Dammarane-type triterpenoids with anti-cancer activity from the leaves of *Cleome gynandra*

Buntubonke Mzondo[a,b,c], Nomusa Dlamini[b], Frederick P. Malan[c], Philip Labuschagne[a], Venugopal R. Bovilla[d], SubbaRao V. Madhunapantula[d], Vinesh Maharaj[c]

[a] Centre for Nanostructured and Advanced Materials, Council for Scientific and Industrial Research, Pretoria 0001, South Africa
[b] Advanced Agriculture and Food, Council for Scientific and Industrial Research, Pretoria 0001, South Africa
[c] Department of Chemistry, Faculty of Natural and Agricultural Science, University of Pretoria, Private Bag x20, Pretoria 0028, South Africa
[d] Centre of Excellence in Molecular Biology and Regenerative Medicine (CEMR) Laboratory, Department of Biochemistry, JSS Medical College, JSS Academy of Higher Education & Research, Mysore - 570015, Karnataka, India

Main author: Buntubonke Mzondo
Corresponding author: Vinesh Maharaj

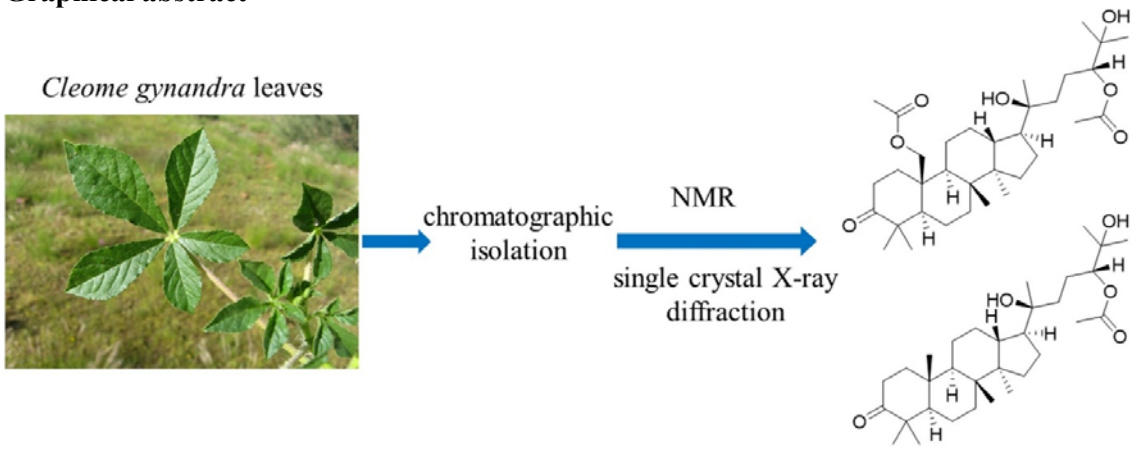
Highlights

- Two new dammarane-type triterpernoids, cleogynones A and B were isolated from *Cleome gynandra.*
- The structures were elucidated by NMR and single crystal X-ray crystallography.
- Cleogynone A and B showed moderate cytotoxicity against breast and colorectal cancer.

ABSTRACT

Three dammarane-type triterpernoids including two new ones, cleogynones A and B (**1** and **2**), were isolated from the leaves of *Cleome gynandra*. The structures of the new triterpenoids were elucidated by spectroscopic data analysis and confirmed by single crystal X-ray crystallography. All three compounds showed moderate cytotoxicity against breast cancer (MDA-MB-468), cleogynone B (**2**) and compound (**3**) further showed cytotoxicity against colorectal cancer (HCT-116 & HCT-15). Cleogynone B was also moderately active against lung cancer (A549).

Graphical abstract



Key words: *Cleome gynandra;* Dammarane triterpernoids; Lung cancer; Breast cancer; Colorectal cancer

1. Introduction

Cleome gynandra L. belongs to the family *Capparaceae* and is reported to have originated in tropical Africa. It is amongst the most nutritious leafy vegetables (Chweya and Mnzava, 1997; Mishra et al., 2011). In some countries, it plays a pivotal role in household food security during drought, as it prevails as a sole leafy vegetable available during the period of vegetable shortages or the relish-gap (Chweya and Mnzava, 1997; Ekpong, 2009). Some traditional uses involve treating medical conditions such as epilepsy, irritable bowel syndrome, protozoal and worm infections, while it is also used to ease labour during childbirth in some communities (Kamatenesi-Mugisha and Oryem-Origa, 2007; Mishra et al., 2011). *C. gynandra* has been reported to have immunomodulatory, antioxidant, analgesic, and lysosomal stability actions, amongst other medicinal properties (Bala et al., 2012; Mishra et al., 2011; Narendhirakannan et al., 2005). The leaves specifically possess medicinal properties for the treatment of hepatic disorders, inflammation, and cancer (Bala et al., 2010; Mishra et al., 2011; Narendhirakannan et al., 2007).

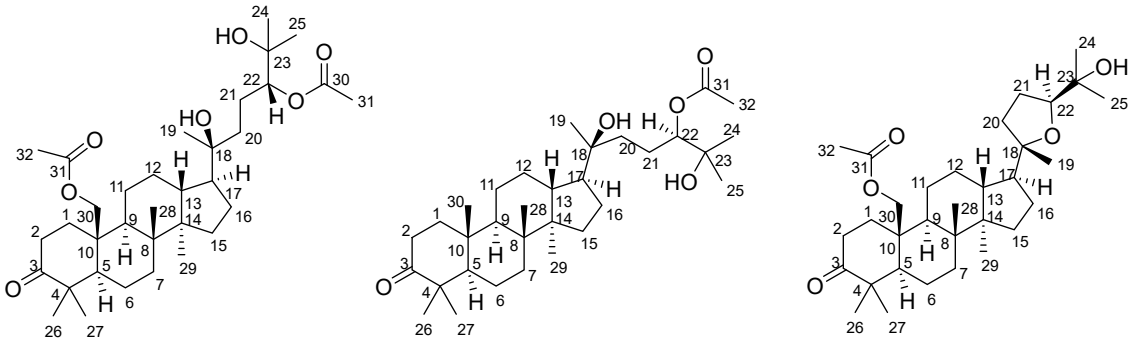
Cancer is the second leading cause of death in the world at large (Gupta et al., 2010; J. Ferlay et al., 2012). The World Health Organization (WHO) stated that cancer accounted for 9.6 million deaths in 2018. Lung-, breast-, and colorectal cancers were the most prevalent, accounting for 2.09, 2.09 and 1.80 million cases, respectively (Ferlay et al., 2012). Towards identifying compounds responsible for biological activity from *C. gynandra*, anti-cancer was one of the biological properties reported for *C. gynandra*, and was hence chosen for this study. In this study, we initially evaluated extracts from the leaves of *C. gynandra* for their *in vitro* anti-cancer activity followed by the isolation and structure elucidation of the compounds responsible for the anti-cancer activity. Cancer cell lines selected for the study were the lung cancer (A549), colorectal cancer (HC-116 & HCT-15) and breast cancer (MDA-MB-468). We herein, therefore, report the isolation, structure elucidation and characterization of compounds with anti-cancer properties from *C. gynandra* leaves.

2. Results and discussion

Sequential extraction of dried and ground leaves of *C. gynandra*, using *n*-hexane, dichloromethane, ethyl acetate and methanol, led to four extracts, which were initially evaluated for their lung cancer cytotoxicity in the A549 cell line, Table 1 shows the anti-cancer screening results for the extracts. This result led to the fractionation of *n*-hexane extract as it showed the best activity (>80% inhibition in the concentration range of 0.13 – 1.0 mg/mL, over 48-hour incubation and >75 % over 24-hour incubation). The fractionation of *n*-hexane extract *via* bioassay guided fractionation using column chromatography eventually led to the isolation of two novel triterpenoids (compound **1** and **2**) and one known compound (**3**) (Fig. 1).

Table 1
^{1}H (400 MHz) and ^{13}C NMR (100 MHz) spectroscopic data for compounds **1** and **2** in $CDCl_3$

Position	**1**		**2**	
	δ_H (mult., *J* in Hz)	δ_C	δ_H (mult., *J* in Hz)	δ_C
1	2.08 (m) 1.56 (m)	34.5	1.93 (m) 1.46 (m)	39.9
2	2.50 (m) 2.28 (m)	34.4	2.48 (m)	34.1
3		216.9		218.2
4		45.9		47.4
5	1.58 (m)	52.6	1.38 (m)	55.3
6	1.47 (m)	19.7	1.52 (m)	19.7
7	1.51 (m) 1.29 (m)	34.1	1.56 (m) 1.33 (m)	34.5
8		40.1		40.3
9	1.57 (m)	51.6	1.44 (m)	50.0
10		38.7		36.81
11	1.53 (m) 1.29 (m)	22.9	1.74 (m) 1.49 (m)	24.8
12	1.66 (m) 1.45 (m)	24.7	1.80 (m) 1.29 (m)	27.6
13	1.54 (m)	42.7	1.79 (m) 1.61 (m)	23.7
14		50.3	1.63 (m)	42.5
15	1.37 (m) 1.05 (m)	31.5		50.4
16	1.72 (m)	23.7	1.46 (m) 1.10 (m)	31.2
17	1.66 (m)	49.8	1.73 (m)	49.9
18		75.1		75.2
19	1.06 (s)	25.0	1.14 (s)	25.0
20	1.37 (m)	37.1	1.46 (m)	37.0
21	1.47 (m)	19.7	1.51 (m) 1.29 (m)	22.0
22	4.74 (dd, 10.3, 2.4)	80.5	4.82 (dd, 10.3, 2.4)	80.5
23		72.5		72.5
24	1.14 (s)	26.7	1.22 (s)	26.6
25	1.14 (s)	25.2	1.22 (s)	25.1
26	1.02 (s)	29.4	1.09 (s)	26.7
27	0.90 (s)	19.5	1.05 (s)	21.1
28	0.88 (s)	15.3	1,00 (s)	15.1
29	0.84 (s)	16.7	0.89 (s)	16.4
30	4.18 (d, 11.7) 4.00 (d, 11.7)	64.5	0.95 (s)	16.0
31	-	171.3		171.4
32	1.90 (s)	21.0	2.12 (s)	21.0
33	-	171.2		
34	2.05 (s)	21.1		



1 **2** **3**

Fig. 1. Structures of compounds **1** - **3**

2.1. Structure elucidation

Compound **1** (Fig. 1)**,** crystallised on drying in a centrifugal evaporator to yield colourless single crystalline material. The IR spectrum of compound **1** revealed absorption at 3450, 1710, and 1270 cm^{-1}, which were attributed to hydroxyl, ketone and ester groups, respectively. Compound **1** was further assigned the molecular formula $C_{34}H_{56}O_7$, from its HRESIMS *m/z* 559.3996, $[M - H_2O + H]^+$ (calculated for 559.3999), ^1H and ^{13}C NMR (including DEPT) and single crystal X-ray crystallographic data (SCXRD) with seven degrees of unsaturation. The ^1H NMR of **1** (Table 2) revealed nine singlet methyl groups at δ_H 2.05 (3H, s), 1.90 (3H, s), 1.14 (6H, s), 1.06 (3H, s), 1.02 (3H, s), 0.90 (3H, s), 0.88 (3H, s) and 0.84 (3H, s), the presence of one oxygenated methine at δ_H 4.74 (1H, dd, J = 10.3, 2.4 Hz) and one oxygenated methylene with protons at δ_H 4.18 (1H, d, J = 11.7 Hz) and 4.00 (1H, d, J = 11.7 Hz). In addition, two proton resonances at δ_H 2.50 (1H, m) and 2.28 (1H, m) from one methylene were seen. Several protons for methylenes and methines overlap in the range between δ_H 1.18 and 1.78, one of the methylenes had one proton at δ_H 2.08 (1H, m) and the other embedded within the overlapping range. Two protons, at δ_H 2.01 (2H, s) further revealed the presence of two hydroxyl groups.

Inspection of the ^{13}C NMR spectrum together with the DEPT spectrum suggested that compound **1** was a triterpenoid derivative with 34 carbons (Table 2), consisting of nine methyl carbons, eleven sp^3 methylenes (including one oxymethylene at δ_C 64.5), five sp^3 methines (including one oxymethine at δ_C 80.5), six sp^3 quaternary carbons (including two oxygen carrying at δ_C 75.1 and 72.5), and three sp^2 quaternary carbons (carbonyls at δ_C 171.2, 171.3 and 216.9). Apart from three carbonyl groups, the remaining elements of unsaturation were attributed to four rings.

Table 2: Comparative assessment of cytotoxic potential of solvent extracts of *Cleome gynandra* against A549 lung cancer cell line

Name of the extract	Concentration (mg/mL)	Growth inhibition (%) at 24 h (Mean±SE)	Growth inhibition (%) at 48 h (Mean±SE)
Respective vehicle control #		9.4±1.96	-2.46±2.15
Hexane extract	0.125	75.05±0.20	84.87±0.05
	0.25	78.61±0.73	87.18±0.13
	0.50	78.10±0.23	87.27±0.18
	1.00	77.53±0.77	87.08±0.18
Dichloromethane extract	0.125	-5.55±7.22	16.01±5.78
	0.25	12.39±2.82	46.07±2.80
	0.50	22.45±0.49	64.64±1.68
	1.00	25.70±2.52	66.44±1.91
Ethyl acetate extract	0.125	16.3 8±2.71	38.29±0.86
	0.25	34.71±1.86	63.46±0.64
	0.50	42.04±1.39	75.24±1.27
	1.00	56.42±1.25	74.35±0.79
Methanol extract	0.125	16.69±3.84	5.55±0.60
	0.25	10.95±3.73	-1.03±1.75
	0.50	9.92±1.20	2.36±1.26
	1.00	12.98±0.78	14.86±1.21

The final concentration of respective vehicle control in the treatment media is 1%

The ^{1}H-^{1}H COSY correlation (Fig. 2) between δ_H 1.56, 2.08 (H-1)/δ_H 2.28, 2.50 (H-2) and the HMBC correlations (Fig. 2) from H-1 (δ_H 2.08, 1.56), H-2 (δ_H 2.50, 2.28), H-26 (δ_H 1.02) and H-27 (δ_H 0.90) to C-3 (δ_C 216.9) of compound **1** confirmed the existence of a ketone at C-3 and that C-4 is substituted by two methyls. Furthermore, the correlations from H-30 (δ_H 4.00, 4.18) to C-1 (δ_C 34.5), C-5 (δ_C 52.6), C-9 (δ_C 51.6) and C-31 (δ_C 171.3), and from CH_3-32 (δ_H 1.90) to C-31 (δ_C 171.3) demonstrated the substitution of C-10 with a methylene further linked to an acetyl group. A second acetyl group is located at C-22 as demonstrated by the HMBC correlations between H-22 (δ_H 4.74) and C-33 (δ_C 171.2), H-34 (δ_H 2.05) and C-33 (δ_C 171.2). The aforementioned information helped establish the vicinity of the three carbonyl groups contained in the structure.

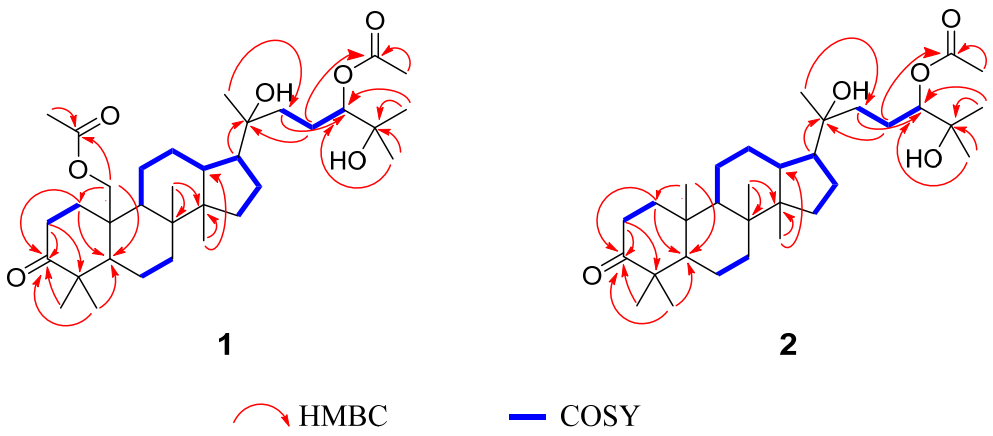


Fig. 2. Key HMBC and ^{1}H-^{1}H COSY correlations of compounds **1** and **2**.

Two hydroxyl groups as part of the structure were further located. The HMBC correlations from H-24 (δ_H 1.14), H-25 (δ_H 1.14) to C-22 (δ_C 80.5) and to C-23 (δ_C 72.5) illustrated that C-23 carries two methyl groups and a possible hydroxyl group which was later confirmed by single crystal X-ray diffraction (SCXRD) analysis. The HMBC correlations from CH_3-19 (δ_H 1.06) to C-17 (δ_C 49.8), C-18 (δ_C 75.1), C-20 (δ_C 37.1) demonstrated that C-18 neighbours a methylene (C-20) on one end, and attaches a hydroxyl group and a methyl group. While the HMBC correlation from H-17 (δ_H 1.66) to C-18 (δ_C 75.1) demonstrated that on the other end C-18 is connected to the D ring at C-17. C-20 shares an HMBC correlation with H-22 while H-22 shares a ^{1}H-^{1}H COSY correlation with H-21, completing the linkage of the structural moieties.

The location of the further two methyl groups is described by the HMBC correlations from CH_3-28 (δ_H 0.88) to C-7 (δ_C 34.1), C-14 (δ_C 50.3), C-8 (δ_C 40.1) and that from CH_3-29 (δ_H 0.84) to C-14 (δ_C 50.3), C-15 (δ_C 31.5), these methyl groups are attached at C8 and C14, respectively.

The NOESY spectrum of compound **1** showed correlation between H-22, H-24 and H-25 as shown in Fig. 3 indicating that these protons are on the same side, further demonstrating that the hydroxyl group on C-23 is at the same side as the acetate group that attaches at position C-22. Furthermore, the correlation between H-30 (δ_H 11.7) and H-28 proved that the methyl group C-28 is in the same face as the oxymethylene at C-30. Compound **1** crystallized in the chiral space group $P2_1$, revealing nine chiral centres present (seven as part of the triterpenoid skeletal structure, and two as part of the heptyl-acetate based-moiety). The final refinement of the CμKα data in SCXRD analysis, based on the seven oxygen atoms in the molecule, resulted in a Flack parameter (Flack, 1983; Flack and Bernardinelli, 2000) of 0.04(10), which not only indicated the relative configuration of two methyl groups (Fig. 4), but also unambiguously determined the absolute stereochemistry of compound **1** to be 5*R*, 8*R*, 9*S*, 10*S*, 13*R*, 14*R*, 17*S*, 18*S*, and 22*R*. All other bond lengths and angles fall within the expected ranges for the respective functional groups.

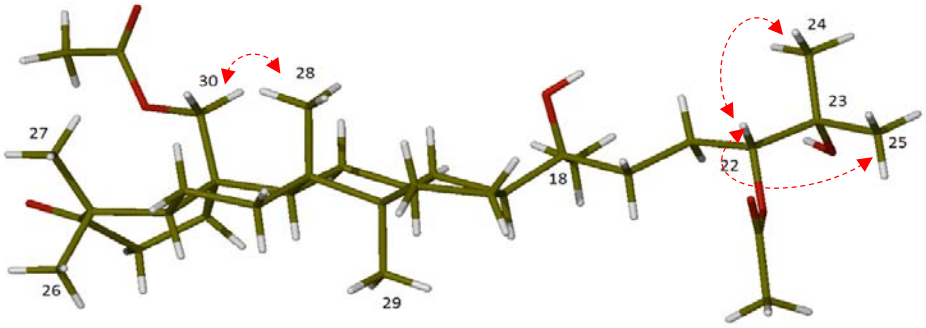


Fig. 3. Key NOESY ( ) correlations of compound **1**.

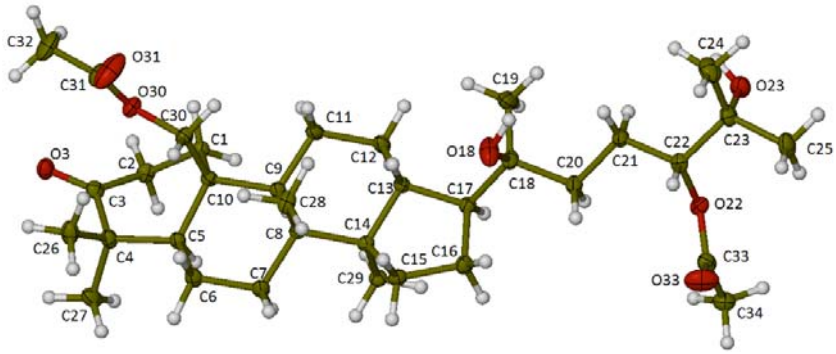


Fig. 4. Pictogram of compound **1** structure (CμKα) (ellipsoids shown at the 50 % probability level). Green represents C-atoms, red represents O-atoms while white represents H-atoms.

Compound **2** was obtained as a white powder and assigned the molecular formula $C_{32}H_{54}O_5$ from its HRESIMS *m/z* 501.3971 $[M - H_2O + H]^+$ (calculated for 501.3944), NMR spectra and single crystal X-ray diffraction, with six degrees of unsaturation.

^{1}H NMR (Table 2) revealed signals for nine singlet methyls (δ_H 0.89, 0.95, 1.00, 1.05, 1.09, 1.14, 1.22, 1.22, 2.12), one oxygenated methine at δ_H 4.82 (1H, dd, J = 10.3, 2.4 Hz) and a proton resonance at δ_H 2.48 (2H, m) for a methylene group. Several proton resonances for methylenes and methyls are embedded in the overlapping range between δ_H 1.25 – 1.99.

The ^{13}C NMR and DEPT spectra (Table 2) suggested that compound **2** was also a triterpenoid derivative, with a total of 32 carbons, assigned to nine methyls, ten methylenes, five methines (including one oxygenated δ_C 4.82), and eight quaternary carbons (including two carbonyl δ_C 171.4 and 218.2). With the exception of the two carbonyls, the remaining elements of unsaturation were attributed to a tetracyclic system.

Similar to compound **1**, the ^{1}H-^{1}H COSY correlation (Fig. 2) between δ_H 1.46, 1.93 (H-1)/δ_H 2.48 (H-2) and the HMBC correlations (Fig. 2) from H-1 (δ_H 1.46, 1.93), H-2 (δ_H 2.48), CH_3-26 (δ_H 1.02) and CH_3-27 (δ_H 0.90) to C-3 (δ_C 216.9) confirmed the presence of a ketone at C-3, and that C-4 is substituted by two methyl groups. As in compound **1** an acetyl group was located at C-22, demonstrated by the HMBC correlations between H-22 (δ_H 4.82) and C-31 (δ_C 171.4), H-32 (δ_H 2.12) and C-31 (δ_C 171.4). Furthermore, the correlations from CH_3-30 (δ_H 0.95) to C-1 (δ_C 39.9), C-5 (δ_C 55.3), C-10 (δ_C 36.8) demonstrated the substitution of C-10 with a methyl, further illustrating the only difference between compounds **1** and **2**, was that while **1**

has an acetate group linked by a methylene to C-10, compound **2** has a methyl group attached at position C-10. The remaining HMBC correlation between compounds **1** and **2** were identical or similar.

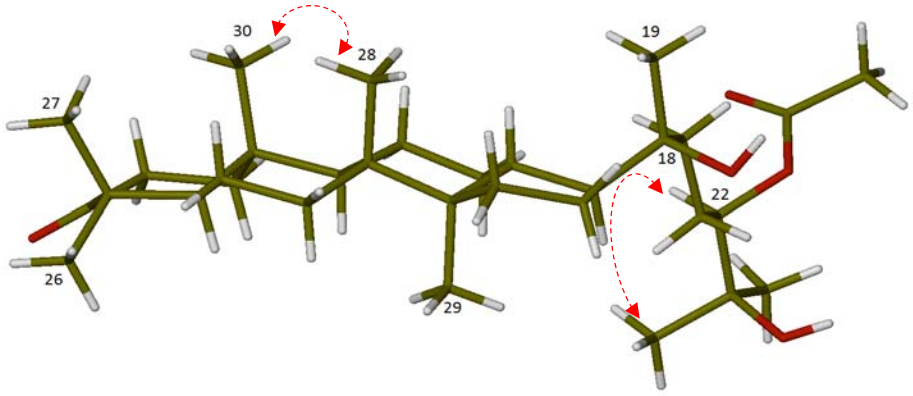


Fig. 5. Key NOESY ( ) correlations of compound **2**.

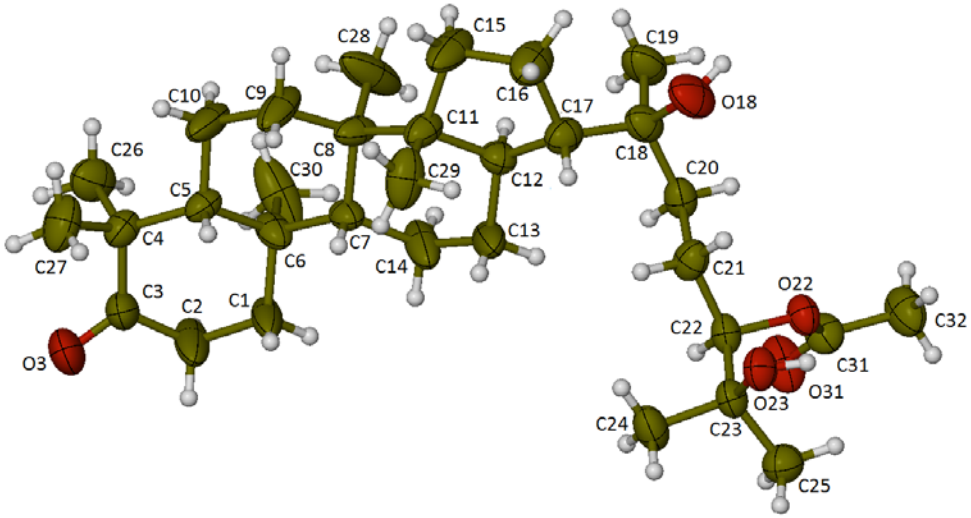


Fig. 6. Pictogram of compound **2** structure (CμKα) (ellipsoids shown at the 50 % probability level). Green represents C-atoms, red represents O-atoms while white represents H-atoms.

The NOESY spectrum of compound **2** (as illustrated in Fig. 5) showed correlation between CH_3-30 and CH_3-28 proving that the methyl group C-28 is in the same side as the methyl at C-30. Compound **2** crystallized in the chiral space group *P*1, revealing nine chiral centres present as before with compound **1**. The final refinement of the CμKα data in SCXRD analysis, based on the five oxygen atoms in the molecule, resulted in a Flack parameter (Flack, 1983; Flack and Bernardinelli, 2000) of 0.06(6), which indicated the relative configuration of the three methyl groups on the infused ring backbone (Fig. 6), while also assisting in unambiguously determining the absolute stereochemistry of compound **2** to be 5*R*, 8*R*, 9*R*, 10*R*, 13*R*, 14*R*, 17*S*, 18*S*, and 22*R*. The equivalent of compound **2** was semi-synthetically prepared and the absolute configuration was the same as compound **2**, except for C-22 which was the *S* isomer as compared to *R* in compound **2** (Mai et.al, 2016). All bond lengths and angles observed in

compound **2** compare well with the corresponding bond lengths and angles observed in compound **1**, and fall within the expected ranges for the respective functional groups.

The structure of compound **3** was confirmed by comparison of the NMR data to those described in literature by Das *et al*., who isolated cleogynol for the first time from *C. gynandra* and acetylation of which resulted in compound **3** without any given name. (Das et al., 1999)

2.2. Biological activity

Compounds **1**-**3** possessed varying anti-cancer activity against lung cancer (A549), breast cancer (MDA-MB-468) and colorectal carcinoma (HCT-116 and HCT-15) cell lines. Tables 3-6 demonstrate the anti-cancer activity of the pure compounds tested against the different cell lines. Compound **2** showed the highest lung cancer inhibition activity, 51 % inhibition over 24-hour treatment, comparable to diallyl disulfide (DADS) 48% inhibition tested at 146.2 µg/mL (Table 3). The breast cancer (MDA-MB-468) cytotoxicity results for the compounds are shown in Table 4, where all three compounds showed concentration-dependent inhibition above 50 % at 25 µg/mL over 48-hour treatment, comparable to DADS at 146.2 µg/mL. Compound **2** and compound **3** showed approx. 80% inhibition activity against HCT-116 over 24-hour treatment at the highest concentration (Table 5). Against HCT-15, cleogynone B (**2**) and compound **3** showed 80% inhibition at 25 µg/mL concentration over both the 24 and 48-hour treatment, while compound **1** only exhibited the 80 % inhibition activity over a 48-hour treatment (Table 6). It is evident that the best overall concentration-dependent activity is demonstrated against breast cancer MDA-MB-468 over 48-hour treatment as inhibition decreases slowly with decreasing concentration over the concentration range 0.78 – 25 µg/mL. The highest activity is observed against colorectal cancer cell line HCT-15, reaching 80% inhibition at 25 µg/mL concentration for all compounds.

Table 3: Comparative assessment of cytotoxic potential of *Cleome gynandra* Compound-2 against A549 lung cancer cell line

Name of the compound	Concentration (µg/mL)	Growth inhibition (%) at 24 h (Mean±SE)	Growth inhibition (%) at 48 h (Mean±SE)
Vehicle DMSO control #		9.89±4.76	4.06±5.58
Positive control DADS	146.2	46.55±2.12	53.23±2.40
Compound 2	0.39	-4.37±3.08	-0.87±4.03
	0.78	-5.78±2.50	0.42±4.38
	1.56	-6.60±1.91	0.44±4.83
	3.12	-5.46±2.21	0.72±4.43
	6.25	-1.8±3.80	1.33±4.55
	12.50	-5.19±3.13	0.53±4.20
	25.00	51.38±0.39	16.14±1.72

The final concentration of vehicle DMSO control in the treatment media is 1%

Table 4: Comparative assessment of cytotoxic potential of *Cleome gynandra* compounds 1, 2, 3 against MDA-MB-468 lung cancer cell line

Name of the compound	Concentration (µg /mL)	Growth inhibition (%) at 24 h (Mean±SE)	Growth inhibition (%) at 48 h (Mean±SE)
Vehicle DMSO control #	0.00	-0.37±1.22	3.22±2.62
Positive control DADS	146.2	60.99±2.06	59.13±0.04
Compound 1	0.78	9.33±7.54	11.14±4.23
	1.56	3.87±6.41	20.88±6.26
	3.12	0.32±5.29	21.6±1.97
	6.25	1.42±5.68	23.73±11.92
	12.50	3.74±6.67	28.12±8.34
	25.00	44.48±1.42	52.55±5.42
Compound 2	0.78	-0.99±4.58	-8.10±6.11
	1.56	1.72±2.89	10.7±3.71
	3.12	0.22±3.14	18.65±4.45
	6.25	2.41±1.60	20.3±3.06
	12.50	6.12±3.51	34.95±6.61
	25.00	43.82±2.38	51.28±1.99
Compound 3	0.78	5.71±1.45	-4.67±5.00
	1.56	7.03±1.80	8.33±2.15
	3.12	8.02±2.29	12.13±1.71
	6.25	6.84±2.63	14.89±0.67
	12.50	10.99±2.69	22.18±2.13
	25.00	48.32±2.19	66.58±1.30

The final concentration of vehicle DMSO control in the treatment media is 1%

Table 5: Comparative assessment of cytotoxic potential of *Cleome gynandra* compounds 2 and 3 against HCT116 lung cancer cell line

Name of the compound	Concentration (µg /mL)	Growth inhibition (%) at 24 h (Mean±SE)	Growth inhibition (%) at 48 h (Mean±SE)
Vehicle DMSO control #	0.00	18.91±3.18	21.87±1.75
Positive controls	30.11 (Cisplatin)	35.8±2.98	26.76±2.05
	146.2 (DADS)	89.17±0.90	79.84±3.63
Compound 2	0.39	23.6±0.89	-1.52±5.44
	0.78	20.33±0.96	11.18±1.67
	1.56	23.74±2.36	16.25±2.84
	3.12	25.08±3.26	18.07±3.06
	6.25	30.27±3.15	22.93±2.46
	12.50	30.58±1.51	22.94±3.68
	25.00	89.34±5.46	55.79±1.23
Compound 3	0.39	18.25±1.71	0.85±7.16
	0.78	15.26±1.81	12.87±4.43
	1.56	17.15±2.17	16.55±2.93
	3.12	22.34±1.80	21.68±3.13
	6.25	27.68±1.39	24.78±1.55
	12.50	28.02±0.48	24.44±1.76
	25.00	87.76±1.22	60.96±4.95

The final concentration of vehicle DMSO control in the treatment media is 1%

Table 6: Comparative assessment of cytotoxic potential of *Cleome gynandra* compounds 1, 2, 3 against HCT15 lung cancer cell line

Name of the compound	Concentration (µg /mL)	Growth inhibition (%) at 24 h (Mean±SE)	Growth inhibition (%) at 48 h (Mean±SE)
Vehicle DMSO control #		18.62±3.57	21.66±1.16
Positive controls	30.11 (Cisplatin)	45.25±1.95	75.92±1.60
	146.2 (DADS)	63.71±2.88	79.98±1.12
Compound 1	0.39	-7.06±1.71	4.01±1.15
	0.78	-3.51±1.54	-3.58±0.88
	1.56	-4.72±3.15	-3.28±0.23
	3.12	-1.88±8.68	-3.84±0.56
	6.25	8.92±4.87	-2.88±1.00
	12.50	16.16±5.63	7.01±2.43
	25.00	51±6.59	82.97±0.56
Compound 2	0.39	16.85±3.06	1.6±1.52
	0.78	9.98±4.18	-4.38±0.87
	1.56	7.11±5.90	-5.04±0.60
	3.12	8.27±5.85	-4.06±0.38
	6.25	15.1±4.42	-2.74±0.48
	12.50	36.65±4.05	8.82±2.00
	25.00	72.27±3.91	81.74±0.34
Compound 3	0.39	20.46±3.28	-0.35±0.60
	0.78	2.83±6.86	-5.01±1.30
	1.56	1.43±6.82	-5.32±0.31
	3.12	-0.17±8.74	-4.27±0.66
	6.25	7.64±7.70	-4.17±1.07
	12.50	30.44±6.39	7.82±2.75
	25.00	64.85±9.03	83.67±3.16

The final concentration of vehicle DMSO control in the treatment media is 1%

3. Experimental

3.1. General experimental procedures

NMR spectra were recorded on a Bruker Avance III-400 instrument (Bruker, Karlsruhe, Germany). ESI-MS data were obtained *via* ultra-performance liquid chromatography with an Acquity UPLC system, coupled to a Waters Synapt G2 quadrupole time-of-flight (QTOF) mass spectrometer using an electrospray ionization technique operating in positive mode (Milford, Massachusetts, USA). IR spectra were recorded by a Perkin Elmer spectrum 100 FT-IR spectrophotometer (Waltham, Massachusetts, USA). Optical rotations were recorded on a

Perkin Elmer Model 341 polarimeter. Single crystal X-ray data were collected using a Bruker D8 Venture diffractometer using monochromated CμKα (k = 1.54178 Å), a Photon 100 detector (Billerica, Massachusetts, USA), and APEX III control software (Bruker AXS Inc.). Data reduction was performed using SAINT+ (Bruker AXS Inc.), and the intensities were corrected for absorption using SADABS (Bruker AXS Inc.). All structures were solved by direct methods using a SHELXS algorithm (Sheldrick, 2015a) and refined using the SHELXL program (Sheldrick, 2015b). All H atoms were placed in geometrically idealised positions and constrained to ride on their parent atoms. The X-ray crystallographic coordinates for all structures have been deposited at the Cambridge Crystallographic Data Centre (CCDC), with deposition numbers CCDC 2016627 (**1**) and 2016628 (**2**). Silica gel 60 (0.063-0.2 mm; Macherey-Nagel GmbH & Co. KG, Düren, Germany) and Silica gel-60 (0.040-0.063 mm; Merck KGaA, Darmstadt, Germany), were used for open column chromatography. Silica gel GF_{254} plates (Merck KGeA, Darmstadt, Germany) were used for thin layer chromatography (TLC) detection, while preparative thin layer chromatography involved the use of TLC Silica gel 60 glass plates (Merck KGaA, Darmstadt, Germany).

3.2. Plant material

C. gynandra was harvested from the Agricultural Research Council – Vegetable and Ornamental Plants (ARC-VOP) (Roodeplaat, Pretoria, South Africa) in November 2016, and identified by Dr Willem Jansen Van Rensburg a voucher specimen (TOT8420) has been deposited at the ARC-VOP laboratory.

3.3. Isolation of compounds

Dried leaves (500 g) of *C. gynandra* were ground to fine particles (~ 40 μm) and soaked in *n*-hexane (5 L × 1h, 5L × 18 and 5L ×1) at room temperature with constant stirring. The *n*-hexane extracts were evaporated to dryness at 40 °C under vacuum by a rotary evaporator to obtain 13 g extract residue. The insoluble residue was further soaked in dichloromethane (5L × 1h, 5L × 18h and 5L × 1h) at room temperature with stirring. The dichloromethane extracts were evaporated to dryness at 40 °C under vacuum by a rotary evaporator to obtain 18 g residue. The resulting insoluble plant residue was further soaked in ethyl acetate (5 L × 1h, 5L × 18h and 5L ×1h) at room temperature with constant stirring. The ethyl acetate extracts were evaporated to dryness at 50 °C under vacuum by a rotary evaporator to obtain 19 g residue. Finally, the insoluble plant residue was further soaked in methanol (5L × 1h, 5L × 18h and 5L × 1h) at room temperature with stirring. The methanol extracts were evaporated to dryness at 60 °C under vacuum by a rotary evaporator to obtain 16 g. The lung cancer cytotoxicity results of the extracts led to further fractionation of only the *n*-hexane extract as this was most active. Of the 13 g *n*-hexane extract yield, 10 g was divided to sixteen fractions (1 – 16) by silica gel column chromatography eluting with gradient *n*-hexane-acetone (1:0 to 1:1, v/v). Fraction 7 (300 mg) was subjected to a silica gel flash column chromatography (*n*-hexane-acetone 1:0 to 7:3, v/v) to yield fifteen subfractions (7_A–7_O), subfraction 7_E (30 mg) was subjected to preparative TLC (*n*-hexane-acetone 8:2, v/v) to yield compound **3** (11 mg). Fraction 10 (200 mg) was chromatographed on a silica gel column (*n*-hexane-acetone 1.0 to 1:1, v/v) to yield eleven subfractions (10_A–10_K), subfraction 10_B was subjected to preparative TLC (*n*-hexane-

acetone 7:3, v/v) to yield compound **2** (24 mg) while preparative TLC of subfraction 10_D yielded compound **1** (29 mg).

*3.3.1. Cleogynone A (**1**)*

White crystalline, mp 180 – 181 °C; $[\alpha]^{20}_D$ +184.5 (c = 0.14, MeOH); UV (MeOH) λ_{max} (log ε): 202 (2.72), 219 (2.50), 282 (1.81) nm; IR ν_{max}: 3466, 2960, 2929, 2860, 1723, 1462, 1376, 1267, 1250, 1122, 1072, 1039, 947, 742 cm^{-1}; HRESIMS *m/z*: 559.3996 $[M - H_2O + H]^+$ (calculated for $C_{34}H_{55}O_6$, 559.3999); ^1H and ^{13}C NMR shown in Table 2.

*X-ray crystallographic data compound **1**:* $C_{34}H_{56}O_7$, M = 576.79, monoclinic, a = 7.8377(5) Å, b = 16.5237(10) Å, c = 12.3065(8) Å, α = 90°, β = 99.120(2)°, γ = 90°, V = 1573.64(17) $Å^3$, T = 150 K, space group $P2_1$, Z = 2, μ(CμKα) = 0.664 mm^{-1}, 21101 reflections measured, 5925 independent reflections (R_{int} = 0.023). The final R_1 values were: R = 0.0280, wR_2 = 0.0749 and S = 1.046.

*3.3.2. Cleogynone B (**2**)*

White crystalline, mp 142 – 144 °C; $[\alpha]^{20}_D$ +50.8 (c = 0.19, MeOH); UV (MeOH) λ_{max} (log ε): 204 (2.00), 220 (2.50), 283 (1.81) nm; IR ν_{max}: 3502, 2941, 2880, 1718, 1699, 1452, 1376, 1350, 1246, 1222, 1173, 1137, 1034, 954, 899 cm^{-1}; HRESIMS m/z 501.3971. $[M - H_2O + H]^+$ (calculated for $C_{32}H_{53}O_4$, 501.3944); ^1H and ^{13}C NMR shown in Table 2.

*X-ray crystallographic data compound **2**:* $C_{32}H_{54}O_5$, M = 518.75, triclinic, a = 7.4475(5) Å, b = 7.7239(5) Å, c = 14.5814(10) Å, α = 97.298(3)°, β = 93.052(3)°, γ = 115.003(3)°, V = 748.66(9) $Å^3$, T = 150 K, space group $P1$, Z = 1, μ(CμKα) = 0.592 mm^{-1}, 39127 reflections measured, 5573 independent reflections (R_{int} = 0.049). The final R_1 values were: R = 0.0479, wR_2 = 0.1278 and S = 1.036.

3.3.3. Anti-cancer assays

All cell lines were cultured at 37°C, 5% CO_2, in Dulbecco's Modified Eagle Medium (DMEM) containing high glucose (containing 4.5 g/L glucose, sodium bicarbonate and phenol red), supplemented with 10 % fetal bovine serum (FBS), Glutamax (2 mM) and PenStrep (0.5 mg/mL).

MDA-MB-468, A549, HCT-15 and HCT-116 cells were trypsinized at 10,000 cells per well in a 100 µL volume plated in 96 well plate. The cells were incubated in CO_2 to reach 60 - 70% confluence, at which point, they were treated with increasing concentration of test sample along with positive control (Cisplatin-30.11 µg/mL and/or diallyl disulfide (DADS) 146.2 µg/mL), incubated over 24 hours and 48 hours and cell viability determined using sulforhodamine B (SRB) assay. After 24- and 48-hour incubation cells were fixed using 50 µl of 50% trichloroacetic acid (TCA) was added to each well, and the plates were kept at 4°C for 60 minutes. After fixation of cells, the wells were carefully washed with slow running tap water, and subsequently the plates were allowed to dry. One hundred microliters of SRB was added to each well and the plates incubated for 30 min. Excess unbound SRB was removed and the plates washed with 1% acetic acid. The washed wells were allowed to dry at room temperature. The protein bound SRB was dissolved in 100µL of 10 mM tris base solution and the absorbance read at 510 nm on a multimode plate reader (PerkinElmer, Massachusetts, USA)

Declaration of Competing Interest

The authors declare that there is no conflict of interest

Acknowledgement

This work was funded by the National Research Foundation (Grant number 96718, 2016 – 2020) and the Department of Science and Innovation (Grant number DST/CON 0060/2019, 2016 – 2020) in collaboration with the Council for Scientific and Industrial Research, South Africa and the University of Pretoria.

References

Bala, A., Kar, B., Haldar, P.K., Mazumder, U.K., Bera, S., 2010. Evaluation of anticancer activity of *Cleome gynandra* on Ehrlich's Ascites Carcinoma treated mice. J. Ethnopharmacol. 129, 131–134. https://doi.org/10.1016/j.jep.2010.03.010

Bala, A., Kar, B., Karmakar, I., Kumar, R.B.S., Haldar, P.K., 2012. Antioxidant activity of Cat's whiskers flavonoid on some reactive oxygen and nitrogen species generating inflammatory cells is mediated by scavenging of free radicals. Chin. J. Nat. Med. 10, 321–327. https://doi.org/10.1016/S1875-5364(12)60065-X

Bruker, 2016. *APEX3*, *SAINT-Plus* and *SADABS*. Bruker AXS Inc., Madison, Wisconsin, USA.

Chweya, J.A., Mnzava, N.A., 1997. Cat ' s whiskers, International Plant Genetc Ressources Institute, 7 - 22.

Das, P.C., Patra, A., Mandal, S., Mallick, B., Das, A., Chatterjee, A., 1999. Cleogynol, a novel dammarane triterpenoid from *Cleome gynandra*. J. Nat. Prod. 62, 616–618. https://doi.org/10.1021/np9803528

Ekpong, B., 2009. Effects of seed maturity, seed storage and pre-germination treatments on seed germination of cleome (*Cleome gynandra* L.). Sci. Hortic. (Amsterdam). 119, 236–240. https://doi.org/10.1016/j.scienta.2008.08.003

Flack, H.D., 1983. On enantiomorph-polarity estimation. Acta Crystallogr. Sect. A A39, 876–881.

Flack, H.D., Bernardinelli, G., 2000. Reporting and evaluating absolute-structure and absolute-configuration determinations. J. Appl. Crystallogr. 33, 1143–1148. https://doi.org/10.1107/S0021889800007184

Ferlay, J., Soerjomataram, I., Ervik, M., Dikshit, R., Eser, S., Mathers, C., 2012. GLOBOCAN v1.0, Cancer Incidence and Mortality Worldwide: IARC CancerBase No. 11

Gupta, S.C., Kim, J.H., Prasad, S., B.B.A., 2010. Regulation of survival, proliferation, invasion, angiogenesis, and Cancer metastasis of tumor cells through modulation of inflammatory pathways by nutraceuticals. Cancer Metastasis Rev 29, 405–434. https://doi.org/10.1007/s10555-010-9235-2.Regulation

Kamatenesi-Mugisha, M., Oryem-Origa, H., 2007. Medicinal plants used to induce labour during childbirth in western Uganda. J. Ethnopharmacol. 109, 1–9. https://doi.org/10.1016/j.jep.2006.06.011

Mai, H.L., Grellier, P., Prost, E., Lemoine P., Poullain C., Dumontet V., Deguin B., Hue Vo T.B., Michel S., Grougnet R., 2016., Triterpenes from the exudate of *Gardenia urvillei.* Phytochemistry, 122, 193-202

Mishra, S.S., Moharana, S.K., Dash, M.R., 2011. Review on cleome gynandra. Int. J. Res. Pharm. Chem. 1, 681–689.

Narendhirakannan, R.T., Subramanian, S., Kandaswamy, M., 2007. Anti-inflammatory and lysosomal stability actions of *Cleome gynandra* L. studied in adjuvant induced arthritic rats. Food Chem. Toxicol. 45, 1001–1012. https://doi.org/10.1016/j.fct.2006.12.009

Narendhirakannan, R.T., Subramanian, S., Kandaswamy, M., 2005. Free radical scavenging activity of *Cleome gynandra* L . leaves on adjuvant induced arthritis in rats, Molecular and Cellular Biochemistry, 276, 71–80.

Sheldrick, G.M., 2015a. Crystal structure refinement with SHELXL. Acta Crystallogr. Sect. C Struct. Chem. C71, 3–8.

Sheldrick, G.M., 2015b. SHELXT - Integrated space-group and crystal-structure determination. Acta Crystallogr. Sect. A Found. Adv. A71, 3–8.