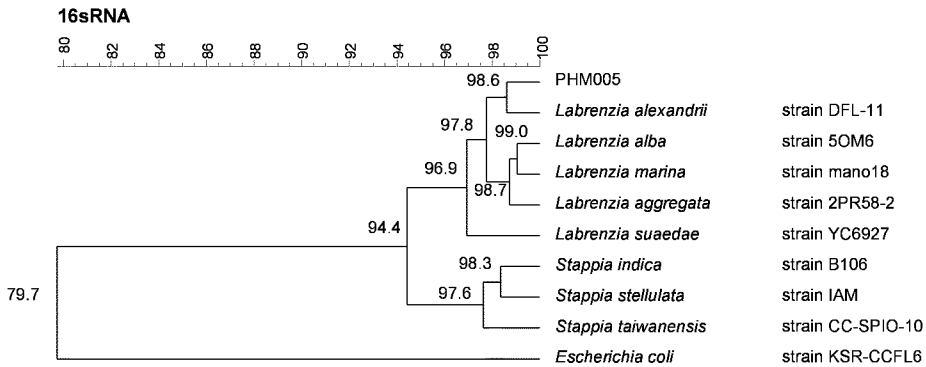


Figure 2

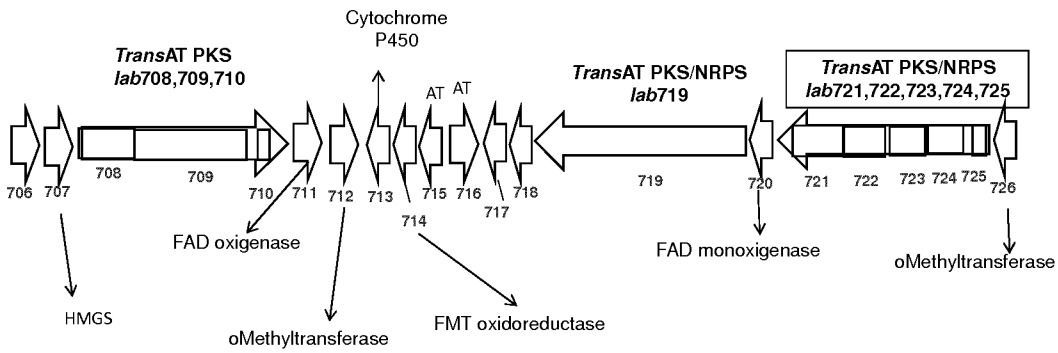


Figure 3

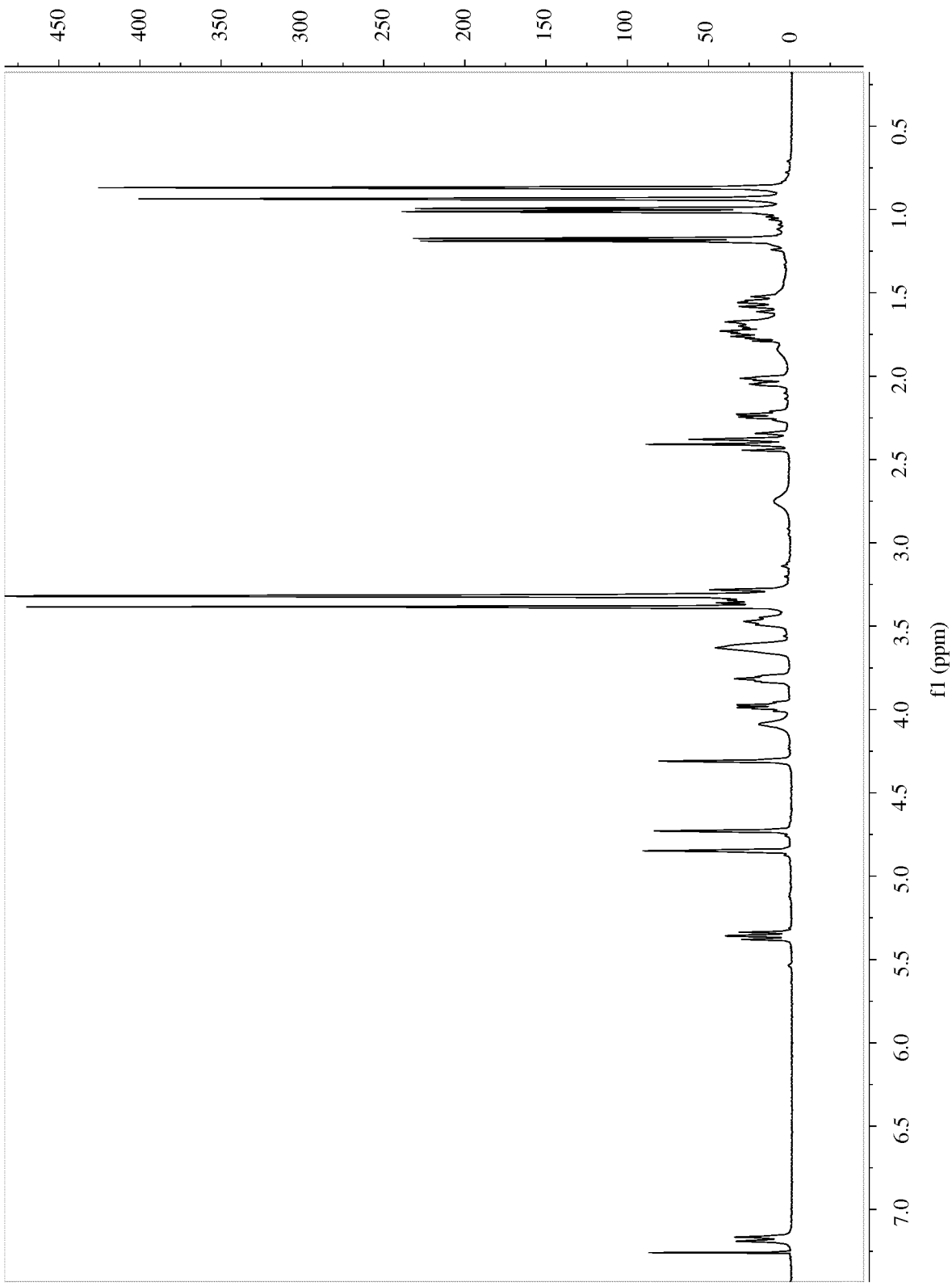

450
400
350
300
250
200
150
100
50
0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
f1 (ppm)

Figure 4

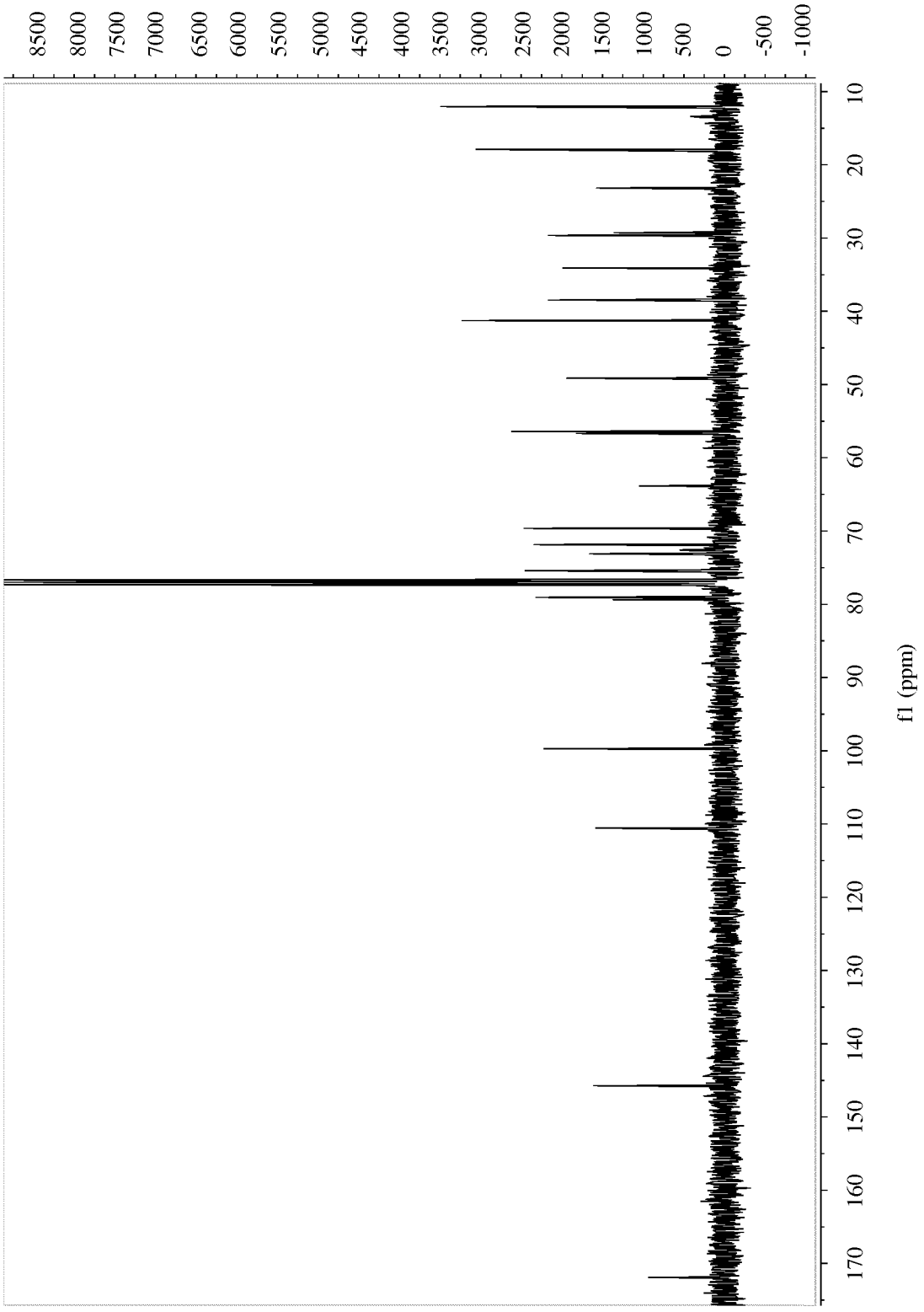

8500
8000
7500
7000
6500
6000
5500
5000
4500
4000
3500
3000
2500
2000
1500
1000
500
0
-500
-1000
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
f1 (ppm)

Figure 5

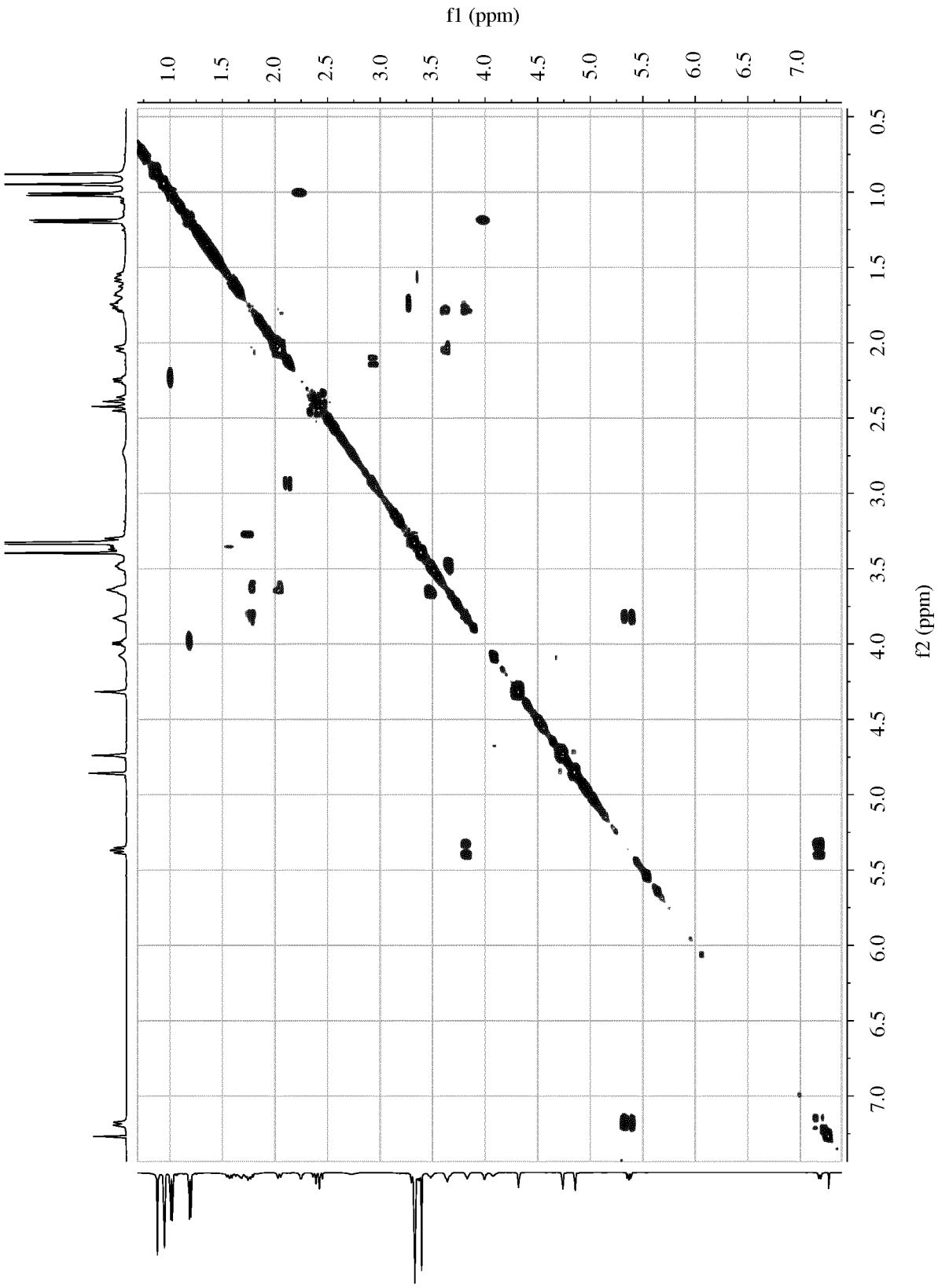

f1 (ppm)
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
f2 (ppm)
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0

Figure 6

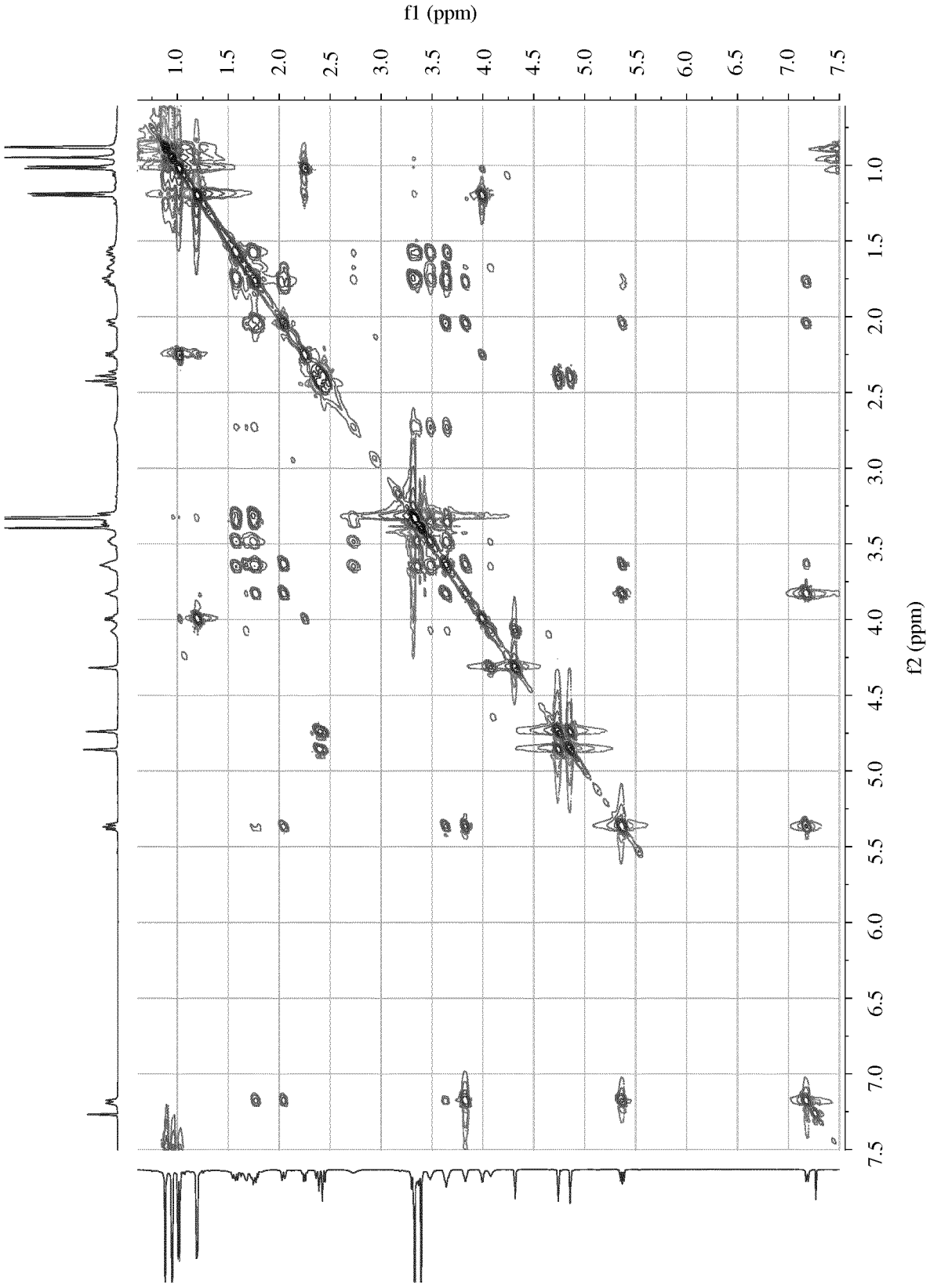


Figure 7

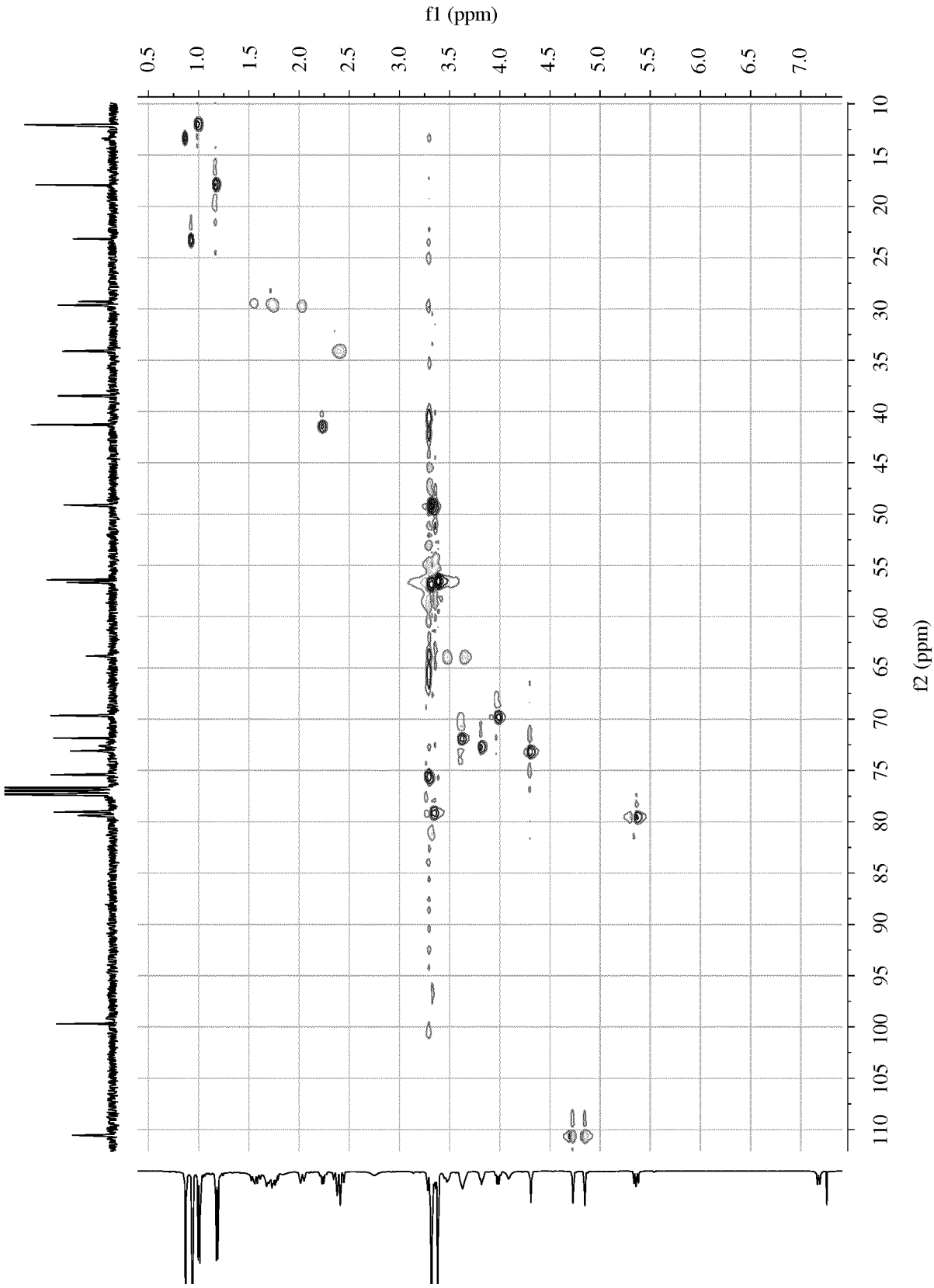

f1 (ppm)
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
f2 (ppm)
10
15
20
25
30
35
40
45
50
55
60
65
70
75
80
85
90
95
100
105
110

Figure 8

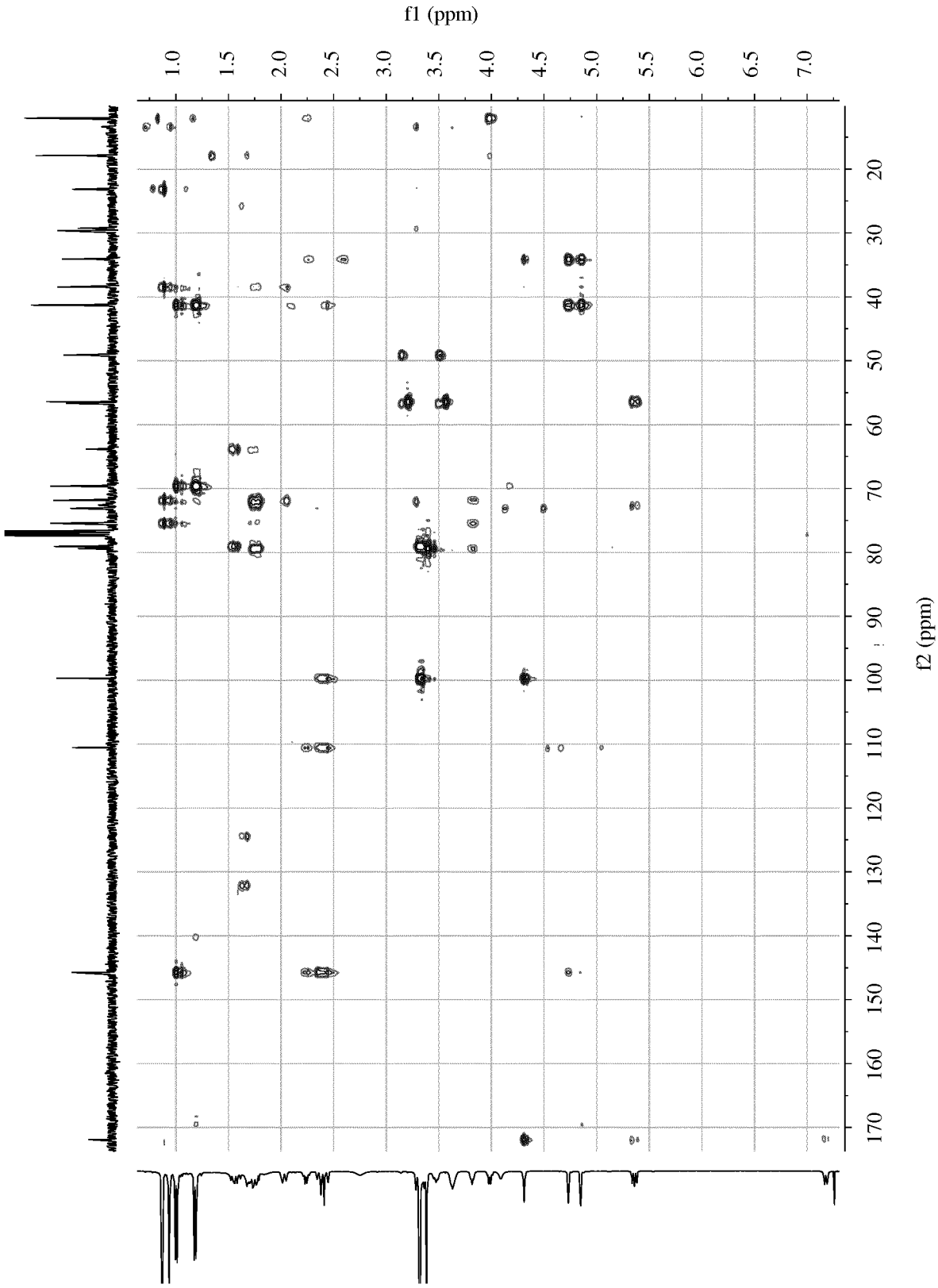

f1 (ppm)
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
f2 (ppm)
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170

Figure 9

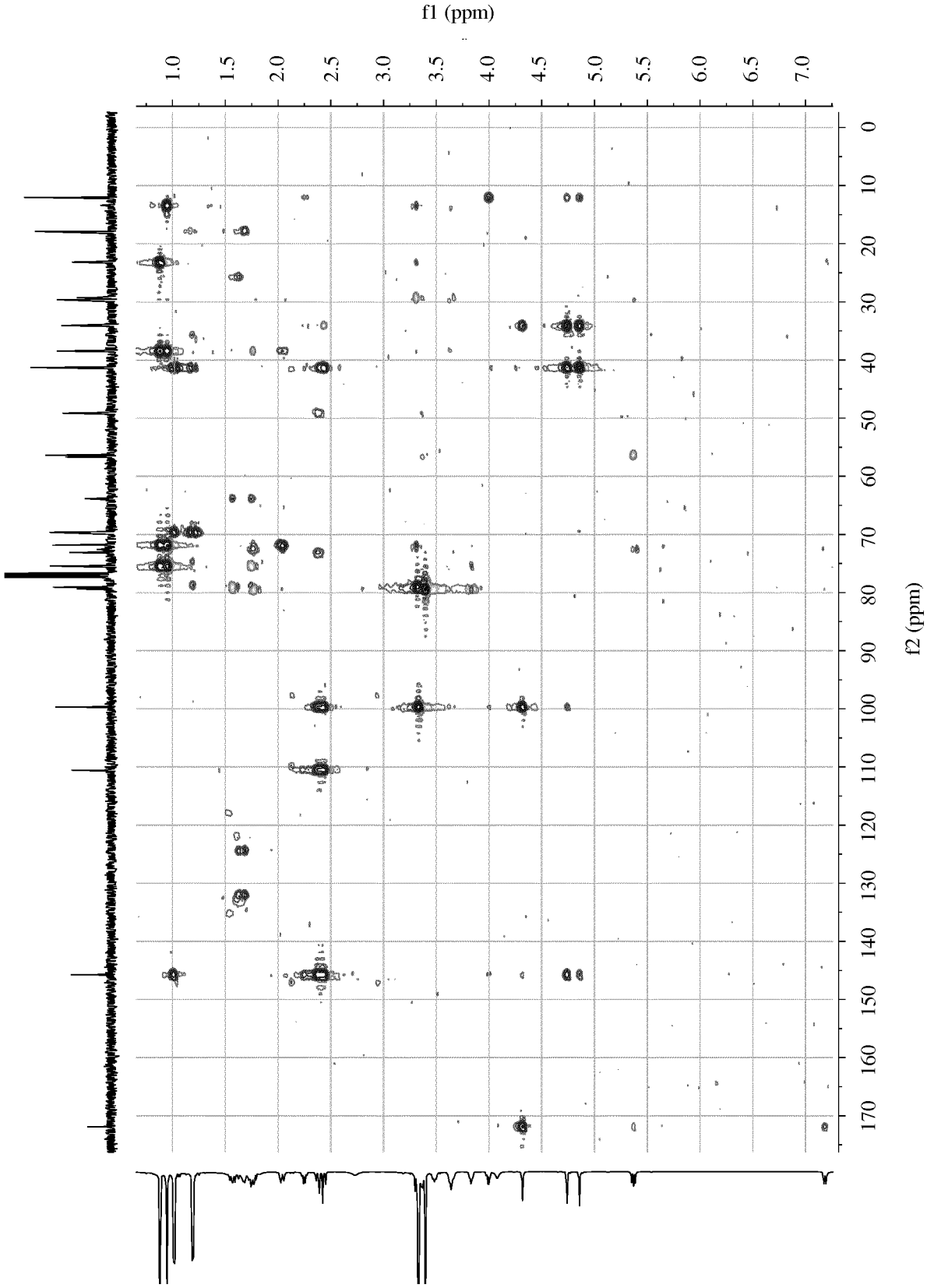

f1 (ppm)
1.0
1.5
2.0
2.5
3.0
3.5
4.0
4.5
5.0
5.5
6.0
6.5
7.0
f2 (ppm)
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170

Figure 10

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2018/056665

A. CLASSIFICATION OF SUBJECT MATTER
INV. C07D407/12 A61K31/351 A61P35/00
ADD.

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
C07D A61K A61P

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)
EPO-Internal, WPI Data, Sequence Search

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	NARQUIZIAN R ET AL: "The pederin family of antitumor agents: structures, synthesis and biological activity", 20000101, no. 32, 1 January 2000 (2000-01-01), pages 25-56, XP001145840,	1-8, 11-15, 25-31
A	chapter 2.2; table 1	9,10,16, 17

-/--

[X] Further documents are listed in the continuation of Box C.

[X] See patent family annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
5 July 2018	03/08/2018

Name and mailing address of the ISA/	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016	Gutke, Hans-Jürgen

4

Form PCT/ISA/210 (second sheet) (April 2005)

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	FANGHUI WU ET AL: "Total Synthesis of Pederin and Analogues", ANGEWANDTE CHEMIE INTERNATIONAL EDITION, vol. 50, no. 5, 1 February 2011 (2011-02-01), pages 1131-1134, XP055469240, ISSN: 1433-7851, DOI: 10.1002/anie.201006438	1-8, 11-15, 25-31
A	schemes 1-4	9,10,16, 17
A	WO 03/044186 A2 (MAX PLANCK GESELLSCHAFT [DE]; PIEL JOERN [DE]) 30 May 2003 (2003-05-30) cited in the application	1-17
X	DATABASE EMBL [Online] 15 August 2016 (2016-08-15), "Labrenzia sp. strain PHM005 16S ribosomal RNA gene, partial sequence.", XP002782808, retrieved from EBI accession no. EM_STD:KX661380 Database accession no. KX661380 sequence	18-24

INTERNATIONAL SEARCH REPORT

International application No.
PCT/EP2018/056665

Box No. II Observations where certain claims were found unsearchable (Continuation of item 2 of first sheet)

This international search report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☐ Claims Nos.:
because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box No. III Observations where unity of invention is lacking (Continuation of item 3 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

see additional sheet

1. ☐ As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fees, this Authority did not invite payment of additional fees.

3. ☒ As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims for which fees were paid, specifically claims Nos.:

25-31

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☒ The additional search fees were accompanied by the applicant's protest and, where applicable, the payment of a protest fee.

☐ The additional search fees were accompanied by the applicant's protest but the applicable protest fee was not paid within the time limit specified in the invitation.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

This International Searching Authority found multiple (groups of) inventions in this international application, as follows:

1. claims: 1-15(completely); 16, 17(partially)

   Pederin derivatives of formula (I)

1.1. claims: 1-17(partially)

   Pederin derivatives of formula (I) wherein R1 = H for use as anticancer agents and preparation process thereof

1.2. claims: 1-17(partially)

   Pederin derivatives of formula (I) wherein R1 is not H for use as anticancer agents and preparation process thereof

---

2. claims: 16, 17(partially)

   Alternative preparation process of pederin

---

3. claims: 18-32

   Biologically pure strain PHM005, isolated nucleic acid thereof and related subject-matter (modular enzymatic system, vector, recombinant host cell or transgenic organism) as well as use of the isolated nucleic acid in the preparation of a modified Lab bioynthetic gene cluster and/or a pederin-like compound

---

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/EP2018/056665

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 03044186 A2	30-05-2003	AT 484580 T	15-10-2010
		AU 2002356704 A1	10-06-2003
		EP 1448767 A2	25-08-2004
		JP 2005509435 A	14-04-2005
		US 2005118590 A1	02-06-2005
		WO 03044186 A2	30-05-2003