# Diterpenoids from the stem bark of Croton megalocarpoides Friis & M. G. Gilbert

Moses Langat, Beth Ndunda, Caitlin Salter, Babatope Odusina, Sani Isyaka, Eduard Mas-Claret, Patricia Onocha, Jacob Midiwo, Jean-Marc Nuzillard, Dulcie Mulholland

**Diterpenoids from the stem bark of *Croton megalocarpoides* Friis & M. G. Gilbert**

Moses K Langat[a,b,c,*], Beth M Ndunda[b,d], Caitlin Salter[b], Babatope O Odusina[e], Sani M Isyaka[b], Eduard Mas-Claret[b], Patricia A Onocha[f], Jacob O Midiwo[c], Jean-Marc Nuzillard[g], Dulcie A Mulholland[b,c]

[a] *Jodrell Laboratory, Natural Capital and Plant Health Department, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3DS, United Kingdom*

[b] *Natural Products Research Group, Department of Chemistry, Faculty of Engineering and Physical Sciences, University of Surrey, Guildford, GU2 7XH, United Kingdom*

[c] *School of Chemistry and Physics, University of KwaZulu-Natal, Durban, 4041, South Africa*

[d] *Department of Chemistry, Nairobi University, P Bag 30197-00100, Nairobi, Kenya*

[e] *Department of Chemical Sciences, Tai Solarin University of Education, Ijagun, Ijebu Ode, Ogun State, Nigeria*

[f] *Department of Chemistry, University of Ibadan, Ibadan, Nigeria*

[g] *Université de Reims Champagne Ardenne, CNRS, ICMR UMR 7312, 51097 Reims, France*

*Keywords:*

*Croton megalocarpoides* Friis & M. G. Gilbert

Euphorbiaceae

Megalocarpoidolide I

Megalocarpoidolide J

Megalocarpoidolide K

12-*epi*-crotonzambefuran A

1-*trans*-*p*-hydroxycoumaroyl–geranylgerani-1-ol

Logic for Structure Determination

*Corresponding author.

*E mail address*: m.langat@kew.org

**Abstract**

Five previously undescribed compounds, megalocarpoidolide I (**1**), megalocarpoidolide J (**3**), 12-*epi*-crotonzambefuran A (**4**), megalocarpoidolide K (**5**), 1-*trans*-*p*-hydroxycoumaroyl–geranylgerani-1-ol (**6**) were isolated from the stem bark of *Croton megalocarpoides* Friis & M. G. Gilbert. The known *ent*-trachyloban-18-ol, megalocarpoidolide B, megalocarpoidolide C (**2**), megalocarpoidolide H, crotocorylifuran, 7,8-dehydrocrotocorylifuran, 1,2-dehydrocrotocorylifuran-2-one, acetyl aleuritolic acid, lupeol, *N*-*trans*-*p*-coumaroyl-3',4'-dihydroxyphenylethylamine, dodecyl *trans*-ferulate and lignoceryl *trans*-ferulate were also isolated. The structures of the compounds were determined using NMR and IR spectroscopy and HRMS. The structure of compound **1** was determined using Logic for Structural Determination (LSD). Compounds **1**, **2** and **3** were evaluated against the NCI60 panel of human tumour cell lines at 10μM level but found to be inactive.

## 1. Introduction

In our continued search for bioactive compounds from African *Croton* species (Aldhaher et al., 2017; Isyaka et al., 2020; Langat et al., 2011; Langat et al., 2012; Mulholland et al., 2010) we report the phytochemistry of the stem bark of the Kenyan endangered arborescent *C. megalocarpoides* Friis & M. G. Gilbert. This plant is endemic to the semi-evergreen coastal bush lands or forests of Kenya and South Somalia (Friis and Gilbert, 1984; Beentje, 1994). It is a monoecious shrub or tree, growing up to 8 meters tall in rocky places. The bark of this species is grey and fissured into rectangular scales. Friis & Gilbert (1984) reported that *C. megalocarpoides* is related to *C. megalocarpus* Hutch., *C. mayumbensis* J. Leonard, and *C. mubango* Mull. Arg (Friis and Gilbert, 1984). We recently reported the chemical constituents of the roots of *C. megalocarpoides*, from which twenty-two compounds were reported. These included twelve *ent*-clerodanes, megalocarpoidolide A, megalocarpoidolide B, megalocarpoidolide C, 12-*epi*-crotocorylifuran, crotocorylifuran, 8β-hydroxycrotocorylifuran, crotocorylifuran-2-one, megalocarpoidolide D, 7,8-dehydrocrotocorylifuran, megalocarpoidolide E, megalocarpoidolide F, megalocarpoidolide G and megalocarpoidolide H. Three abietane diterpenoids isolophanthin A, isolophanthin E and abietic acid, four *ent*-trachylobane diterpenoids *ent*-3α,18-dihydroxytrachylobane, *ent*-trachyloban-18-ol, *ent*-trachyloban-18-oic acid, *ent*-3α-hydroxytrachyloban-18-al and the triterpenoids, lupeol and acetyl aleurotolic acid were also isolated from the roots (Ndunda et al., 2016). These compounds were found to be inactive against a panel of human pathogenic bacterial and fungal

strains including *Candida albicans*, *Candida glabrata*, *Candida krusei*, *Aspergillus fumigatus*, *Cryptococcus neoformans*, *Staphylococcus aureus*, methicillin-resistant *S. aureus*, *Escherichia coli*, *Pseudomonas aeruginosa*, and *Mycobacterium intracellulare* at the highest tested concentration of 20 μg/mL (~ 50 μM). The compounds were also inactive against *Plasmodium falciparum* D6 and W2 strains at 4.76 μg/mL (~ 12 μM) and did not exhibit cytotoxic activity towards VERO cells up to 4.76 μg/mL (~ 12 μM) (Ndunda et al., 2016).

## 2. Results and discussion

Five previously undescribed diterpenoid derivatives, megalocarpoidolide I (**1**), megalocarpoidolide J (**3**), 12-*epi*-crotonzambefuran A (**4**), megalocarpoidolide K (**5**), 1-*trans*-*p*-hydroxycoumaroyl–geranylgerani-1-ol (**6**) were extracted from the stem bark of *Croton megalocarpoides* Friis & M. G. Gilbert. The known compounds megalocarpoidolide B, megalocarpoidolide C (**2**), megalocarpoidolide H, crotocorylifuran, 7,8-dehydrocrotocorylifuran, 1,2-dehydrocrotocorylifuran-2-one, 18,19,-dimethoxycarbonyl-3α-acetoxy-4β-hydroxy-15,16–epoxy-clerodane-7,13(16),14-triene-12,20-olide, *ent*-trachyloban-18-ol (Ndunda et al., 2016), acetyl aleuritolic acid, lupeol, *N*-*trans*-*p*-coumaroyl-3',4'-dihydroxyphenylethylamine, dodecyl *trans*-ferulate and lignoceryl *trans*-ferulate were isolated alongside the novel compounds. The structures of the compounds were determined using NMR, IR and HRMS techniques. The structure of compound **1** was determined using the Logic for Structure Determination (LSD) protocol (Nuzillard and Massiot, 1991; Nuzillard et al., 2018). Compounds **1**, **2** and **3** were evaluated against the NCI60 panel of human tumour cell lines at 10μM level but found to be inactive.

Compound **1** was isolated as a yellow oil and determined to have a molecular formula of $C_{21}H_{24}O_5$ from HR-ESIMS and 10 degrees of unsaturation. The IR spectrum showed absorption bands for two ester carbonyl groups at 1712 $cm^{-1}$ and 1764 $cm^{-1}$ (α, β-unsaturated). The $^{1}H$, $^{1}H$-$^{1}H$ COSY, $^{13}C$, $^{1}H$-$^{13}C$ HSQC and $^{1}H$-$^{13}C$ HMBC spectroscopic data were used as input for the automatic structure elucidation software, LSD (Nuzillard and Massiot, 1991; Nuzillard et al., 2018). The LSD protocol proposed the structure of **1** (Figure 1), which is postulated to arise from megalocarpolide C (**2**) (Figure 1) that was also isolated in this study and previously from the roots of this plant (Ndunda et al., 2016). Raw NMR data, 1D and 2D spectra, annotated drawings of spectra, input file to LSD and drawing of the solution structures are permanently available at https://doi.org/10.5281/zenodo.3757401 and temporarily at https://www.dropbox.com/s/hgdeo719uiska1r/Megalocarpoidolide_I.zip?dl=0. LSD proposes four solutions for the chemical structure of **1**, with only one being acceptable according to the

comparison between the experimental $^{13}C$ NMR chemical shifts and those predicted by nmrshiftdb2 (https://nmrshiftdb.nmr.uni-koeln.de/, Steinbeck and Kuhn, 2004). The $^{1}H$ NMR spectrum gave characteristic resonances of a β-substituted furan ring with proton resonances at $\delta_H$ 6.38 (br, d, $W_{1/2}$ = 3.7 Hz), $\delta_H$ 7.42 (dd, $J$ = 1.7, 1.7 Hz) and $\delta_H$ 7.35 (br, d, $W_{1/2}$ = 3.5 Hz) for H-14, H-15 and H-16 respectively (Table 1) and the $sp^2$ carbon resonances at $\delta_C$ 126.0 (C-13), 112.2 (C-14), 144.5 (C-15) and 140.3 (C-16). The $^{13}C$ NMR spectrum showed twenty-one carbon resonances including additional $sp^2$ carbon resonances at $\delta_C$ 137.4 (C-3), 137.1 (C-1), 131.6 (C-4) and 129.5 (C-2), carbonyl carbon resonances at $\delta_C$ 178.5 and 168.5, and an oxygenated carbon resonance at $\delta_C$ 95.2 (C-4). The two carbonyl groups and the four double bonds accounted for 6 DBE suggesting that compound **1** was tetracyclic (Table 2). As with megalocarpoidolide C (**2**), the 3H-17 methyl group doublet proton resonance at $\delta_H$ 0.90 (d, 6.5 Hz) showed correlations in the HMBC spectrum with C-9 ($\delta_C$ 54.8, C), C-8 ($\delta_C$ 35.6, CH) and $\delta_C$ 30.3 (C-7, $CH_2$) resonances. In addition, three coupled double bond proton resonances at $\delta_H$ 6.31 (br d, $W_{1/2}$ = 1.8 Hz), $\delta_H$ 6.28 (dd, $J$ = 1.5, 5.1 Hz) and $\delta_H$ 7.13 (d, $J$ = 5.1 Hz) for H-1, H-2 and H-3 respectively were observed in the $^{1}H$ NMR spectrum for **1**. The corresponding carbon resonances, as seen in the multiplicity-edited HSQC spectrum, for C-1, C-2 and C-3 occurred at $\delta_C$ 137.1, 129.5 and 137.4 respectively. The H-3 proton resonance at $\delta_H$ 7.13 (d, $J$ = 5.1 Hz) showed correlations in the HMBC spectrum with the C-1 carbon resonance ($\delta_C$ 137.1), a fully substituted carbon resonance at $\delta_C$ 131.6 for C-4, and a carbon resonances at $\delta_C$ 36.6 ($CH_2$) and $\delta_C$ 168.5 that were assigned to C-19 and C-18 respectively. The C-19 methylene carbon resonance corresponded in the HSQC spectrum with a pair of proton doublets at $\delta_H$ 2.79 and $\delta_H$ 2.83 (d, 13.0 Hz). These two proton resonances showed correlations with the C-3, C-4, C-5, C-6, C-10 and C-18 resonances in the HMBC spectrum which also showed correlations between the H-1 and the C-5 ($\delta_C$ 95.2) and C-9 ($\delta_C$ 54.8), between the H-2 and C-10 ($\delta_C$ 56.3), between the two H-11 and C-20, between the two H-19 and C-10 ($\delta_C$ 56.3), C-18 ($\delta_C$ 168.5) and C-6 ($\delta_C$ 36.6), between the two H-11 and H-8 and the C-20 ($\delta_C$ 178.5) and between the twoH-12 and C-14 ($\delta_C$ 112.2) and C-16 ($\delta_C$ 140.3) resonances. The NOESY spectrum was used to assign the relative configuration of **1**. Correlations were seen between the H-8 and H-10, andH-8 and 2H-11 resonances and determined to be on the same face of the molecule (Figure 3). A specific rotation value of $[\alpha]_{22,D}$ -53° was measured for this compound indicating that compound **1** belonged to the *ent*-series as with the reported megalocarpoidolide C and the *ent*-clerodane diterpenoids earlier isolated from the roots of *C. megalocarpoides* and laevinoid from *Croton laevigatus* (Ndunda et al., 2016; Wang et al., 2013). *Ent*-clerodanes with the rare bicyclo[5.4.0]undecane skeleton have been reported from

the Mexican *Croton cortesianus* (Siems et al., 1992), *Portulaca pilosa* (Portulacaceae) (Ohsaki et al., 1991; Ohsaki et al., 1995), *Dodonaea viscosa* (Sapindaceae) (Niu et al., 2010) and *cis*-clerodane analogues from the liverwort *Scapania pava* (Guo et al., 2012). We propose a biosynthesis of compound **1** (Figure 2) starting from the 19-hydroxy precursor to megalocarpoidolide C (**2**).

Compound **3** was assigned a molecular formula of $C_{22}H_{24}O_9$ from its HR-ESIMS. The IR spectrum gave absorption bands at 1772 $cm^{-1}$, 1738 $cm^{-1}$, 1700 $cm^{-1}$ (C=O stretches) and 1644 $cm^{-1}$ (C=C stretch). Compound **3** was the 1,2-dehydro-3-deacetyl derivative of megalocarpoidolide H (Ndunda et al., 2016), also isolated in this work. As for **1** and **2**, the $^{1}H$ and $^{13}C$ NMR spectra of compound **3** showed characteristic resonances of a β-substituted furan ring. The H-3 proton resonance occurred at $\delta_H$ 3.93 (dd, $J$ = 12.4, 1.8 Hz) in compound **3** but at $\delta_H$ 5.08 (t, $J$ = 2.6 Hz) for megalocarpoidolide H with the 3α-acetate. The COSY spectrum showed the following coupled sequence: H-3 / H-2 ($\delta_H$ 6.08, dt, $J$ = 10.4, 12.4 Hz) / H-1 ($\delta_H$ 5.66, dd, $J$ = 10.4, 1.7 Hz) / H-10 ($\delta_H$ 3.06, m). The $^{1}H$ and $^{13}C$ NMR resonances of ring B of compound **3** were comparable to megalocarpoidolide H with H-7 of the $\Delta^7$-double bond at $\delta_H$ 5.73 (m) and the 3H-17 vinyl methyl group proton resonance at $\delta_H$ 1.85, (br dd, $W_{1/2}$ = 4.4 Hz). The NOESY spectrum showed correlations between the H-3 and H-10 and H-10 and 2H-11 resonances confirming the relative configurations at C-3 and C-9. The configuration at C-12 was assigned as *R* due to the correlation between the H-12 and 3H-17 proton resonances (Figure 4). A specific rotation value of $[\alpha]_{22,D}$ -63° was measured for this compound and indicated that compound **3** belonged to the *ent*-series as with the reported compound **1**, megalocarpoidolide C (**2**) and megalocarpoidolide H (Ndunda et al., 2016).

Compound **4** was determined to have a molecular formula of $C_{22}H_{24}O_7$ from HR-ESIMS. The IR spectrum showed absorption bands for two ester carbonyl groups at 1699 $cm^{-1}$ and 1730 $cm^{-1}$ (α, β-unsaturated). As for compound **3**, compound **4** had a 20,12*R*-lactone group, as shown by correlations in the HMBC spectrum between the C-9 ($\delta_C$ 52.1), C-13 ($\delta_c$ 125.1), C-14 ($\delta_c$ 108.4), C-16 ($\delta_c$ 139.8) and C-20 ($\delta_c$ 176.3) and the H-12 ($\delta_H$ 5.45, t $J$ = 8.6 Hz) resonances and in the NOESY spectrum by correlations between the H-12 and 3H-17 resonances ($\delta_H$ 1.12, d $J$ = 7.0 Hz). The HMBC spectrum showed correlations between the C-9 and the 3H-17 methyl group proton doublet and the alkene H-1 resonance ($\delta_H$ 5.97, dd $J$ = 3.4, 9.4 Hz) and the COSY spectrum showed coupling between the H-10 (($\delta_H$ 2.87, dd $J$ = 3.4, 3.3) / H-1 / H-2 ($\delta_H$ 6.17, ddd $J$=3.3, 5.2, 9.4 Hz) / H-3 ($\delta_H$ 6.98, d $J$=5.2 Hz) resonances, confirming the presence of a 1,3-diene as in compound **1**. The presence of methyl esters at C-18 and C-19 were indicated by resonances at $\delta_C$ 166.7 (C-18) and $\delta_C$ 171.4 (C-19) which showed correlations in the HMBC

spectrum with methoxy group proton resonances at $\delta_H$ 3.72 (s) and $\delta_H$ 3.60 (s) respectively. Compound **4** is a 12-epimer of crotozambefuran A (Ngadjui et al., 2002). Slight differences in the $^{13}C$ NMR chemical shifts between compound **4** and crotozambefuran A were observed for C-12, C-17, C-19 and C-20 which occurred at $\delta_C$ 72.1, $\delta_C$ 17.4, $\delta_C$ 171.4 and $\delta_C$ 176.3 respectively for **4** and $\delta_C$ 73.0, $\delta_C$ 16.5, $\delta_C$ 172.5 and $\delta_C$ 177.8 for crotozambefuran A. The NOESY spectrum for compound **4**, showed correlations between H-12 and H-1, and H-10 and H-8 proton resonances. The NOESY spectrum also showed a correlation between the H-12 and 3H-17 resonances ($\delta_H$ 1.12, d *J* = 7.0 Hz) (Figure 5). The structure of compound **4** was determined to be 12-*epi*-crotonzambefuran A.

Compound **5** was assigned a molecular formula of $C_{20}H_{20}O_5$ from its HR-ESIMS with 11 degrees of unsaturation. The IR spectrum showed absorption bands at 1759 $cm^{-1}$, 1717 $cm^{-1}$ (C=O stretches) and 1643 $cm^{-1}$ (C=C stretch). This compound only differed from compound **4** in having lost the C-19 methyl ester and in the formation of a 5,10-double bond shown by the loss of the H-10 resonance and the presence of two fully substituted carbons at $\delta_C$ 138.1 and $\delta_C$ 130.4 for C-5 and C-10 respectively, to give an aromatic ring A. The 3H-17 resonance showed correlations with the proton resonances of the β-substituted furan ring (Figure 6). A specific rotation value of $[\alpha]_{22,D}$ -33˚ was measured for this compound. Compound **5** was named megalocarpoidolide K and would have arisen by oxidation of the 19-methyl group to a carboxylic acid, followed by decarboxylation and aromatization. A related 19-nor clerodane was isolated from *Croton cortesianus* (Siems et al., 1992), but they were not able to confirm the structure due to the small amount of material available.

Compound **6**, isolated as a white solid, was assigned a molecular formula of $C_{29}H_{40}O_3$ from LRMS and had 10 degrees of unsaturation. The IR spectrum showed absorption bands at 3429 $cm^{-1}$ (O-H alkene stretch) and 1712 $cm^{-1}$ (C=O stretches). The NMR spectra showed this was a *p*-coumaryl ester of geranyl geraniol. The *p*-coumaryl part was shown by a pair of *ortho*-coupled doublets at $\delta_H$ 7.42 (d *J* = 8.6 Hz) and $\delta_H$ 6.84 (d *J* = 8.6 Hz) ascribable to H-5'/9' and H-6'/8' respectively and $\delta_H$ 7.63 (H-2', d *J* = 16.0 Hz) and $\delta_H$ 6.31 (H-3', d *J* = 16.0 Hz) for protons of a *trans* double bond (Renata et al., 2012). An ester carbonyl carbon resonance ($\delta_C$ 167.8) was assigned as C-1' as it showed correlations in the HMBC spectrum with the two H-1 resonances $\delta_H$ (4.72, d J = 7.1) of geranyl geraniol. The $^1H$ NMR spectrum showed the typical five downfield vinyl methyl proton resonances at $\delta_H$ 1.74, $\delta_H$ 1.68, $\delta_H$ 1.60, $\delta_H$ 1.59 and $\delta_H$ 1.59 and alkene proton resonances at $\delta_H$ 5.42 (dd *J* = 7.1, 1.0 Hz), $\delta_H$ 5.10 (m 4H) consistent with geranyl geraniol (Fedeli et al., 1966). Thus compound **6** was determined to be 1-*trans*-*p*-hydroxycoumaroyl–geranylgerani-1-ol.

Compounds **1**, **2** and **3** were evaluated against the NCI 60 panel of human tumour cell lines (National Cancer Institute, 2019) which is derived from nine cancer cell types including leukaemia, lung, melanoma, colon, CNS, ovary, renal, prostate and breast cancers (Shoemaker, 2006). Compounds were evaluated at a single dose of $10^{-5}$ M. Results for the single dose screen are given in the Supplementary Data (Figures S4 to S6). The three compounds were found to be inactive at this concentration.

## 3. Experimental section

### 3.1 *General experimental procedure*

Optical rotations were measured on a JASCO P-1020 polarimeter. FTIR spectra were recorded using a Perkin-Elmer (2000) spectrometer. 1D and 2D NMR spectra were recorded in $CDCl_3$ on a 500 or 400 MHz Bruker AVANCE NMR instruments at room temperature. Chemical shifts (δ) are expressed in ppm and were referenced against the solvent resonances at $\delta_H$ 7.26 and $\delta_C$ 77.23 ppm for $^1H$ and $^{13}C$ NMR for $CHCl_3$, at $\delta_H$ 1.94 and $\delta_C$ 128.4 ppm for $^1H$ and $^{13}C$ NMR for $CD_3CN$ and at $\delta_H$ 4.87 and $\delta_C$ 49.15 ppm for $^1H$ and $^{13}C$ NMR for $CD_3OD$ respectively. HR-ESI mass spectra were recorded a Micromass Quattro Ultima mass spectrometer using a Waters Alliance HPLC to introduce samples (University of Surrey) or a Waters Xevo G2-S (National Mass Spectrometry Facility, Swansea). Flash chromatography were undertaken using Grace Reveleris X2 Flash Column Chromatography instrument. Grace Reveleris © silica 120 g cartridges was used during fractionation. Purity of compounds was monitored *via* thin layer chromatography (TLC) using pre-coated aluminium-backed plates (silica gel 60 $F_{254}$, Merck) and compounds were visualised by UV radiation at 254 nm and then using an anisaldehyde spray reagent (1% *p*-anisaldehyde:2% $H_2SO_4$: 97% cold MeOH) followed by heating. Final purifications used preparative thin layer chromatography (Merck 818133) and gravity column chromatography that was carried out using different column sizes (1-2 cm diameter), which were packed with silica gel (Merck Art. 9385) in selected solvent systems or Sephadex (LH 20) in $CH_3OH/CH_2Cl_2$.

*3.2 Plant material*

The stem bark of *Croton megalocarpoides* Friis & M. G. Gilbert was collected from the Kenya coastal region. The plant was authenticated at the University of Nairobi herbarium in the school of Biological Science by Patrick K Mutiso. A voucher specimen BN 2009/8 was lodged at the herbarium.

*3.3 Extraction and isolation*

The stem bark (700 g) of *C. megalocarpoides* was cut into small pieces, air dried, then ground using a laboratory hammer mill with a sieve diameter of 1 mm and stored in a well-ventilated environment. Ground powder (700 g) was sequentially extracted on a shaker at room temperature, with *n*-hexane, $CH_2Cl_2$ and $CH_3OH$ (2 L) each for 24 h each and then filtered. The extracts were concentrated and evaporated to dryness using a rotary evaporator to yield 21.6 g (3.1 %) *n*-hexane, 15.0 g (2.1 %) $CH_2Cl_2$ and 27.5 g (3.9 %) $CH_3OH$ extracts. The extracts were each adsorbed onto 30 g silica gel and subjected to flash chromatographic separation. Flash chromatography were undertaken using Grace Reveleris ® X2 Flash Column Chromatography instrument. Grace Reveleris ® silica 120 g cartridges were used during fractionation. Flash chromatography was carried out firstly using a hexane/$CH_2Cl_2$ step gradient starting with 100% hexane and gradually increasing the concentration of $CH_2Cl_2$ to 100% and secondly, starting from 100% $CH_2Cl_2$ and increasing the polarity to 80:20 $CH_2Cl_2$: $CH_3OH$ for the *n*-hexane and $CH_2Cl_2$ extracts. The gradient elution for the $CH_3OH$ extract for the second solvent system's polarity was increased from 100% $CH_2Cl_2$ to 60:40 $CH_2Cl_2$: $CH_3OH$. 75 mL fractions were collected at a time. Compounds were detected using a wavelength of 254 nm. A detailed purification protocol is provided in the Supplementary Information (Figures S1 – S3). From the *n*-hexane extract, fractions 14-20 obtained by flash chromatography were pooled and fractioned using a 1cm silica gel column to give subfractions 15-17 determined as *ent*-trachyloban-18-ol (1.6 mg) (Kapingu et al., 2000), subfractions 10-12 determined as dodecyl *trans*-ferulate (1.6 mg), subfraction 22 determined as lupeol (22.0 mg) and subfractions 30-33 determined as acetyl aleuritolic acid (1.6 mg). From the $CH_2Cl_2$ extract, fractions 72-73 obtained by flash chromatography were pooled and determined to be lignoceryl *trans*-ferulate (8.3 mg). Fractions 78-79 obtained by flash chromatography were pooled and fractioned using a 1cm silica gel column to give subfractions 54-56 determined as lupeol (14.7 mg). Fractions 92-93 obtained by flash chromatography were pooled and fractioned using a 1cm silica gel column to give subfractions 4-5 determined as 1-*trans*-*p*-hydroxycoumaroyl–geranylgerani-1-ol (**6**, 1.7 mg), subfractions 16–19 determined as 1,2-

dehydrocrotocorylifuran-2-one (6.1 mg) (Ndunda et al., 2016), and subfractions 24-27 determined as compound **4**, 12-*epi*-crotozambefuran A (1.6 mg). Fractions 94-95 obtained by flash chromatography were pooled and fractioned using a 1cm silica gel column to give subfractions 53-59, determined as crotocorylifuran (1.3 mg) (Ngadjui et al., 2002). Fractions 98–99 were pooled and fractioned using 2 cm silica gel column to give subfractions 59-61, determined as 7,8-dehydrocrotocorylifuran (1.8 mg) (Ndunda et al., 2016). Fractions 106–109 were pooled and fractioned using 2 cm silica gel column to give subfractions 29-32 which was determined as 18,19-dimethoxycarbonyl-3α-acetoxy-4β-hydroxy-15,16–epoxy-clerodane-7,13(16),14-triene-12,20-olide (5.4 mg) (Ndunda et al., 2016). From the $CH_3OH$ extract fractions 43-44 obtained by flash chromatography were pooled and fractioned using a 1cm silica gel column to give subfractions 16-19 determined as compound **5** (2.5 mg), subfractions 25-28 determined as compound **1** (2.5 mg), subfractions 36-44 determined as megalocarpoidolide C (**2**) (8.1 mg) (Ndunda et al., 2016) and subfraction 61 was determined to be megalocarpoidolide B (10.9 mg) (Ndunda et al., 2016). Fractions 77-78 obtained by flash chromatography were pooled and fractioned using a 1cm silica gel column to give subfraction 31 determined as megalocarpoidolide H (29.0 mg) (Ndunda et al., 2016). Fractions 103-106 obtained by flash chromatography were pooled and fractioned using a 1cm Sephadex column to give subfractions 24-25 determined as compound **3** (3.1 mg). Fraction 152 obtained by the flash chromatography was determined to be *N-trans-p*-coumaroyl-3',4'-dihydroxyphenylethylamine (7.3 mg) (El Gamaal et al., 1994). Structures of known compounds were confirmed as referenced above.

*3.4 Compound characterisation*

*Megalocarpoidolide I* (**1**): yellow oil; (2.5 mg) $[\alpha]^{22,D}$ -53 (*c* 0.001 $CH_2Cl_2$); IR (NaCl) $\nu_{max}$ ($cm^{-1}$): 2923, 2848, 1764, 1712, 1642, 1454; $^1H$ and $^{13}C$ NMR are given in Tables 1 and 2; HRESIMS *m/z* 357.1696 $[M+H]^-$ ($C_{21}H_{25}O_5$ requires 357.1702).

*Megalocarpoidolide J* (**3**): white solid; (3.1 mg) $[\alpha]^{22,D}$ -63 (*c* 0.002 $CH_2Cl_2$); IR (NaCl) $\nu_{max}$ ($cm^{-1}$): 2947, 1772, 1738,1700, 1644, 1443, 1363, 1265, 1222, 1172, 1102, 1040, 872, 817, 732; $^1H$ and $^{13}C$ NMR are given in Tables 1 and 2; HRESIMS *m/z* 431.1350 $[M-H]^-$ ($C_{22}H_{23}O_9$ requires 431.1348).

*12-Epi-crotozambefuran* (**4**)*:* yellow solid; (1.6 mg) $[\alpha]^{22,D}$ -25 (*c* 0.0015 $CH_2Cl_2$); IR (NaCl) $\nu_{max}$ ($cm^{-1}$): 2919, 2843, 1730, 1699, 1649; $^1H$ and $^{13}C$ NMR are given in Tables 1 and 2; HRESIMS *m/z* 399.1088 $[M-H]^+$ ($C_{22}H_{23}O_7$ requires 399.1085).

*Megalocarpoidolide K* (**5**): white solid; (2.5 mg) $[\alpha]^{22,D}$ -33 (*c* 0.001 $CH_2Cl_2$); IR (NaCl) $\nu_{max}$ ($cm^{-1}$): 2923, 2848, 1759, 1717, 1643, 1454, 1380, 1272, 1250, 1150, 1036, 976, 871; $^1H$ and $^{13}C$ NMR are given in Tables 1 and 2; HRESIMS *m/z* 341.1385 $[M+H]^+$ ($C_{20}H_{21}O_5$ requires 341.1389).

*1-Trans-p-hydroxycoumaroyl–geranylgerani-1-ol* (**6**): yellow solid; (1.7 mg) IR (NaCl) $\nu_{max}$ ($cm^{-1}$): 3429, 2940, 2856, 1712, 1636; $^1H$ and $^{13}C$ NMR are given in Tables 1 and 2; LRMS *m/z* 436.23 $[M]^+$ ($C_{29}H_{40}O_3$ requires 436.30).

**4. Acknowledgements**

BN wishes to thank the University of Surrey, Guildford, UK, for providing laboratory space to undertake this research and also for NMR, IR and MS analysis. BN also thanks the University of Nairobi, International Foundation of Sciences (IFS), International Science Program (ISP-Sweden), Organization for Prevention of Chemical Weapons (OPCW), Natural Products Research Network for East and Central Africa (NAPRECA) for funds to support her PhD studies and supporting the research work that produced these results. The authors thank P.C. Mutiso for collection of the plant material. BOO wishes to thank Tai Solarin University of Education, Nigeria for a study leave during his research visit at the University of Surrey, UK.

**Appendix A. Supplementary data**

Supplementary data relating to this article can be found at xxxxxxx

## References


Aldhaher, A., Langat, M.K., Ndunda, B., Chirchir, D., Midiwo, J.O., Njue, A., Schwikkard, S., Carew, M., Mulholland, D., 2017. Diterpenoids from the roots of *Croton dichogamus* pax. Phytochemistry 144, 1–8.

Beentje H.J. Kenyan Trees, Shrubs and Lianas. Nairobi: Majestic Printing Works Ltd; 1994: 190–192.

El-Gamaal A.A., Takeya K., Itokawa H., Halim A.F., Amer M.M., Saad, H.-E.A., Awad S.A. 1994. Studies on the chemical constituents of *Atraphaxis spinosa* L. var. sinaica Boiss. Natural Medicines 48, 304-306

Fedeli E., Capella P., Cirmelle M., Jacini G., 1966. Isolation of geranyl geraniol from the unsaponifiable fraction of linseed oil. J. Lipid Res. 7, 437–441.

Friis I., Gilbert M.G., 1984. *Croton megalocarpoides* sp. nov. An arborescent Euphorbiaceae from semi-evergreen vegetation in S. Somalia and E. Kenya. Nord J Bot. 4, 327–331.

https://doi.org/10.5281/zenodo.3757401

https://www.dropbox.com/s/hgdeo719uiska1r/Megalocarpoidolide_I.zip?dl=0

https://nmrshiftdb.nmr.uni-koeln.de/

Isyaka M.S., Langat M.K., Mas-Claret E., Mbala B.M., Mvingu B.K., Mulholland D.A., 2020. *Ent*-abietane and *ent*-pimarane diterpenoids from *Croton mubango* (Euphorbiaceae). Phytochemistry 170, 112217

Kapingu M.C., Guillaume D., Mbwambo Z.H., Moshi M.J., Uliso F.C., Mahunnah R.L., 2000. Diterpenoids from the roots of *Croton macrostachys*. Phytochemistry 54, 767–770.

Langat M.K., Crouch N.R., Pohjala L., Tammela P., Smith P.J., Mulholland D.A., 2012. *Ent*-kauren-19-oic acid derivatives from the stem bark of *Croton pseudopulchellus* Pax. Phytochem. Lett. 5, 414–418.

Langat, M.K., Crouch, N.R., Smith, P.J., Mulholland, D.A., 2011. Cembranolides from the leaves of *Croton gratissimus*. J. Nat. Prod. 74, 2349–2355.

Mulholland, D.A., Langat, M.K., Crouch, N.R., Coley, H.M., Mutambi, E.M., Nuzillard, J.M., 2010. Cembranolides from the stem bark of the southern African medicinal plant, *Croton gratissimus* (Euphorbiaceae). Phytochemistry 71, 1381–1386.

National Cancer Institute, 2019. NCI DTP Developmental Therapeutics Program. https://dtp.cancer.gov/discovery_development/nci-60/methodology.htm, Accessed date: 10 January 2020.

Ndunda, B., Langat, M.K., Mulholland, D.A., Eastman, H., Jacob, M.R., Khan, S.I., Walker, L.A., Muhammad, I., Kerubo, L.O., Midiwo, J.O., 2016. New *ent*-clerodane and abietane

diterpenoids from the roots of Kenyan *Croton megalocarpoides* Friis. & M. G. Gilbert. Planta Med. 82, 1079–1086.

Ngadjui, B.T., Abegaz, B.M., Keumedjio, F., Folefoc, G.N., Kapche, G.W.F., 2002. Diterpenoids from the stem bark of *Croton zambesicus*. Phytochemistry 60, 345–349.

Nuzillard, J.-M., Plainchont, B., 2018. Tutorial for the structure elucidation of small molecules by means of the LSD software. Magn Reson Chem. 56, 458–468.

Nuzillard J.-M., Massiot G., 1991. Logic for Structure Determination. Tetrahedron 47, 3655–3664.

Swislocka R., Kowczyk-Sadowy M., Kalinowska M., Lewandowski W., 2012. Spectroscopic (FTIR, IR-Raman, $^{1}H$, $^{13}C$ NMR) and theoretical studies of *p*-coumaric acid and alkali metal *p*-coumarates. Journal of Spectroscopy 27, 546146.

Steinbeck C., Kuhn S., 2004. NMRShiftDB – compound identification and structure elucidation support through a free community-built web database. Phytochemistry 65, 2711–2717.

Shoemaker, R.H., 2006. The NCI60 human tumour cell line anticancer drug screen. Nat. Rev. Cancer 6, 813–823.

Wang G., Zhang H., Liu H., Yue J., 2013. Laevinoids A and B: Two diterpenoids with an unprecedented backbone from *Croton laevigatus*. Org Lett., 15, 4880–4883.

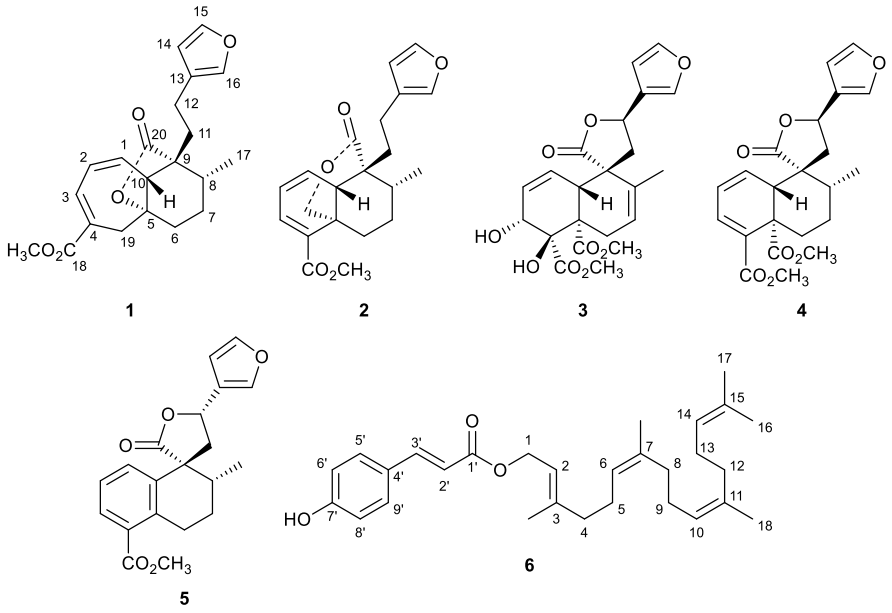


Figure 1. Compounds **1**-**6** from the stem bark of *Croton megalocarpoides*

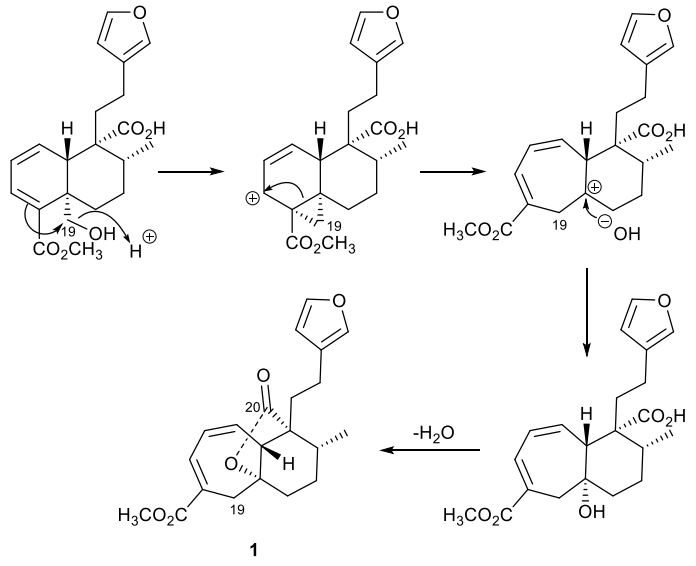


Figure 2. Proposed biosynthesis of compound **1**

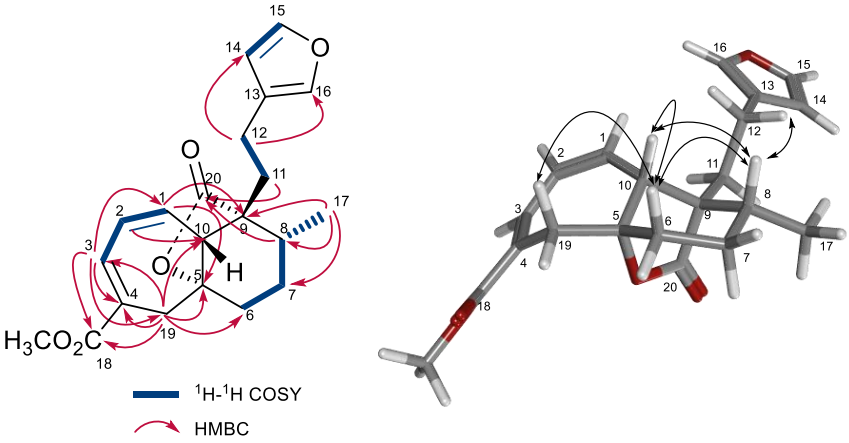


Figure 3. (a) $^1$H-$^1$H COSY and key HMBC (H→C) correlations of compound **1**; (b) Key NOESY correlations of compound **1**

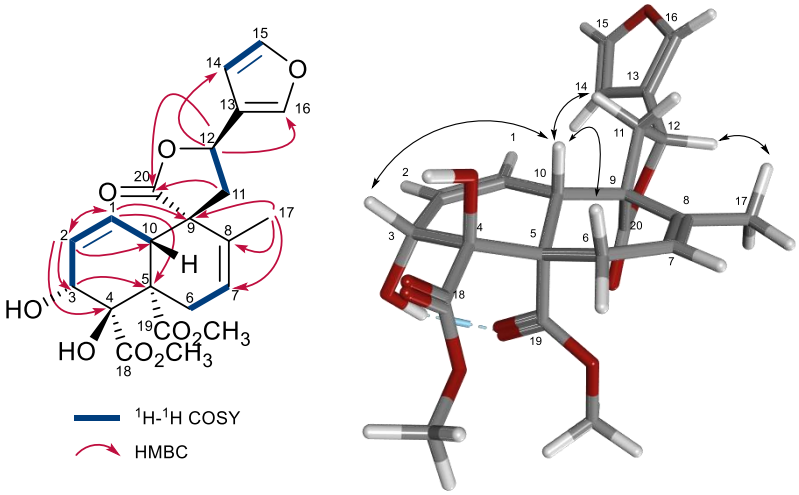


Figure 4. (a) $^1$H-$^1$H COSY and key HMBC (H→C) correlations of compound **3**; (b) Key NOESY correlations of compound **3**

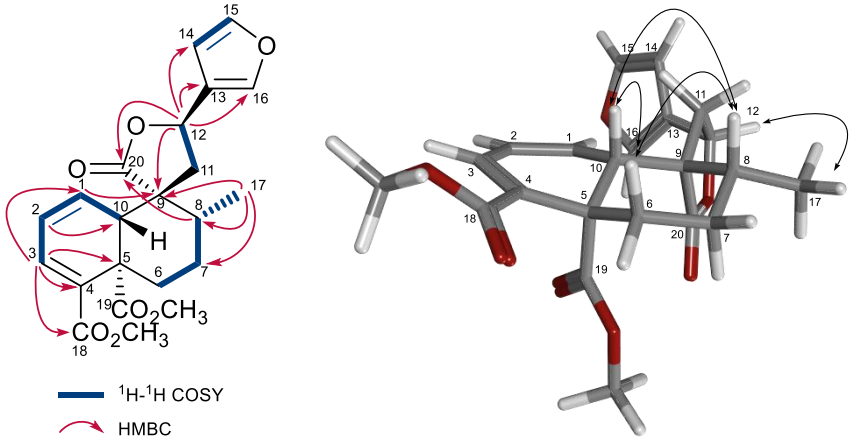

Figure 5. (a) $^{1}$H-$^{1}$H COSY and key HMBC (H→C) correlations of compound **4**; (b) Key NOESY correlations of compound **4**

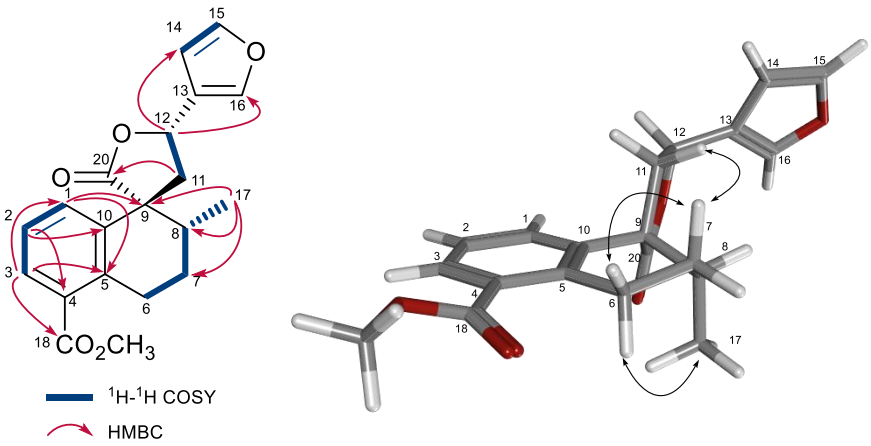


Figure 6. (a) $^{1}$H-$^{1}$H COSY and key HMBC (H→C) correlations of compound **5**; (b) Key NOESY correlations of compound **5**

Table 1: $^{1}$H NMR chemical shift s for compounds **1**, **3** – **6**, δ in ppm and values in parenthesis are *J* in Hz, values with * are overlapped and their *J* values were not determined

| **No** | **1 ($CD_3CN$)** | **1 ($CDCl_3$)** | **3 ($CDCl_3$)** | **4 ($CDCl_3$)** | **5 ($CDCl_3$)** | **6 ($CDCl_3$)** |
|---|---|---|---|---|---|---|
| 1 | 6.31, br d, ($W_{1/2}$ = 1.8) | 6.29, br d, ($W_{1/2}$ = 1.7) | 5.66, dt, (10.4, 1.8) | 5.97, dd, (3.3, 9.4) | 7.26, m* | 4.72, d, (7.1) |
| 2 | 6.28, dd, (1.5, 5.1) | 6.28, br d, ($W_{1/2}$ = 1.7) | 6.08, dt, (10.4, 12.4) | 6.17, ddd, (3.3, 5.2, 9.4) | 7.76, m | 5.42, d, (7.1, 1.0) |
| 3 | 7.13, d, (5.1) | 7.20, br d, ($W_{1/2}$ = 6.1) | 3.93, dd, (12.4, 1.8) | 6.98, dd, (5.2) | 7.27, m* | - |
| 4a | - | - | - | - | - | 1.97, m |
| B | | | | | | 1.97, m |
| 5 | - | - | - | - | - | 2.12, m |
| b | | | | | | 2.06, m |
| 6α | 1.82, m | 1.85, m | 2.42, m | 3.15, dt, (13.2, 3.0) | 3.15, m | 5.10, m |
| β | 1.89, m | 1.92, m | 2.46, t, (2.4) | 1.30, dt, (3.0, 13.2) | 3.15, m | - |
| 7α | 1.30, m | 1.42, m | 5.73, m | 2.15, m | 1.98, m | - |
| β | 1.89, m | 1.89, m | | 1.56, m | 1.98, m | - |
| 8a | 2.00, m | 1.88, m | - | 1.71, m | - | 2.08, m |
| b | | | | | | 1.97, m |
| 9a | - | - | - | - | - | 2.06, m |
| b | | | | | | 2.06, m |
| 10 | 2.73, d (4.0) | 2.50, d (2.5) | 3.06, m | 2.87, dd, (3.3, 3.3) | - | 5.10, m |
| 11α | 1.69, m | 1.83, m | 2.41, m | 2.42, m | 2.54, dd, (10.1, 14.0) | - |
| β | 2.03, m | 2.21, m | 2.72, dd, (7.9, 14.3) | 2.51, m | 2.90, dd, (6.6, 14.0) | |
| 12α | 2.28, m | 2.21, m | 5.50, t, (8.5) | 5.45, t, (8.6) | 5.56, dd, (6.6, 10.1) | 1.97, m |
| β | 2.49, m | 2.45, m | | | | 1.97, m |
| 13a | - | - | - | - | - | 2.06, m |
| b | | | | | | 2.06, m |
| 14 | 6.38, br d, ($W_{1/2}$ = 3.7) | 6.27, d, (1.0) | 6.42, dd, (0.8, 1.7) | 6.43, dd, (0.7, 1.7) | 6.49, dd, (0.9, 1.7) | 5.10, m |
| 15 | 7.42, dd, (1.7, 1.7) | 7.37, dd, (1.7, 1.7) | 7.46, dd, (1.7, 1.7) | 7.47, dd, (1.7, 1.7) | 7.48, dd, (1.7, 1.7) | - |
| 16 | 7.35, br d, ($W_{1/2}$ = 3.5) | 7.25, br s, ($W_{1/2}$ = 3.5) | 7.48, br s, ($W_{1/2}$ = 2.8) | 7.49, br s, ($W_{1/2}$ = 3.6) | 7.54, br s, ($W_{1/2}$ = 3.3) | 1.60, s |
| 17 | 0.90, d, (6.5) | 0.98, d, (6.3) | 1.85, br dd, ($W_{1/2}$ = 4.4) | 1.12, d, (7.0) | 1.16, d, (7.0) | 1.68, s |
| 18 | - | - | - | - | - | 1.59, s |
| 19α | 2.79, d (13.0) | 2.78, d, (13.4) | - | - | - | 1.59, s |
| β | 2.83, d, (13.0) | 2.88, d, (13.4) | | - | | |
| 20 | - | - | - | - | - | 1.74, s |
| 18-$OCH_3$ | 3.74, s | 3.80, s | 3.83, s | 3.72, s | 3.88, s | |
| 19-$OCH_3$ | | | 3.74, s | 3.60, s | - | |
| 5-$OCH_3$ | | | | | | |
| 1’ | | | | | | - |
| 2’ | | | | | | 7.63, d, (16.0) |
| 3’ | | | | | | 6.31, d, (16.0) |
| 4’ | | | | | | - |
| 5’/9’ | | | | | | 7.42, d, (8.6) |
| 6’/8’ | | | | | | 6.84, d, (8.6) |
| 7’ | | | | | | |

Table 1: $^{13}C$ NMR chemical shift s for compounds **1**, **3 – 6**, δ in ppm

| **No** | **1 ($CD_3CN$)** | **1 ($CDCl_3$)** | **3 ($CDCl_3$)** | **4 ($CDCl_3$)** | **5 ($CDCl_3$)** | **6 ($CDCl_3$)** |
|---|---|---|---|---|---|---|
| 1 | 137.1 | 135.0 | 123.6 | 132.8 | 133.5 | 61.0 |
| 2 | 129.5 | 129.8 | 131.1 | 125.7 | 129.7 | 118.6 |
| 3 | 137.4 | 137.6 | 73.6 | 135.6 | 126.7 | 142.6 |
| 4 | 131.6 | 131.4 | 80.3 | 136.4 | 138.5 | 39.8 |
| 5 | 95.2 | 97.0 | 46.6 | 47.1 | 138.1 | 26.5 |
| 6 | 36.6 | 34.9 | 27.8 | 32.0 | 24.3 | 123.8 |
| 7 | 30.3 | 29.1 | 126.2 | 28.3 | 26.6 | 135.7 |
| 8 | 35.6 | 35.3 | 129.3 | 43.0 | 33.3 | 39.9 |
| 9 | 54.8 | 53.4 | 52.1 | 50.8 | 54.1 | 26.7 |
| 10 | 56.3 | 55.0 | 43.3 | 50.1 | 130.4 | 124.6 |
| 11 | 29.9 | 27.7 | 42.0 | 41.9 | 48.7 | 135.2 |
| 12 | 19.8 | 19.1 | 72.3 | 72.1 | 71.6 | 39.9 |
| 13 | 126.0 | 124.2 | 125.1 | 125.6 | 124.5 | 26.4 |
| 14 | 112.2 | 110.8 | 108.2 | 108.4 | 108.5 | 124.4 |
| 15 | 144.5 | 143.3 | 144.5 | 144.5 | 144.4 | 131.5 |
| 16 | 140.3 | 138.8 | 139.8 | 139.8 | 140.2 | 25.9 |
| 17 | 16.8 | 16.3 | 20.0 | 17.4 | 16.2 | 17.8 |
| 18 | 168.5 | 167.3 | 173.1 | 166.7 | 168.3 | 16.2 |
| 19 | 36.6 | 35.9 | 174.1 | 171.4 | 178.4 | 16.3 |
| 20 | 178.5 | 177.4 | 176.4 | 176.3 | - | 16.8 |
| 18-$OCH_3$ | 53.1 | 52.5 | 53.3 | 51.9 | 52.2 | - |
| 19-$OCH_3$ | | | 53.1 | 51.7 | - | - |
| 5-$OCH_3$ | | | | | | - |
| 1’ | | | | | | 167.8 |
| 2’ | | | | | | 144.6 |
| 3’ | | | | | | 115.9 |
| 4’ | | | | | | 127.5 |
| 5’/9’ | | | | | | 130.2 |
| 6’/8’ | | | | | | 116.1 |
| 7’ | | | | | | 157.9 |