Talanta 120 (2014) 181–190

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Chromatographic analysis with different detectors in the chemical characterisation and dereplication of African propolis

Tong Zhang [a,*], Ruwida Omar [a], Weam Siheri [a], Sultan Al Mutairi [a], Carol Clements [a], James Fearnley [b], RuAngelie Edrada-Ebel [a], David Watson [a]

[a] *University of Strathclyde, Department of Pharmaceutical Sciences, Strathclyde Institute of Pharmacy and Biomedical Science, 27 Taylor Street, Glasgow G4 0NR, UK*
[b] *BeeVital, Whitby, North Yorkshire YO22 5JR, UK*

ARTICLE INFO

Article history:
Received 6 October 2013
Received in revised form
26 November 2013
Accepted 30 November 2013
Available online 11 December 2013

Keywords:
Propolis
Chemical profiling
Dereplication
Liquid Chromatography–High Resolution Mass Spectometry (LC–HRMS)
Stilbenoid
Principal Component Analysis (PCA)

ABSTRACT

Propolis or bee glue has very diverse composition and is potentially a source of biologically active compounds. Comprehensive chemical profiling was performed on 22 African propolis samples collected from the sub-Saharan region of Africa by using various hyphenated analytical techniques including Liquid Chromatography (LC)–UltraViolet Detection (UV)–Evaporative Light Scattering Detection (ELSD), LC–High Resolution Mass Spectrometry (HRMS), Gas Chromatography (GC)–MS and LC–Diode Array Detector (DAD)–HRMS/MS. The diversity of the composition of these African propolis samples could be observed by heat mapping the LC–UV and ELSD data. The characteristic chemical components were uncovered by applying Principal Component Analysis (PCA) to the LC–HRMS data and a preliminary dereplication was carried out by searching their accurate masses in the Dictionary of Natural Products (DNP). A further identification was achieved by comparing their GC–MS or LC–DAD–HRMS/MS spectra with previously published data. Generally no clear geographic delineation was observed in the classification of these African propolis samples. Triterpenoids were found as the major chemical components in more than half of the propolis samples analysed in this study and some others were classified as temperate and Eastern Mediterranean type of propolis. Based on the comparative chemical profiling and dereplication studies one uncommon propolis from southern Nigeria stood out from others by presenting prenylated isoflavonoids, which indicated that it was more like Brazilian red propolis, and more significantly a high abundance of stilbenoid compounds which could be novel in propolis.

© 2013 Elsevier B.V. All rights reserved.

1. Introduction

Rapid identification of known compounds from chemical profiling of natural resources, referred to as dereplication, is becoming increasingly important for phytochemistry based drug discovery [1–3] via the targeting of novel compounds. With the development of hyphenated analytical techniques the ability to analyse individual components in a complex mixture has been significantly improved [1,3,4]. In the case of LC–MS the enhancement of LC column efficiency has delivered higher peak capacity allowing separation of greater numbers of components and High Resolution MS (HRMS) and Collision Induced Dissociation (CID) techniques offer the measurement of accurate masses of the molecular ions and fragments of the individually eluting molecules providing valuable information for their structure elucidation [5]. Electron Ionisation (EI), as the interface technology of GC–MS, is able to generate reproducible fragmentation patterns for gaseous molecules. Based on this feature an EI–MS standard spectral library including hundreds of thousands of compounds has been built up by National Institute of Standards and Technology (NIST) and is widely used in identification of unknown compounds in many laboratories. LC–Solid Phase Extraction (SPE)–NMR has also been reported as a powerful analytical platform for dereplication [6]. It is able to generate more accurate structural information than MS based hyphenated techniques. However, the sensitivity and the intricate instrumentation configuration limits its wide application for this purpose.

Propolis (bee glue) is a sticky and dark-coloured material harvested by honey bees in order to seal cracks of the hives and more importantly eliminate biological contamination in the colony. It has been reported to have various biological and pharmacological properties, attracting intense interest in its medicinal applications [7–9]. It has the advantage as source of biologically active compounds that it has already been selected by bees for its biological activity and is collected from plants in a non-destructive way. The complex chemical composition of propolis remains as a

* Correspondence to: SIPBS, 161 Cathedral Street, Glasgow G4 0RE, UK. Tel.: +44 14 15482651; fax: +44 14 15526443.
E-mail address: tong.zhang.101@strath.ac.uk (T. Zhang).

0039-9140/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.talanta.2013.11.094

challenge for purification of bioactive components by conventional phytochemistry. In addition, the chemical and bioactive characteristics of propolis are highly dependent on its geographic origin [9–12]. Up to now the propolis from Europe, Asia, North and South America has been well studied presenting a comprehensive information resource for bioactive natural products found in propolis. For instance, polyphenols including flavonoids, phenolic acids and their esters are general major bioactive chemical components in the poplar propolis from the temperate zones of both Northern and Southern hemispheres [9,11] and prenylated benzophenones and diterpenes are characteristic chemical components of the propolis from tropical zones (Northeast Brazil and Cuba) [13] and Eastern Mediterranean regions (Greece, Crete and Turkey) [14], respectively.

Recently chemical profiling and dereplication using LC–MS or GC–MS has been employed for chemical evaluation of commercial propolis products or wild propolis [13–22]. Common chemical components in propolis crude extracts were quickly recognised by comparing DAD, CID–MS^n or EI–MS spectra of the chromatographic peaks with authentic standard compounds' or previously published data leading to uncommon ones being uncovered and then putatively identified by interpreting the DAD and MS data. By applying this strategy a large number of samples can be quickly characterised in chemical composition and unlike in conventional phytochemistry only small amount of propolis is consumed for analysis. By dereplication of the chemical components and comparison with known types of propolis the analysed samples can be categorised or even defined as new types if novel compounds are discovered [13]. From our literature review there is only limited research data on the African propolis but the uniqueness of its chemical composition has been reported [11,23,24]. The aim of this study was to investigate the basic chemical composition of the African propolis collected from 9 countries in the wide region of sub-Saharan Africa. Using various hyphenated analytical techniques we intended to characterise these propolis samples by a comprehensive chemical profiling of their ethanolic extracts and hoped to uncover some novel compounds by dereplication studies for targeted isolation in the future.

2. Experimental

2.1. Chemicals and solvents

HPLC grade acetonitrile (ACN) and ethyl acetate was purchased from Fisher Scientific, UK. HPLC grade water was produced by a Direct-Q 3 Ultrapure Water System from Millipore, UK. AnalaR grade formic acid (98%) and ethanol were obtained from BDH-Merck, UK. N-Methyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) was purchased from Sigma-Aldrich, UK.

2.2. Propolis sample collection and preparation

The propolis samples were gradually collected by BeeVital from sub-Saharan African countries in the past 10 years (Supporting information 1) and stored at room temperature with dark and dry conditions before the extraction. Approximately 50 mg of propolis was cut off from the core of each sample and extracted with ethanol by ultra-sonication with heat at 40 °C for 2 hours. The filtered solution was dried by nitrogen flow and the amount of residue was measured by subtracting the weight of empty vial from the total weight. Finally each residue was reconstituted with ethanol at a concentration of 5 mg/ml as the stock solution for the following analysis.

2.3. LC–UV–ELSD

1 ml of each stock solution was dried by nitrogen flow and reconstituted as 5 mg/ml with the mobile phase at the ratio of the initial composition of the LC gradient programme. The LC–UV–ELSD analysis was performed on an Agilent 1100 system (Agilent Technologies, Germany) consisting of a quaternary pump, an autosampler, a degasser and a UV single channel (290 nm) detector coupled with an Evaporative Light Scattering Detector (ELSD) (model: SEDEX75, SEDERE France) at 30 °C. An ACE C18 column (150 × 3 mm, 3 μm) (HiChrom, Reading UK) was employed for separation with 0.1% v/v formic acid in water as mobile phases A and 0.1% v/v formic acid in ACN as B at the flow rate of 300 μl/min. The injection volume was 10 μl. The gradient elution was programmed as follows: 0–15 min linear gradient from 30% to 50% of B, 15–25 min at 50% of B, 25–40 min linear gradient from 50% to 80% of B, 40–50 min at 80% of B, 50–51 min increasing to 100% of B, 51–59 min at 100% of B with the flow rate increasing to 500 μl/min for cleaning the column and 60–70 min back to 30% of B for re-equilibration for next run.

2.4. LC–HRMS and DAD–$HRMS^n$

The same samples and the chromatographic conditions described in Section 2.3 were used on an Accela 600 HPLC system combined with an Exactive (Orbitrap) mass spectrometer from Thermo Fisher Scientific (Bremen, Germany). The MS detection range was from 100 to 1500 m/z and the scanning was performed under ESI polarity switching mode. All detailed MS settings were used the same as described in our previous study [25]. The data dependent MS^n fragmentation was carried out by using Collision Induced Dissociation (CID) at 35 V on a LTQ-Orbitrap mass spectrometer combined with a Surveyor HPLC system from Thermo Fisher Scientific (Bremen, Germany) including on-line DAD (200–600 nm) and UV at 290 nm analysis. Again the chromatographic conditions described in Section 2.3 were used. Each whole MS scan consisted of three segments: MS full scan from 100 to 1500 m/z; MS/MS on the most intense m/z signal in the full scan; MS/MS on the most intense fragment generated in the MS^2. By switching on the dynamic exclusion function each m/z signal would be selected for MS/MS fragmentation only three times within 1 minute and then the selection would move to the next most intense m/z signal and so on.

2.5. GC–Ms

1 ml of stock solution was dried by nitrogen flow and reconstituted with 1 ml of ethyl acetate. For derivatisation evaluation the sample was prepared as described in Ref. [14] with MSTFA instead of bis(trimethylsilyl)-trifluoroacetamide (BSTFA). 1 μl of each prepared sample was injected in splitless mode at 280 °C into the GC–MS (Focus GC-DSQ2) system from Thermo Fisher Scientific (Bremen, Germany) equipped with a 30 m long, 0.25 mm i.d., and 0.25 μm film thickness InertCap 1 MS capillary column from GL Sciences (Japan). The temperature gradient was programmed as follows: initially holding at 100 °C for 2 min, linearly increasing to 280 °C at the rate of 5 °C/min, holding at 280 °C for 15 mins and linearly increasing to 320 °C at the rate of 10 °C/min and holding for 10 mins. The temperature was 250 °C in the ionisation source and the ionisation voltage was 70 eV for EI–MS in positive mode.

2.6. Software and data processing

MZMine 2.10 [26] was used for LC–HRMS data processing. The procedure and the settings were the same as described in our previous study [27]. The generated peak lists from both ESI

positive and negative modes were combined and imported to SIMCA-P 13 (Umetrics, Sweden) for Principal Component Analysis (PCA) with Pareto scaling. Using an in-house macro coded by Visual Basic Application in Excel (MicroSoft Office 2010) the first 100 LC–HRMS features from each sample were selected based on the peak area and putatively identified by searching for the accurate mass in Dictionary of Natural Products (version 2013). Xcalibur 2.2 from Thermo Fisher Scientific was used to check the raw LC–HRMS and GC–MS data and generate the MS based chromatograms shown in the manuscript. Clarity from DataApex was used to handle the LC–UV–ELSD data and the heatmaps were produced by using the heatmap.2 function in the Gplot package from R language.

3. Results and discussion

3.1. Preliminary characterisation by HPLC–UV–ELSD

It is well known that poplar type propolis is the most extensively researched type and its major characteristic chemical components are phenolic compounds [9]. Therefore our investigation started with a HPLC gradient method coupled with one selective (UV at 290 nm) and followed by one universal (ELSD) detector which can detect almost every component in a sample. The outcomes of analysis using LC–UV and ELSD are illustrated in heatmap format in Fig. 1 to facilitate the visualisation of chromatographic results between samples which are sorted in the same order for both ELSD and UV figures based on origin countries. For each sample the signal responses were collected at every retention time point with 0.5 min intervals across 22 samples the responses at the same retention time were calculated as Z-scores which are reflected by the gradient coloured bars from red (low) to blue (high) and samples with the same chemical composition should show a similar colour pattern. Generally these propolis samples demonstrated variable chromatographic data although some similarities could be observed between a few samples collected from the same countries, e.g. 20-5U and 23-2U. It should be noticed that some samples (S61K, S151K and D46SA for instance) showed intense responses with the ELSD detector at certain retention times but not with UV implying the presence of non-phenolic compounds lacking chromophores such as terpenoids. Some samples such as 9SA and RS-N showed highly rich chromatographic data with both detectors in the middle of the retention window, in contrast others such as S149U and S40Z only showed responses at both the ends which was also observed in our previous study [11]. There was no clear chromatographic similarity shown between the five samples from South Africa and sample RS-N could be easily distinguished from the other two Nigerian samples by its intensely rich chromatographic responses with both ELSD and UV detectors.

3.2. Comprehensive characterisation by HPLC–HRMS

The above chromatographic data from LC–UV and ELSD was not enough on its own to completely characterise the chemical composition of these propolis samples. LC–HRMS analysis was then performed using the same LC gradient programme and in order to correlate with the previous LC–ELSD data on-line UV detection at 290 nm was also conducted before LC eluent flowed into ESI source. From our literature review only limited research data are available from chemical profiling of propolis using LC–HRMS. In this study by taking the advantages of high sensitivity and mass accuracy of HRMS we expected a comprehensive detection not only for the major but also for the minor chemical components in propolis. The base peak chromatograms of two propolis samples in both ESI positive and negative modes are shown as examples in Fig. 2. Previously sample RS-N showed rich chromatographic responses with both UV and ELSD but sample S61K showed almost nothing with UV. As can be observed in Fig. 2 both propolis samples showed extremely complicated chromatograms with HRMS in both positive and negative ESI modes. From manually checking the raw LC–HRMS data most Individual chromatographic peaks were dominated by single m/z signals which were generated by the protonated/deprotonated molecular ions. This observation indicated a good separation of the chemical components of the propolis which correlated with the ELSD traces

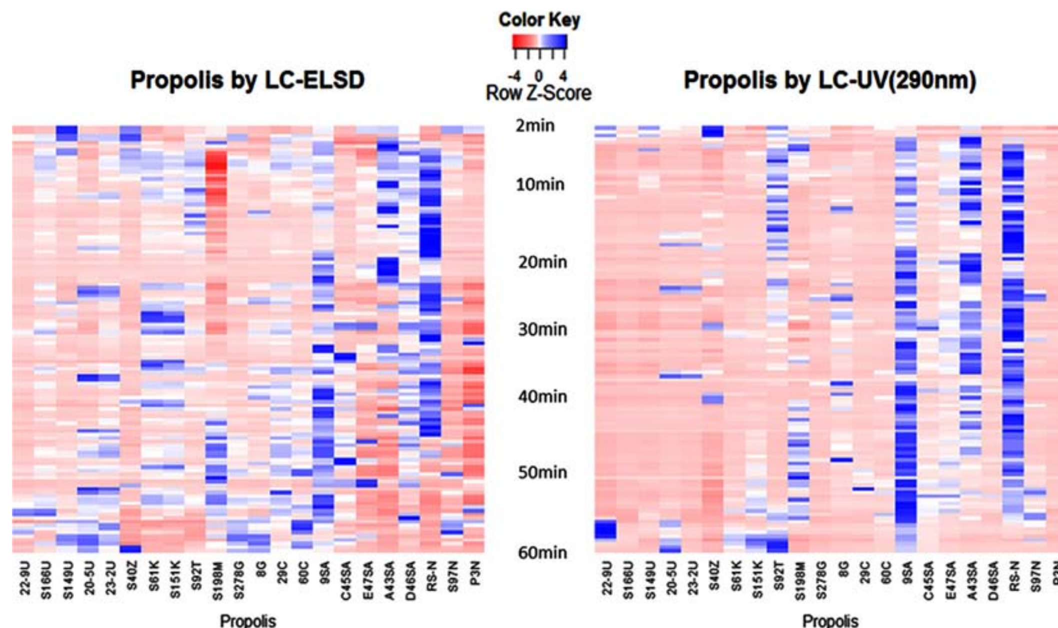

Fig. 1. Heat mapped LC–ELSD and LC–UV (290 nm) data of 22 propolis samples.

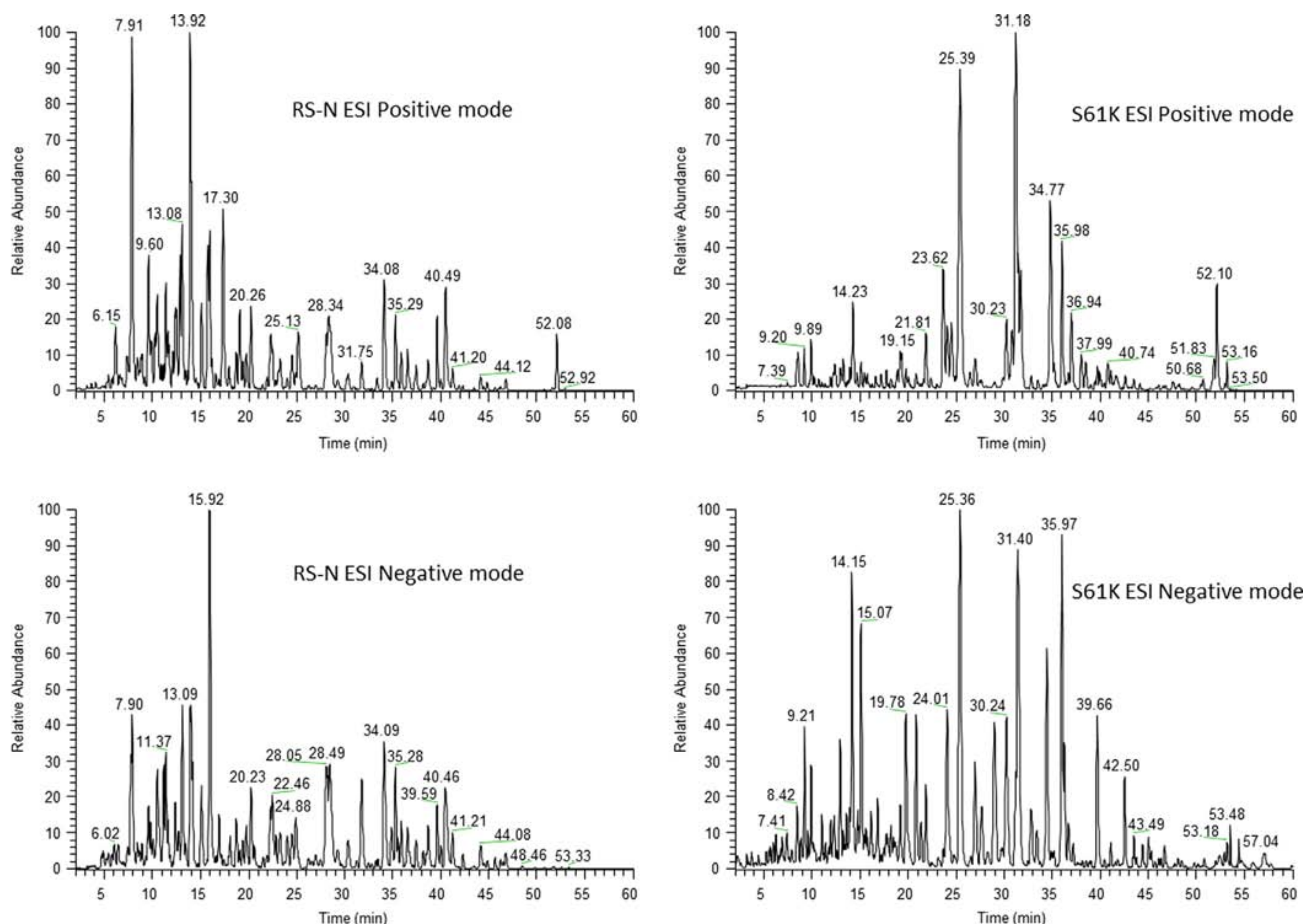

Fig. 2. Representative LC–HRMS chromatograms for two propolis samples (RS-N and S61K) in both ESI positive and negative modes.

and indicating that the occurrence of ion suppression in the ESI source was avoided to a great extent.

Principle component analysis (PCA) has usually been applied for interpreting complex data matrices generated by LC–HRMS. Here we expected to classify the African propolis samples and target the major characteristic chemical components responsible for the classification. However, it is important to bear in mind that ESI-MS is not a universal detector and it does not accurately reflect abundance of the chemical components in the sample. The generated PCA score plots and their corresponding loading plots are shown in Fig. 3. The overlap of replicate data of one sample (S198M, coloured in purple) measured at the beginning and the end of the experiment gave an indication that the LC–HRMS profiling was stable over the period of the experiment. At the first glance at Fig. 3A two outliers (RS-N and A43SA) were recognised and the samples from South Africa gradually shifted increasingly along one direction away from the other samples which were clustered close to each other. These observations were well matched with the LC–ELSD and UV results shown in Fig. 1. More importantly by checking the corresponding loading plot (Fig. 3B) it was straightforward to target the LC–HRMS features resulting in the distribution of propolis samples in the score plot (Fig. 3A) and consequently they were determined as major characteristic chemical components in the relevant samples (Fig. 3B). It is believed that sample S198M could share some common chemical components with RS-N but in low abundance because in Fig. 3A they are the only samples at the same quarter but are far away from each other. It is the same case for South African samples. In order to obtain a more specific classification on the samples clustered in the middle of Fig. 3A the four outlying samples (RS-N and three South African samples) were excluded and a new PCA was performed on the rest (Fig. 3C). 18 samples were seperated as three clusters. It is not surprising to see the isolation of Malawian sample because it was separated from the main cluster in Fig. 3A. The closeness of Kenyan and South African samples indicated that they shared some common characteristic chemical components. This was also reflected by the fact that they demonstrated several blue bars with ELSD at the same retention times but no UV absorption in Fig. 1. The rest of samples comprised the major group and the absence of LC–HRMS features with high MS response could be the main reason for their clustering. It should be noted that these samples also showed lower ELSD and even zero UV responses in Fig. 1. However, a few blue bars can be observed at early retention times in sample S92T, S40Z and S149U which also slightly removed from the core of the cluster.

It is believed that no nitrogenous compounds have been isolated from propolis so far [9]. With the limitation on the number of nitrogen to 1, just in case the compounds generated a NH_4^+ rather than a proton adduct, elemental composition prediction was performed on all the LC–HRMS features used in the PCA based on their accurate mass. Most features circled in Fig. 3 were exclusively assigned to single chemical formulas within a 3 ppm mass error window and associated with knowledge of propolis chemistry the classifications of these characteristic chemical components are proposed according to the putative identifications given in Table 1. The deduction was generally based on the number of carbons and oxygens in the molecules and the Ring and Double Bonds (RDB) equivalence. For example, flavonoids

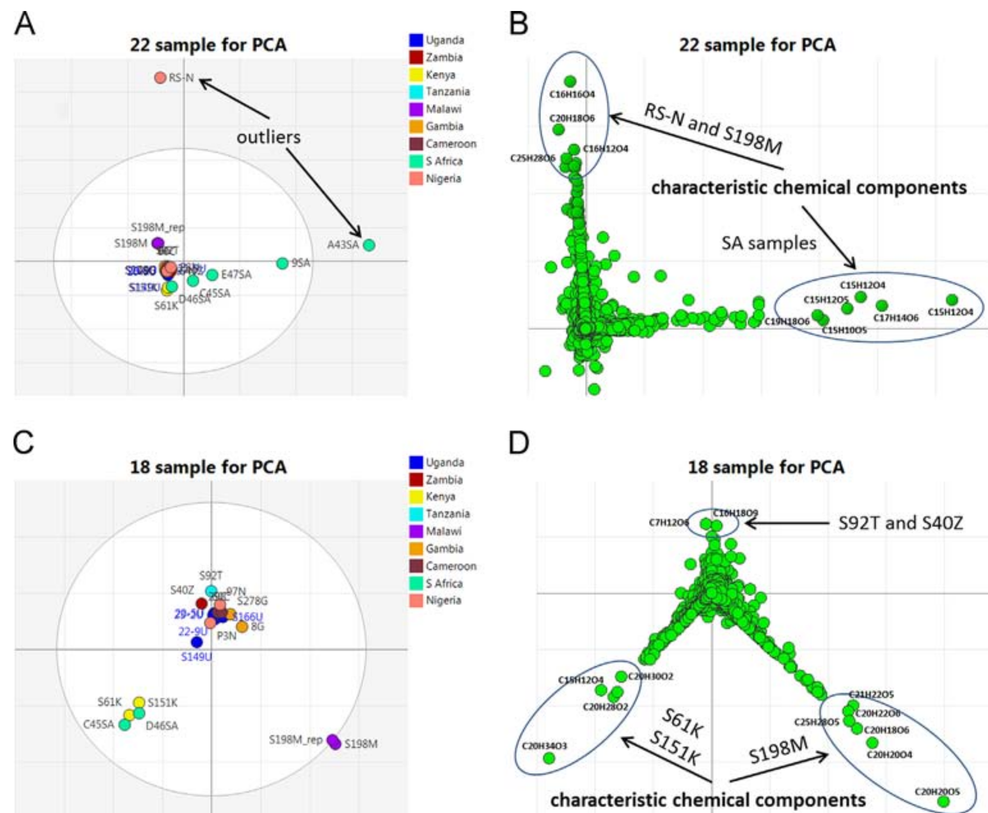

Fig. 3. PCA score plots generated by LC–HRMS data of (A) 22 samples and (C) 18 samples and their corresponding loading plots (B) and (D).

Table 1
The categorisation of propolis samples in Fig. 3A and C and the LC–HRMS information of the circled features in Fig. 3B and D.

Propolis sample	MZMIne ID	m/z	Rt (min)	Formula	RDB	Chemical classification
RS-N and S198M	N_198	353.103	28.50	C20H18O6	12	Prenylated flavonoid
	N_431	339.124	28.16	C20H20O5	11	Prenylated flavonoid
	N_2778	271.098	15.98	C16H16O4	9	Flavonoid (Flavan)
	N_4574	267.066	13.98	C16H12O4	11	Flavonoid
	N_14795	315.088	7.82	C17H16O6	10	Flavonoid
	N_32266	255.066	13.32	C15H12O4	10	Flavonoid
	N_32267	283.061	7.86	C16H12O5	11	Flavonoid
	P_15613	269.081	13.98	C16H12O4	11	Flavonoid
	P_24850	285.076	7.89	C16H12O5	11	Flavonoid
	N_2233	423.182	28.31	C25H28O6	12	Diprenylated flavonoid
	N_469	423.182	34.14	C25H28O6	12	Geranylated flavonoid
	N_3437	447.254	40.51	C29H36O4	12	***Unknown***
A43SA, 9SA and E47SA	N_94	255.066	18.73	C15H12O4	10	Flavonoid
	P_287	257.081	18.74	C15H12O4	10	Flavonoid
	N_150	313.072	19.46	C17H14O6	11	Flavonoid
	N_184	271.061	11.54	C15H12O5	10	Flavonoid
	N_165	269.046	18.94	C15H10O5	11	Flavonoid
	N_2702	341.103	31.07	C19H18O6	11	Flavonoid
S198M	N_24	339.124	19.78	C20H20O5	11	Prenylated flavonoid
	N_30	323.129	36.63	C20H20O4	11	Prenylated flavonoid
	N_36	353.14	35.14	C21H22O5	11	Prenylated flavonoid
	N_116	357.134	11.03	C20H22O6	10	Prenylated flavonoid
	N_16	407.187	43.37	C25H28O5	12	Geranylated flavonoid
S61K, S151K, D46SA, C45SA	N_464	299.202	35.99	C20H28O2	7	Diterpenoid
	P_234	323.258	31.21	C20H34O3	4	Diterpenoid
	P_549	303.232	25.37	C20H30O2	6	Diterpenoid
	P_637	301.216	35.96	C20H28O2	7	Diterpenoid
S92T, S40Z, S149U	N_1	353.088	3.19	C16H18O9	8	Chlorogenic acid
	N_5	191.056	2.36	C7H12O6	2	Hydroxyl acid

MZMine ID: "P" means the feature is generated in ESI positive mode and "N" in ESI negative mode. The following number was automatically generated in MZMine as an index.
RDB: Ring and Double Bonds (RDB) equivalence.

have the carbon backbone of C6 (A ring)–C3 (C ring)–C6 (B ring) therefore the number of carbons would be at least 15. The additional carbons could be contributed from the methyl and/or acyl groups in the molecules. The RDB of flavonoids should be about 10 depending on the C ring structure. When the number of carbons reaches 20 with the RDB value increasing it is a high possibility that the compound is a prenylated flavonoid or a diterpenoid if the RDB value is lower. For further characterisation of these propolis samples the first 100 LC–HRMS features of each sample based on the peak area were putatively identified by searching for the accurate mass (± 3 ppm) in the Dictionary of Natural Products (DNP) using an in-house Microsoft Excel macro.

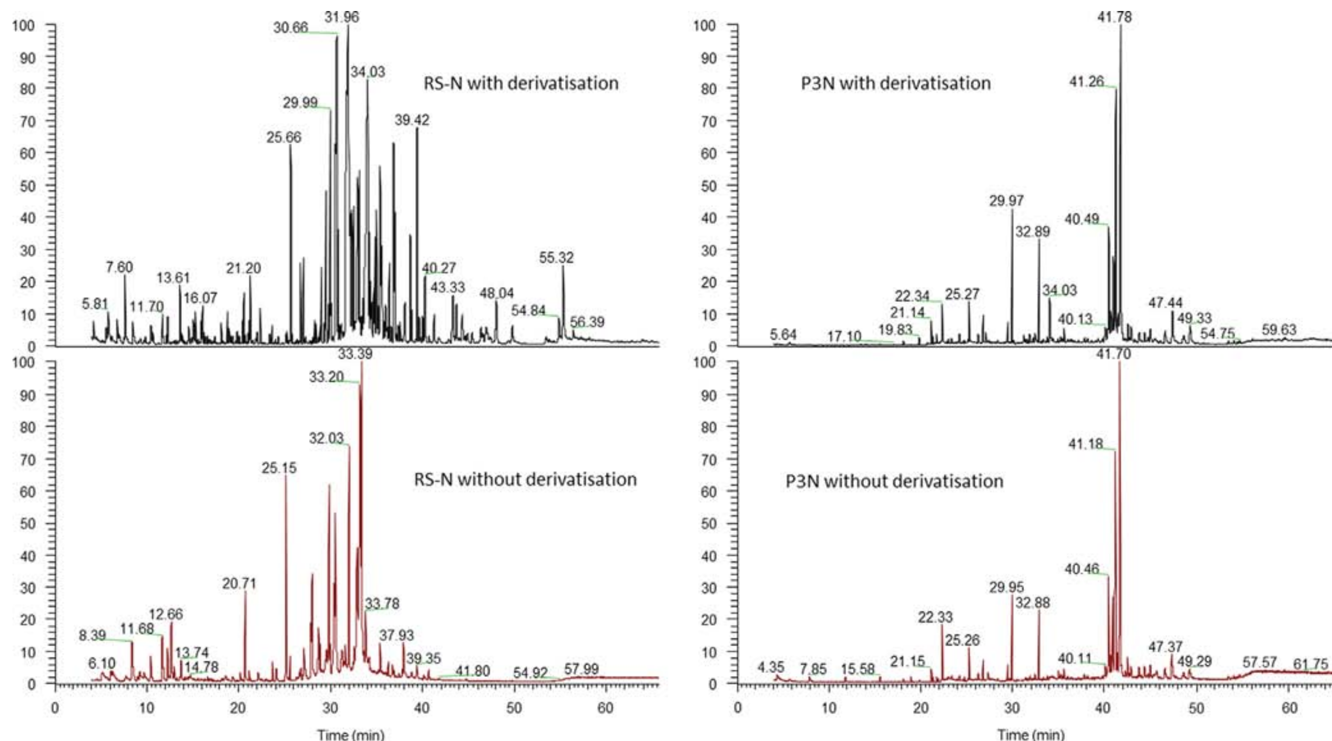

Fig. 4. Representative GC–MS chromatograms for two propolis samples (RS-N and P3N) with and without derivatisation by MST.

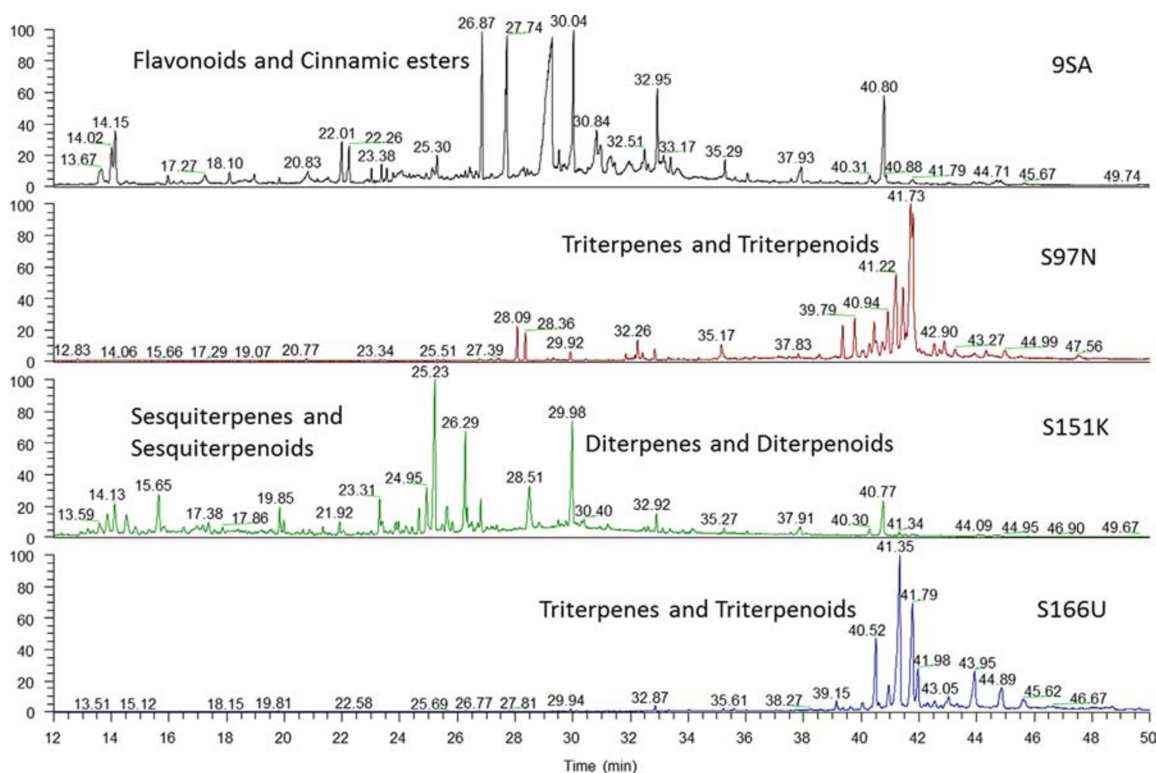

Fig. 5. Representative GC–MS chromatograms for four propolis samples (9SA, S97N, S151K and S166U) and the chemical classification of their major peaks.

By browsing the results (SI 2) attention should be drawn to some indicative information such as the possibility of prenylated benzophenones in sample RS-N and S198M, phenolic acids in South African samples and triterpenoids in Ugandan samples. However, further identification of individual chemical components could not be achieved because one feature was usually linked to multiple isomers in DNP. Therefore it was necessary to carry out a more selective analysis.

3.3. Complementary characterisation by GC–MS

Prior to analysing all samples with GC–MS the performance of the derivatisation procedure was evaluated by running two samples with a completely different LC–UV and LC–HRMS traces. The GC–MS chromatograms of sample RS-N and P3N with and without silylation by MSTFA are shown in Fig. 4. It is clear that sample RS-N demonstrated a rather more complicated chromatogram following derivatisation resulting from the detection of the less volatile polar phenolic compounds after silylation. In contrast almost no change could be found in sample P3N when it was analysed with and without derivatisation. In consideration of the fact that phenolic compounds should be readily detected by LC–UV and LC–HRMS, plus their presence would increase the difficulty of data interpretation. GC–MS analysis was carried out without derivatisation. In order to display the distinctions between different types of samples profiled by GC–MS some typical samples were selected and their GC–MS chromatograms are shown in Fig. 5. S97N and S166U and all the samples from the major group in Fig. 3C demonstrated similar chromatograms showing a bunch of intense peaks around 41 min, on the other hand almost nothing could be observed after 40 min in the rest samples except for S198M. By searching the NIST library those intense peaks were putatively identified as various triterpenes and triterpenoids with similarity scores of more than 800 and they are believed to be the major chemical components in these samples because as shown in Fig. 4 almost no polar compounds were detected after derivatisation of these samples. This result explained the observation that these samples presented lower ESI–HRMS and zero UV (290 nm) responses in the previous LC experiments and the majority of them gave late running peaks detectable by ELSD. For a typical example as shown in Fig. 6 sample S166U showed no UV and extremely low ESI-MS responses in both ionisation modes paralleled with sample RS-N, whereas its ELSD trace displayed intense signals only after 55 min. For other types of samples the GC–MS data further confirmed the chemical characterisation deduced by LC–HRMS data.

3.4. Specific characterisation by HPLC–DAD–HRMS^n

Based on the data analysis above these African propolis samples could have been grouped by the characteristic chemical components and in order to ultimately identify these important compounds some samples were selected for LC-DAD–HRMS^n analysis. As described in experimental section data dependent scanning was used to collect as many MS signals as possible at full scan level for fragmentation in each sample. By comparing the obtained MS^2 and MS^3 spectra with the published data [13,18,19,22,28–33] many characteristic chemical components were putatively identified. For

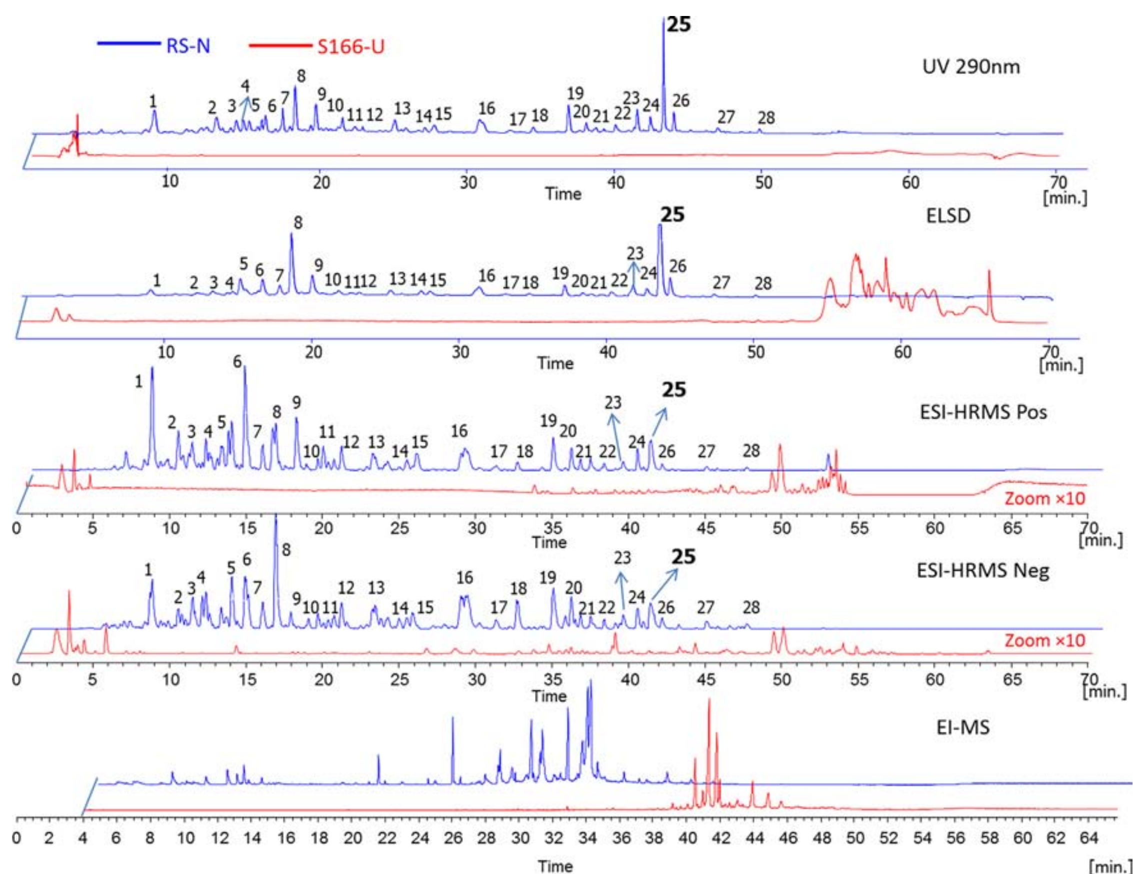

Fig. 6. Paralleled chromatograms for sample RS-N and S166U with various detectors. The Y-axis was set in the same scale for each individual paralleled pair except for ESI–HRMS Pos and Neg in which the trace of S166U was zoomed in by 10 times.

Table 2
Examples of identified characteristic chemical components in different types of propolis and the suitable analytical techniques for their detection.

Propolis sample	Formula	Compound name	Analytical technique	m/z (ESI polarity)	MS/MS [MS3]	References
RS-N	C20H20O5	Demethylxanthohumol	LC–HRMS/MS	341.1381 (+)	285(100) [165(100),191,179,121]	[31]
	C16H16O4	Neovestitol		273.1118 (+)	137(100),123,149,151,163	[13]
	C16H12O4	Formononetin		269.0803 (+)	254(100),237,213,241,136	[22]
	C15H12O4	Liquiritigenin		257.0805 (+)	147(100),239,137,211,242,163	[22]
	C16H12O5	7,3′-Dihydroxy-5′-Methoxy-isoflavon		285.0754 (+)	270(100),253,225,229,137,257	[13]
	C25H28O6	Demethylkuraridin		425.1953 (+)	285(100)[165(100),191,179,121]	[31]
A43SA, 9SA, E47SA	C15H12O4	Pinocembrin	LC–HRMS/MS	257.0804 (+)	153(100),131,173,215,179	[18,19]
	C17H14O6	Pinobanksin-3-O-acetate		313.0715 (−)	253(100) [209 (100),181,180,143,165,151]	[18,19]
	C15H12O5	Pinobanksin		271.0612 (−)	253(100),225,215,197,151,125,165, 185,157	[18,19]
	C15H10O5	Galangin		269.0454 (−)	197(100),213,227,169,241,225,223, 153,143,181	[18,19]
	C16H14O5	Pinobanksin-5-methyl ether		285.0706 (−)	239(100),267,224,253,179	[18,19]
	C16H14O4	Pinocembrin-5-methyl ether		269.0822 (−)	227(100),254,165	[18,19]
	C15H10O4	Chrysin		253.0505 (−)	209(100),211,181,179, 165,151,143,225	[18,19]
	C19H18O6	Pinobanksin-3-O-butyrate		341.1029 (−)	253(100) [209 (100),181,180,165,143,151]	[18,19]
S198M	C33H42O4	Nemorosone	LC–HRMS/MS	501.3001 (−)	432(100) [363 (100),271,417,327,234,389,377]	[33]
	C25H28O5	Prenylated benzophenone		407.1861 (−)	338(100) [232(100),218,295, 270,282]	
S61K, S151K, D46SA, C45SA	C20H28O2	Diterpenic acid	GC–MS	Similarity score was not high enough to exclusively identify the compound		
	C20H34O3	Diterpenic acid				
	C15H24O	sesquiterpene alcohol				
Others	C30H50O	Lupeol/Amyrin	GC–MS	218 and 203 are common major fragments		NIST with similarity more than 800
	C32H52O2	Lupeol/Amyrin acetates				
S92T, S40Z, S149U	C16H18O9	Caffeoylquinic acid	LC–HRMS/MS	353.0876 (−)	191(100)[127,173,85,93,111, 109,171],179	[15]
	C7H12O6	Quinic acid		179.0563 (−)	143(100),134,113,101,89,161	

those compounds without any published data to compare with their MS/MS spectra could also help confirm the chemical classification to some extent. Associated with the identification results achieved by GC–MS the final classification of these African propolis samples and the examples of putatively identified characteristic chemical components are listed in Table 2.

There is no doubt that the three South African samples should be classified as temperate poplar propolis by the typical chemical components of flavonoids without B-ring substituents. The other two South African and two Kenyan samples consist mainly of different diterpenoid acids which are the characteristic chemical components in the propolis from the eastern Mediterranean regions. More than half of the African propolis samples analysed in this study showed triterpenoids as the major chemical components and interestingly some of them also contained polar compounds like caffeoylquinic acids. RS-N and S198M are the most unique samples found in this study. Both of them demonstrated distinct chromatographic outcomes in comparison to the other samples with all analytical techniques used above. Basically due to the detection of isoflavonoids, (poly)prenylated isoflavonoids and benzophenones it is believed that they are more like Brazilian red propolis as reported in references of 13 and 22. A more detailed data analysis was performed on the sample RS-N by integrating the UV, ELSD, DAD and $HRMS^n$ data. In Fig. 6 the major UV (290 nm) and ELSD chromatographic peaks in sample RS-N are numbered and their corresponding LC–HRMS features were also correlated. There is a small shift of retention times between UV/ELSD and HRMS data for sample RS-N and this might be due to some difference in the gradient forming by the two different HPLC systems. The retention time shift is even more marked for the highly lipophilic compounds in S166-U. By manually checking the LC–HRMS data it was found that some UV and ELSD chromatographic peaks were not exclusively constructed by single LC–HRMS features. With dynamic exclusion in data dependent scan the MS^n fragmentation was also achieved on these overlapped eluting compounds. Many of the labelled peaks were identified by accurate match of the obtained MS^n fragment patterns and maximum UV absorbance with the data reported in Refs. [13] and [22] (Table 3 and Supporting information 3). For those without obtaining well matched data a quick recognition of polyprenylated or/and geranylated compounds could also be achieved by the clear observation of loss of C4H8 or/and C9H16 at the MS^2 level. It should be noted that some late eluting components were putatively identified as polyprenylated stilbenoids including the largest peak (25) in the ELSD trace. By searching the accurate mass in DNP there was a hit on mappain for the LC–HRMS feature representing peak 25. Based on the structure of mappain the major fragments at MS^2 and MS^3 levels could be precisely explained. As can been seen in Fig. 7 those polyprenylated stilbenoids could be distinguished from polyprenylated flavonoids by losing H_2O or/and CO rather than breaking the C-ring bonds by retro Diels-Alder (RDA) fragmentation at the MS^3 level. In addition, all these compounds showed maximum UV absorbance at wavelength of 320 nm or higher which is the evidence of an extended conjugated system like stilbene in the molecule. However, their chemical structures will not be ultimately determined until the NMR analysis is carried out. According to the literature stilbenoids are not common components in propolis and only few have recently been isolated from the propolis collected in Vio, Kenya [23] and in Kangaroo Island, Australia [34]. The plant *Macaranga. schweinfurthii* was believed as the source of the propolis from Vio, Kenya and it widely exists in the forest zone in south Nigeria and West Cameroon [23] where the RS-N sample was collected. The abbreviation “RS” refers to

Table 3
Identification of the major components in sample RS-N.

Peak no	ELSD (min)	UV (min)	MS (min)	m/z	Formula	Compound name
1	8.7	8.3	7.63	285.0754 (+)	C16H12O5	7,3′-Dihydroxy-5′-methoxy-isoflavon
2	13.0	12.6	11.16	283.0612 (−)	C16H12O5	Galangin-5-methyl ether
3(1)	13.9	13.5	12.14	271.0962 (+)	C16H14O4	Hydroxy flavone [a]–Me ether
3(3)			12.27	301.1068 (+)	C17H16O5	3-Hydroxy-8,9-dimethoxy-pterocarpan
3(2)			12.2	333.0964 (+)	C17H16O7	3(3) with two more hydroxyl groups
4	14.3	13.9	12.68	303.1222 (+)	C17H18O5	Mucronulatol
5	14.8	14.4	12.92	257.0805 (+)	C15H12O4	Liquiritigenin
6(1)	16.3	15.7	13.7	269.0803 (+)	C16H12O4	Formononetin
6(2)		15.9	13.95	273.1118 (+)	C16H16O4	Vestitol
7(1)	17.5	16.8	14.94	241.0857 (+)	C15H12O3	Dihydroxychalcone or hydroxyflavanone
7(2)		17.1	15	271.0960 (+)	C16H14O4	Flavone
8(1)	18.3	17.9	15.6	271.0961 (+)	C16H14O4	Pinostrobin[a]
8(3)			15.84	273.1118 (+)	C16H16O4	Neovestitol
8(2)			15.72	523.1747 (+)	C32H26O7	Retusapurpurin A
9	19.7	19.3	17.21	271.0961 (+)	C16H14O4	Medicarpin
10	21.5	21.1	18.94	287.0911 (+)	C16H14O5	Trihydroxy flavanone[a]–Me ether
11	22.3	22.0	19.76	341.1381 (+)	C20H20O5	Sorphoraflavanone B
12	22.9	22.5	20.25	355.1173 (+)	C20H18O6	Prenyl flavane
13	25.0	24.6	22.47	371.1122 (+)	C20H18O7	Prenyl flavanone
14	27.1	26.7	24.61	287.1274 (+)	C17H18O4	7-O-metilvestitol
15	27.7	27.3	25.56	479.2785 (+)	C30H38O5	Polyprenyl flavonoid
16(1)	31.0	30.4	28.48	341.1381 (+)	C20H20O5	Demethylxanthohumol
16(2)			28.73	355.1173 (+)	C20H18O6	Luteone[a]
17	32.9	32.5	30.82	425.1953 (+)	C25H28O6	Kushenol F
18	34.5	34.0	32.1	381.2059 (+)	C24H28O4	Geranyl flavonoid
19	36.9	36.5	34.38	425.1953 (+)	C25H28O6	Geranyl flavonoid
20	38.1	37.7	35.52	409.2001 (+)	C25H28O5	Kushenol A
21	38.7	38.3	36.17	517.3302 (+)	C34H44O4	***Schweinfurthin C***[a]
22	40.1	39.6	37.6	563.3364 (+)	C35H46O6	***Schweinfurthin B***[a]
23	41.5	41.1	38.98	425.1953 (+)	C25H28O6	Demethylkuraridin
24	42.5	42.0	39.97	493.2580 (+)	C30H36O6	Sophoraisoflavanone D[a]
25	43.4	42.9	40.88	449.2681 (+)	C29H36O4	***Mappain***[a]
26	44.0	43.6	41.62	423.1795 (+)	C25H26O6	Macarangin[a]
27	47.0	46.6	44.84	517.3307 (+)	C34H44O4	***Geranyl stilbenoid***
28	49.8	49.4	48.05	517.3307 (+)	C34H44O4	Geranyl benzophenone

Stilbenoids are bolded and italicised.

[a] The compound was putatively identified by analysing the MS/MS spectra.

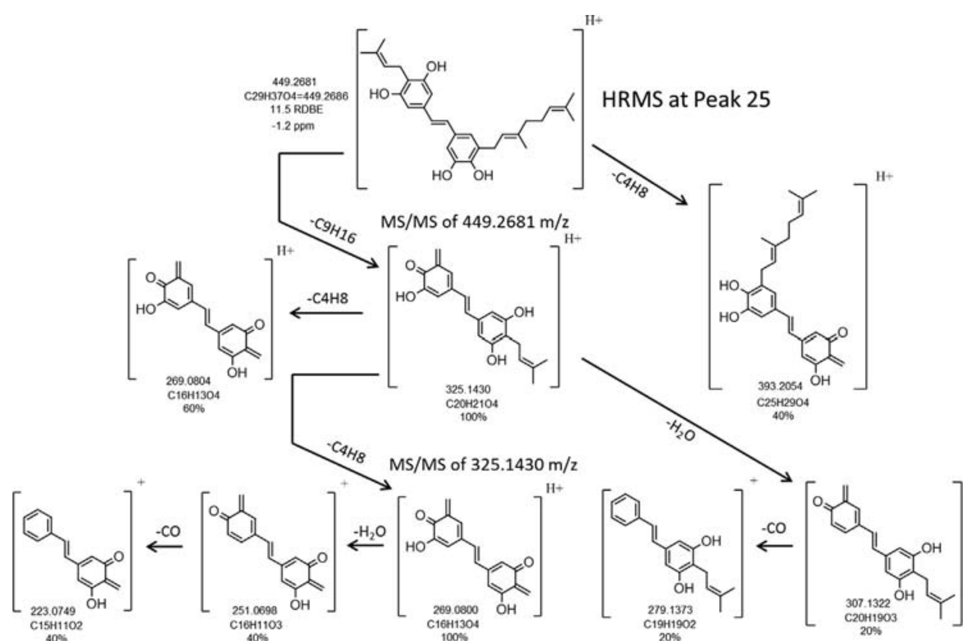

Fig. 7. The proposed fragmentation path way of C29H37O4 on assumption of the structure of mappain.

River State which is the most southern state in Nigeria. This information further supports the presence of stilbenoids in this sample. More attractively according to the predicted formulas some stilbenoid compounds found in this study could be new at least to propolis. As can be observed with their ELSD responses (Fig. 6) they are relatively abundant components in the crude extract and their separation with the current eluting method means that a quick targeted purification might be realised by

using automatic preparative LC with a UV detector by using an optimised eluting method.

4. Conclusion and future work

The African propolis analysed in this study showed high diversity in chemical composition and no clear geographic delineation was found for the classification of these samples. By comparative chemical profiling a propolis sample from South Nigeria stood out as being different and several novel compounds were filtered out through dereplication among the most abundant components in the crude extract. It was proposed that they were stilbenoid compounds from analysing their MS^n spectra. More interestingly the ethanol extract of this sample also presented high biological activity against *Trypanosoma brucei* (data not shown). Future work will be concentrating the labour and the time on purification of these targeted compounds from the crude extract and bioassay and NMR analysis would be performed on these purified compounds. In the meantime more propolis samples from the same region will be collected to confirm the consistency of its chemical composition. This study demonstrated a successful example of utilising chemistry based strategy to discover novel compounds in a large number of samples at an early stage. HRMS based hyphenated analytical techniques offered the measurement of accurate mass and MS^n for individual components in complicated crude extracts of natural products. By searching databases and previously published data a quick dereplication could be achieved. From our experience in this study, however DNP is too comprehensive for an accurate dereplication if only searching by accurate mass. Therefore it is necessary to establish a more specific database only for the natural products isolated from propolis including MS^n data for additional structural confirmation.

Acknowledgements

We acknowledge Scottish Life Sciences Alliance for support for mass spectrometry equipment and BeeVital for providing the propolis samples.

Appendix. Supplementary materials

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.talanta.2013.11.094.

References

[1] T. Michel, M. Halabalaki, A.L. Skaltsounis, Planta Med. 79 (2013) 514–532.
[2] F.E. Koehn, G.T. Carter, Nat. Rev. Drug Discov. 4 (2005) 206–220.
[3] F.E. Koehn, Prog. Drug. Res. 65 (175) (2008) 177–210.
[4] A.M. Gomez-Caravaca, M. Gomez-Romero, D. Arraez-Roman, A. Segura-Carretero, A. Fernandez-Gutierrez, J. Pharm. Biomed. Anal. 41 (2006) 1220–1234.
[5] W.F. Smyth, T.J.P. Smyth, V.N. Ramachandran, F. O'Donnell, P. Brooks, Trac-Trends Anal. Chem. 33 (2012) 46–54.
[6] K.T. Johansen, S.G. Wubshet, N.T. Nyberg, Anal. Chem. 85 (2013) 3183–3189.
[7] J.M. Sforcin, V. Bankova, J. Ethnopharmacol., 133, 253–260.
[8] V.S. Bankova, S.L. de Castro, M.C. Marcucci, Apidologie 31 (2000) 3–15.
[9] A. Salatino, C.C. Fernandes-Silva, A.A. Righi, M.L.F. Salatino, Nat. Product Rep. 28 (2011) 925–936.
[10] A. Salatino, E.W. Teixeira, G. Negri, D. Message, Evid.-Based Complement. Altern. Med. 2 (2005) 33–38.
[11] D.G. Watson, E. Peyfoon, L. Zheng, D. Lu, V. Seidel, et al., Phytochem. Anal. 17 (2006) 323–331.
[12] V. Seidel, E. Peyfoon, D.G. Watson, J. Fearnley, Phytother. Res. 22 (2008) 1256–1263.
[13] A.L. Piccinelli, C. Lotti, L. Campone, O. Cuesta-Rubio, M.C. Fernandez, et al., J. Agric. Food Chem. 59 (2011) 6484–6491.
[14] M.P. Popova, K. Graikou, I. Chinou, V.S. Bankova, J. Agric. Food Chem. 58 (2010) 3167–3176.
[15] A.A. Righi, G. Negri, A. Salatino, Evid.-Based Complement. Altern. Med. 2013 (2013) 14. (Article ID 267878).
[16] C.C. Fernandes-Silva, A. Salatino, M.L.F. Salatino, E.D.H. Breyer, G. Negri, Quimica Nova 36 (2013) 237–240.
[17] C.A. Nunes, M.C. Guerreiro, J. Sci. Food Agric. 92 (2012) 433–438.
[18] S.I. Falcao, N. Vale, P. Gomes, M.R. Domingues, C. Freire, et al., Phytochem. Anal. 24 (2012) 309–318.
[19] F. Pellati, G. Orlandini, D. Pinetti, S. Benvenuti, Journal of Pharm. Biomed. Anal. 55 (2011) 934–948.
[20] A. Aliboni, A. D'Andrea, P. Massanisso, J. Agric. Food Chem. 59 (2011) 282–288.
[21] O. Cuesta-Rubio, A.L. Piccinelli, M.C. Fernandez, I.M. Hernandez, A. Rosado, et al., J. Agric. Food Chem. 55 (2007) 7502–7509.
[22] C.O. Frozza, C.S. Garcia, G. Gambato, M.D. de Souza, M. Salvador, et al., Food Chem. Toxicol. 52 (2013) 137–142.
[23] A. Petrova, M. Popova, C. Kuzmanova, I. Tsvetkova, H. Naydenski, et al., Fitoterapia 81 (2010) 509–514.
[24] A.C. Sawaya, D.M. Tomazela, I.B. Cunha, V.S. Bankova, M.C. Marcucci, et al., Analyst 129 (2004) 739–744.
[25] T. Zhang, D.J. Creek, M.P. Barrett, G. Blackburn, D.G. Watson, Anal. Chem. 84 (2012) 1994–2001.
[26] T. Pluskal, S. Castillo, A. Villar-Briones, M. Oresic, Bmc Bioinf. (2010) 11.
[27] T. Zhang, D.G. Watson, L.J. Wang, M. Abbas, L. Murdoch, et al., Plos One (2013) 8.
[28] S.I. Falcao, M. Vilas-Boas, L.M. Estevinho, C. Barros, M.R. Domingues, et al., Anal. Bioanal. Chem. 396 (2010) 887–897.
[29] R. Simons, J.P. Vincken, E.J. Bakx, M.A. Verbruggen, H. Gruppen, Rapid Commun. Mass Spectrom. 23 (2009) 3083–3093.
[30] A.L. Piccinelli, L. Campone, F. Dal Piaz, O. Cuesta-Rubio, L. Rastrelli, J. Am. Soc. Mass Spectrom. 20 (2009) 1688–1698.
[31] Y. Zhang, P. Zhang, Y. Cheng, J. Mass Spectrom. 43 (2008) 1421–1431.
[32] L. Zhang, L. Xu, S.S. Xiao, Q.F. Liao, Q. Li, et al., J. Pharm. Biomed. Anal. 44 (2007) 1019–1028.
[33] V.F.D. Ishida, G. Negri, A. Salatino, M.F.C.L. Bandeir, Food Chem. 125 (2011) 966–972.
[34] A. Abu-Mellal, N. Koolaji, R.K. Duke, V.H. Tran, C.C. Duke, Phytochemistry 77 (2012) 251–259.