Novel colon-available triterpenoids identified in raspberry fruits exhibit antigenotoxic activities in vitro


Published in:
Molecular Nutrition & Food Research

Document Version:
Peer reviewed version

Queen's University Belfast - Research Portal:
Link to publication record in Queen's University Belfast Research Portal

Publisher rights
© 2016 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim. This work is made available online in accordance with the publisher's policies. Please refer to any applicable terms of use of the publisher.

General rights
Copyright for the publications made accessible via the Queen's University Belfast Research Portal is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
The Research Portal is Queen's institutional repository that provides access to Queen's research output. Every effort has been made to ensure that content in the Research Portal does not infringe any person's rights, or applicable UK laws. If you discover content in the Research Portal that you believe breaches copyright or violates any law, please contact openaccess@qub.ac.uk.

Download date: 11. Mar. 2019
Novel colon-available triterpenoids identified in raspberry fruits exhibit anti-
genotoxic activities in vitro.

Gordon J. McDougall*, J. William Allwood¹, Gema Pereira-Caro², Emma M. Brown³, Susan Verrall¹, Derek Stewart¹, Derek Stewart¹, Derek Stewart¹, Cheryl Latimer³, Geoff McMullan³, Roger Lawther⁵, Gloria O'Connor⁵, Ian Rowland⁶, Alan Crozier⁷ and Chris I. R. Gill³

¹Environmental and Biochemical Sciences Group, Enhancing Crop Productivity and Utilisation Theme, The James Hutton Institute, Invergowrie, Dundee DD2 5DA, Scotland
²Postharvest, Technology and Agrifood Industry Area, IFAPA, Córdoba, Spain
³Northern Ireland Centre for Food and Health, Centre for Molecular Biosciences, University of Ulster, Cromore Road, Coleraine, N. Ireland, UK
⁴NIBIO, Norsk Institut for Bioøkonomi, Pb 115, NO-1431 Ås,
⁵Altnagelvin Area Hospital, Western Health and Social Care Trust, Glenshane Road, Londonderry
⁶Hugh Sinclair Unit of Human Nutrition, Department of Food and Nutritional Sciences, University of Reading, P.O. Box 226, Whiteknights, Reading, UK
⁷Department of Nutrition at the University of California, Davis, California 95616, USA

Correspondence: Gordon J. McDougall
Email: Gordon.mcdougall@hutton.ac.uk
FAX: +44 (0)844 928 5429; TEL: +44 (0)1382 568782

Abbreviations: CDDO = 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid; FAV = fruit and vegetables; GI = gastrointestinal; TRF = triterpenoid-rich fraction;

Key words: Raspberries, metabolic profiling, triterpenoids, colon cancer, ileal fluid
ABSTRACT

Scope: Ileostomy studies provide a unique insight into digestion of food, allowing identification of physiologically relevant dietary phytochemicals and their metabolites important to gut health. We previously reported the consistent increase of components in ileal fluids of ileostomates after consumption of raspberries with use of non-targeted LC-MS\textsuperscript{n} techniques and data deconvolution software highlighting two major unknown components (m/z 355 and 679).

Methods and results: In-depth LC-MS\textsuperscript{n} analyses suggested that the ileal m/z 355 components were p-coumaroyl glucarates. These compounds have not been identified previously and were confirmed in raspberry extracts after partial purification. The major ileal component with m/z 679 was a glycoside with an aglycone of m/z 517 and was present as two peaks in extracts of whole puree, unseeded puree and isolated seeds. These components were purified using Sephadex LH20 and C18 SPE units and identified as major, novel raspberry triterpenoid glycosides. This triterpenoid-enriched fraction (100 nM) protected against H\textsubscript{2}O\textsubscript{2}-induced DNA damage in both colon cancer and normal cell lines and altered expression of cytoprotective genes.

Conclusion: The presence of these novel raspberry triterpenoid components in ileal fluids indicates that they would be colon-available in vivo, so confirmation of their anti-cancer bioactivities is of key physiological relevance.
1 INTRODUCTION

There is great interest in the possibility that non-nutritive components could directly contribute to the health benefits attributable to a diet rich in fruit and vegetables (FAV) [1-3].

One set of phytochemicals of specific interest are the (poly)phenols as are found in high concentrations in certain beverages and FAV, such as berries, and have been attributed a range of important bioactivities [4-9]. However, an important issue with ascribing any potential health benefit is the long held view that the bioavailability of many (poly)phenol classes is low. For example, urinary recoveries of anthocyanins, which reflect passage through the circulatory system, are typically <<1% of intake [10], although recent studies with \[^{13}\text{C}]\text{cyanidin-3-O-glucoside}\ [11, 12] and raspberries [13], which took into account metabolites and the parent anthocyanins, established that cyanidin-based anthocyanins are much more bioavailable than previously envisaged. Likewise with orange juice flavanones which have recently been shown to be much more bioavailable than previously envisaged [14, 15].

It is becoming increasingly evident that (poly)phenols may also contribute to health benefits through interactions within the gastrointestinal (GI) tract, including protection against epithelial damage associated with GI cancers [16, 17], maintenance of a beneficial microbiota and its associated biotransformation of (poly)phenolic metabolites [18] and inhibition of digestive enzymes that can slow nutrient release with benefits for glycemic control and/or obesity [19-21].

Investigations that simulate digestive processes in vitro [e.g. 22] have provided evidence that berry (poly)phenols have different stabilities in the GI tract and may be available in amounts that could beneficially influence important physiological processes. Although these studies can define and compare potential in-gut availability, they cannot adequately mimic the dynamic, active processes of digestion [23]. Ileostomy studies provide a unique insight and can identify phytochemicals and their metabolites which, in volunteers with an intact colon, would pass from the small to the large intestine [24-26].
Although previous targeted liquid chromatography mass spectrometry (LC-MS^n) analyses [27] confirmed that the major (poly)phenol components of raspberries, anthocyanins and ellagitannins, were present in ileal fluid after raspberry supplementation [26], non-targeted LC-MS^n analyses selected other components that consistently increased in abundance after raspberry supplementation including some previously unreported components [27]. In this study, we sought evidence for the nature of two of the major novel components and extracted and identified these components from raspberries. Through their survival in ileal fluids, these novel components are confirmed as colon-available so we also explored the potential bioactivity of the novel triterpenoids identified using cellular models relevant to colon cancer and discuss the physiological impact of survival in the GI tract of these bioactive components.

2 MATERIALS AND METHODS

2.1 Chemicals

Chemicals were obtained as described previously [27]. Tenuifolin was purchased from Stanford Chemicals Ltd (Irvine CA, USA). 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid (CDDO) was purchased from (Cambridge Bioscience Ltd, UK).

2.2 Plant material and processing

Raspberries (Rubus idaeus var. Glen Ample), grown and prepared at the James Hutton Research Institute in 2014 were stored at −80°C. A 8 kg aliquot was defrosted, pureed [7], frozen and transported to University of Ulster for the ileostomy feeding studies.

2.3 Ileostomy feeding study
The ileal fluid samples were collected from the raspberry puree ileostomy feeding study (Ref No. 11/NI/0112) described in full previously [27]. In brief, following a diet low in (poly)phenolic compounds, 11 ileostomates provided a baseline ileal fluid sample (T = 0 h) then consumed 300 g of pureed raspberries and a second ileal fluid sample collected at T= 8 h. The ileal fluid samples were collected, processed within 30 min and stored as aliquots at −80°C.

2.4 Non-targeted LC-MS\textsuperscript{n} analysis

LC-MS\textsuperscript{n} analysis of ileal samples was performed on an HPLC system consisting of an Accella 600 quaternary pump, Accella PDA detector coupled to an LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific, San Jose, CA, USA) operated under Xcalibur software [27]. The same programme was used to analyse samples from the raspberries. Certain samples (e.g. seed and puree extracts) were also analyzed using an LCQ-DECA system with an ion trap mass spectrometer [27].

2.5 MS data handling and analysis

The non-targeted LC-MS data from the Orbitrap analysis was first deconvolved using the SIEVETM software programme which produces a list of retention time (RT)-m/z pairs along with an associated extracted ion chromatogram based on peak area for each sample [28]. The data were analysed statistically [27] and a subset was defined of components whose patterns of abundance were increased after supplementation in all 11 subjects.

2.6 Raspberry puree and seed extractions

Frozen raspberries (~500g) were thawed on ice then pureed in a Waring blender (top speed, 3 x 15 s). A portion of the puree was sieved to remove seeds (0.5 mm size) and the
unseeded puree collected. The seeds were washed with ice-cold distilled water then dried on paper towels. Seeded and unseeded purees (5 mL) and seeds (1 g fresh weight) were extracted with 10 mL of 0.1 % aqueous formic acid for 30 min at 5 ºC with end-over-end mixing in a blood rotator then centrifuged (2 500g, 5 min 5 ºC) and the extracts removed to new tubes. The extraction was repeated with 0.1% aqueous formic acid then sequential extractions were carried with 80% acetonitrile containing 0.1% formic acid and finally with 10 mL 50 % aqueous acetone. Aliquots,1mL, of the extracts were dried in a Speed-Vac, re-suspended in 5% aqueous acetonitrile containing 0.2 % formic acid before LC-MSn analysis.

2.7 Fractionation

Raspberry puree (15 mL) was diluted in an equal volume of 0.2 M HCl, vortexed well and incubated in a blood rotator (100 rpm) for 10 min at 5 ºC. After centrifugation (2500 g, 10 min, 5 ºC), the supernatant was used for fractionation. Strata X-C solid phase extraction (SPE) units (100 mg/3 mL units; Phenomenex Ltd, Macclesfield, U.K.) were pre-equilibrated with 10 mL methanol then washed with 10 mL water. Extracts (5 mL) were applied to X-cartridges and unbound material was collected. After a wash with 2 x 5 mL of water, the unbound eluates were combined. The units were eluted with 5 mL methanol which was collected. The cartridges were still red as anthocyanins remained bound. The fractions were tested for total phenol content by the Folin method [29] and only the methanol fraction contained appreciable phenolics. Aliquots (1 mL) were dried, re-suspended as above prior to LC-MSn analysis.

2.8 Bulk seed extraction and purification

Eight kg of raspberries were pureed in 250 g batches and the seeds separated (seed yield ~4% w/w). Seeds (250 g) were extracted with 1 L of 0.1% aqueous formic acid for 60 min at 5 ºC with orbital rotation at 90 rpm and the extract obtained by filtering through a glass sinter
(porosity 3). The seeds were then extracted twice with 500 mL of 80% ethanol and finally with 500 mL of 50% aqueous acetone. These fractions were assayed for total phenol content and aliquots (1 mL) dried and re-suspended as above for LC-MS\textsuperscript{n} analysis. The aqueous extract and the first ethanol extract were combined and diluted to 10 % aqueous ethanol then applied to a 70 mL column of Sephadex LH20 (GE Healthcare, Buckinghamshire, UK) which had been equilibrated with 80 % aqueous acetone then 10 % aqueous ethanol. The unbound fraction was collected and the column was washed with 2 column volumes of 10 % ethanol before being eluted with 80% ethanol then 50 % acetone. Aliquots of each fraction were dried using the Speed-Vac prior to LC-MS analysis. The bulk of the unbound and wash fractions were combined and the ethanol removed by rotary evaporation. The fraction was made up to 0.1 % aqueous formic acid and applied to a C18 solid phase extraction units (Strata C18-E, GIGA units, 10 g capacity; Phenomenex, Ltd., Macclesfield, U.K.) that had been treated with 80 % aqueous acetonitrile containing 0.1% formic acid then 0.1% aqueous formic acid. The unbound fraction was collected with a wash fraction of 75 mL 0.1% aqueous formic acid. Components were eluted with 75 mL of 15 % aqueous acetonitrile containing 0.1% formic acid followed by 75 ml 80 % aqueous acetonitrile containing 0.1% formic acid. Fractions were assayed for total phenol content and aliquots dried and re-suspended for LC-MS analysis. Enrichment in triterpenoids was followed using the red-brown colour reaction of the Lieberman-Burchard method [31]. The content of triterpenoids in the final fraction was estimated as tenuifolin equivalents by peak areas.

2.9 Tissue culture

Human colon cells HT29 (adenocarcinoma) and CCD 841 CoN (normal epithelial) (32) were acquired from European Collection of Cell Cultures (ECACC) and American Type Culture Collection (ATCC) respectively. HT29 cells were cultured in DMEM supplemented with 10% FBS and 100U/L penicillin/streptomycin. CCD 841 CoN cells were maintained in MEM supplemented with 10% FBS, 100U/L penicillin/streptomycin, 1% sodium pyruvate, 1%
Both cell lines were incubated at 37°C with 5% CO₂ and grown as monolayers in roux flasks. Cells were sub-cultured every 3-4 days by the addition of trypsin (0.25% trypsin-EDTA) at 37°C for 5 min. Cells were centrifuged at 1200 rpm for 3 min, the supernatant decanted and cells re-suspended in the appropriate medium.

For the purposes of this study both cell lines were treated with either a pure synthetic triterpenoid 2-cyano-3,12-dioxoolean-1,9-dien-28-oic acid (CDDO), a known inducer of the Nrf2/ARE pathway (33, 34), or a triterpenoid-rich fraction (TRF). Both treatments were prepared fresh in the respective cell media to a final concentration of 100 nM for use in the in vitro experiments.

2.10 Cytotoxicity and genotoxicity of triterpenoids

2.10.1 Cytotoxicity assay. The effects of CDDO and TRF on the viability of HT29 and CCD841 cells were determined using the MTT assay (35). The assay is based on the ability of living cells to metabolize the tetrazolium salt, MTT (3-(4,5-dimethylthaizaol-2-yl)-2,5-diphenyl tetrazolium bromide, Promega, Madison, USA) by mitochondrial activity to formazan, a blue dye (36) which can be measured spectrophotometrically. Cells were seeded in 96 multi-well plates (Costar, Cambridge, MA, USA) at a concentration of 1.5 x 10⁴ HT29 and 3.0 x 10⁴ CCD 841 cells per well respectively. After 2 days incubation at 37°C, media was replaced with 100 nM of either TRF or CDDO, then incubated for 24 h. The wells were washed and cells incubated for a further 48 h in fresh media. Thereafter 15 μL of MTT were added to each well. After 4 h, lysis was carried out with 100 μL solubilizing solution to free the product formazan. Formazan was measured using a microtiter plate reader (Alpha, SLT Rainbow Thermo, Antrim, UK) at a wavelength of 560 nm. The survival of the cells
treated with cell media only was set as 100% viability. Each treatment was performed in octuple and the experiment was carried out on 3 separate occasions.

2.10.2 COMET assay. The assay, as described previously [37] using the well-established HT29 cell model for colonic DNA damage (17) and the normal colonocyte CCD841. In brief, both cell lines were incubated for 24 h with 100 nM of either TRF or CDDO. To assess the anti-genotoxic potential of the treatments, the cells were treated with hydrogen peroxide (75μM, H₂O₂ for HT29 and 25 μM, H₂O₂ for CCD 841) for 5 min at 4°C, then centrifuged for 5 min at 258 x g. The supernatant was discarded and the cell pellet re-suspended in 85 μL of 0.85% low melting point agarose (LMPA) in PBS and maintained in a water bath at 40°C. The suspension was added to previously prepared gels (1% normal agarose) on frosted slides and coverslips were added. The slides were immersed in lysis buffer (2.5 M NaCl, 100 mM Na₂EDTA, 10mM TRIS, pH10) for 1 h at 4°C and then placed in electrophoresis buffer and allowed to unwind for 20 min before running at 26 V (300 mA) for 20 min. After electrophoresis, gels were washed 3 times for five min in neutralisation buffer (0.4 M Tris, pH 7.5) at 4°C. All gels were stained with 20 μL of ethidium bromide (2 μg/mL in double distilled water prior to scoring. Images were analysed at 400 times magnification using a Nikon eclipse 600 epi-fluorescence microscope. The percentage DNA tail was recorded using Komet 5.0 image analysis software (Kinetic Imaging Ltd, Liverpool, UK). For each slide, 50 cells were scored. Data represents the mean percent tail DNA of triplicate gels per treatment from three independent experiments. To assess the genotoxic potential of the compounds, cells were treated as above omitting the H₂O₂ challenge. Positive (H₂O₂) and negative controls (PBS) were included in all experiments (cells without CDDO/TRF pre-treatment).

2.11 RNA isolation and cDNA synthesis
Cell pellets were collected from both treated HT29 and CCD 841 CoN cells and homogenised using QIAshredder (Qiagen) prior to extraction of total RNA with RNeasy Mini Plus Kit (Qiagen) as per the manufacturer’s instructions. Standard PCR of isolated RNA confirmed the absence of genomic contamination. Quantification of RNA was determined via NanoDrop™ ND-100 UV/VIS spectrophotometer with quality verified by gel electrophoresis. Only RNA samples with $A_{260}/A_{280}$ and $A_{260}/A_{230}$ between 1.8–2.0 were deemed acceptable for downstream applications.

cDNA was produced using anchored-oligo(dT)$_{18}$ and Transcriptor First Strand cDNA synthesis Kit (Roche). Using 1 µg of total RNA, 20 µL of cDNA was synthesised as per the manufacturer’s instructions for reverse transcription. Standard PCR using 1 µL cDNA as a template and the housekeeping gene GAPDH (primers shown in Table S1) as a control, confirmed the absence of unspecific products.

### 2.12 Real-time PCR (qPCR)

**Real-time** qPCR was performed using the Lightcycler 480 II (Roche) in accordance with the manufacturer’s instructions. Each 10 µL PCR reaction contained 0.5 µM each primer, 5 µL LightCycler 480 SYBR Green I Master (Roche), 2 µL nuclease-free water and 1 µL cDNA template. Cycling conditions were as follows: 95°C for 10 min, 50 cycles of 95°C for 10 s, 57°C for 10 s and 72°C for 10 s. Melt curve analysis for each gene confirmed product specificity, and only artefact free reactions were considered valid. For all experiments negative control reactions (no template control and negative reverse transcriptase) were run on the same plate.

cDNA standards (in triplicate) were produced by the 5-fold dilution of pooled cDNA, and used to produce cDNA calibration curve slopes. LightCycler software (version 1.5) generated primer efficiencies for each gene (primers shown in Table 1), calculated by using the equation $E=10^{[\text{slope}]}$ [38]. cDNA templates used for target runs were diluted to within the
standard curve linear range, in this case 1:10. Only primers with an efficiency of 90-100% were used for target analysis (Table S1). Relative expression is calculated using the following equation:

\[
\frac{\text{Conc. target gene}}{\text{Conc. reference genes}} \text{(control)} : \frac{\text{Conc. target gene}}{\text{Conc. reference genes}} \text{(sample)}
\]

Each cDNA target sample was normalised to 3 reference genes, (HPRT, β-Actin, GAPDH) and calculated as a ratio of the untreated control samples. All target cDNA samples were performed as technical triplicates, with biological replicates for each sample.

2.13 Statistical Analysis

The mean of each data set was used for statistical analysis and experiments were carried out as independent triplicates. The Shapiro-Wilk test was used to test for normality. Analysis of variance was applied to test for significant differences between means compared to control using Dunnett T post hoc test. Significance was accepted at p < 0.05. Analysis was carried out using SPSS (version 20 for Windows).

3 Results

Two major unknown compounds (m/z 355 and 679) were consistently identified in ileal fluids from the 11 volunteers [27] after raspberry supplementation (Fig. 1). The compound at m/z 679 was present at a similar intensity to the major ellagitannin peak, sanguiin H-6 (compare Figs. 1A and 1B). The compound at m/z 355 (Fig. 1C) was present in three separate peaks (see Supplementary data; Fig. S1). Initial [M-H]- MS data (Table 1) [m/z = 355.0648 and MS² fragments of 337, 209 and 191] did not match with any berry component or database entry [27]. However, positive mode data gave m/z M+H = 357.0809, predicted formula of C_{15}H_{17}O_{10} with MS² of m/z 339 (loss of H₂O) and 147 (loss of 210). In-source fragments mirrored the MS² fragments and gave predicted formulae for m/z M+H⁺ = 339 of C_{15}H_{15}O_{9}
and \( m/z \) M+H = 147 of C₉H₇O₂. The neutral loss of 146 amu in negative mode and the MS² fragment at 147 in positive mode suggests a \( p \)-coumaric acid derivative. Indeed, the peak had an absorbance maximum at \( \sim \)310 nm, characteristic of hydroxycinnamates (results not shown). The neutral loss of 210 in positive mode and the MS² fragment at \( m/z \) 209 in negative mode can be assigned to glucaric acid which strongly suggests that these components are \( p \)-coumaroyl glucarate isomers (see diagram 1, structure 1)) as described in dog’s mercury [39]. These compounds have not been previously identified in raspberry but caffeoyl glucarates have been identified in calafate berries [40]. \( p \)-Coumaroyl glucarate components were confirmed in extracts of Glen Ample raspberries (see Supplementary data; Fig. S2A & B) and were enriched by fractionation on strong ion-exchange SPE cartridges which retained anthocyanins and ellagitannins but released the coumaroyl glucarates (Supplementary data; Fig. S2C & D).

The \( m/z \) 679 signal gave no UV absorbance above 240 nm and so was unlikely to be a (poly)phenolic compound. Exact mass of 679.3648 yielded a predicted formula of C₃₆H₅₅O₁₂ (Table 1) and the major MS² fragment at 517 (loss of 162) suggested a glycoside with an aglycone of 517. However, the \( m/z \) 679 signal had 0.5 amu variants (Table 1) which suggests a doubly charged entity with a true mass of \( \sim \)1360. A signal at \( m/z \) at 1359.7403 was present (predicted formula C₇₂H₁₁₁O₂₄). Positive mode data confirmed these results [exact mass 681.3833; predicted formula of C₃₆H₅₇O₁₂; exact mass 1361.7603, predicted formula of C₇₂H₁₁₃O₂₄]. However, no MS² of the M+H 1361.7 or M-H 1359.7 were obtained. Compounds which were good matches for the predicted formula obtained for [M-H]- \( m/z \) 679 (MW C₃₆H₅₆O₁₂) and were glycosides were all isomers of ursolic acid-based triterpenoid glucosides that differed in the position of attachment of glucose, hydroxyl or methyl groups etc. (Table 1; Diagram 1; structure 2). Triterpenoids have been identified in raspberry leaves and flowers, including in commercial varieties of Rubus idaeus [41]. Ursolic acid based triterpenoid glycosides have been noted in leaves of R. coreanus [e.g. suavissimoside R1 [42], whereas trachelosperoside B1 (MW = 682) and nigaichigoside F1 (MW = 666) have
also been identified in *Rubus rosifolius* fruits [43]. Ilexoside XLV has been reported to give a main MS² fragment at 517 [44] from loss of hexose but any of the isomers reported in Table 1 could give the same fragmentation.

Two separate peaks (T1 and T2) with apparent *m/z* 679 were apparent in methanol extracts of seeds and were major contributors to the total MS signal (Fig. 2). Only peak T1 was detected in the ileal fluids. Sequential extractions of raspberry puree, isolated seeds and "unseeded" puree (see supplementary material Fig. S3) showed that Peaks T1 and T2 were present in both whole and unseeded purees but with higher levels in the whole puree (Fig. 3) and seeds. Peak T2 was more abundant in the acetonitrile and acetone extracts than water extracts suggesting that it was more hydrophobic than peak T1. Therefore, although the MS data suggests that peaks T1 and T2 may be related to the ursolic acid triterpenoid glycosides noted in Table 1, the nature of these components was examined further after purification.

Fractionation of seed extracts removed polyphenols and enriched peaks T1 and T2 (Fig. 4). This enrichment was accompanied by reduced total phenol content, enhanced response to the Liebermann–Burchard reaction for triterpenoids (results not shown) and an enrichment of other putative triterpenoid peaks (Fig. 4, Table 2; peaks 13, 14, A & B). Indeed, alkaline hydrolysis of the initial seed extract produced *m/z* signals from simple phenolics from degradation of the polyphenols and a range of triterpenoid aglycones consistent with peaks T1 and T2 (i.e. *m/z* 517) but also other putative triterpenoids (see supplementary data Fig. S4, Table S2 for data and method).

The MS properties of peak T1 were similar to those of the *m/z* 679 signal in ileal samples having 0.5 amu variants suggesting doubly charged status and an actual mass of ~1360 (Fig. 5; Table 2). The signal at *m/z* 1359.7643 (predicted formula C_{72}H_{111}O_{24}) gave a single MS² fragment at 679.3. Therefore, Peak T1 may be similar to the ester-linked dimeric triterpenoid, Coreanoside F1 (C_{72}H_{110}O_{24}, MW = 1358.7; Diagram 1; structure 3) extracted from *Rubus coreanus* leaves [45], but this compound has an MW 2 amu less than the
apparent MW for peak T1. This “extra 2H” could occur by (e.g.) substitution with –CH₂OH
and –COOH groups at one position (+14 amu) and –OH and -H groups at another position (-
16; net difference +2 amu).

Peak T2 had a major m/z at 1357.7 (Fig. 5; Table 2), by analogy with m/z 1359 in peak
T1, this could result from a component with C₇₂H₁₀₉O₂₄ but this was not predicted from the
accurate mass data. The signal at m/z 1357.7 gave no MS² but the signal at 1358.7 yielded
a major MS² fragment at 679.4 (loss of 679) with fragments at 1313.4 (loss of 46); 1196.4
(loss of 162), 1151.7 (loss of 162 & 46) 1018.7 (loss of 340) and also 559.7 and 517.3.
Considering that peak T2 also showed 0.5 amu variants suggesting doubly charged status
and an nominal actual mass of ~ 2716, the MS data strongly suggests that peak T1 and T2
are structurally related, and that peak T2 could be a dimer of peak T1. Initial ¹H NMR spectra
were consistent with peak T1 being a ursolic acid based triterpenoid but further NMR studies
and use of alternative MALDI-TOF MS techniques with higher m/z ranges [46] will be
required to confirm the nature of these putative triterpenoid derivatives.

After purification, the triterpenoid-rich fraction (TRF) was effective in preventing H₂O₂-
induced DNA damage to HT29 adenocarcinoma cells and also to the normal epithelial colon
cell line CCD 841 CoN (Fig. 5). At 100 nM, both TRF and the synthetic triterpenoid 2-cyano-
3,12-dioxoolean-1,9-dien-28-oic acid (CDDO) exerted a significant anti-genotoxic effect
against H₂O₂ challenge in the two cell lines (Fig. 5). In HT 29 cells, both treatments reduced
tail DNA by ~40-45 % compared to the untreated control, whilst in CCD 841 CoN cells, DNA
damage was reduced by 50-55%. Efficacy between cell lines was not significantly different.
No cytotoxic or genotoxic activity was observed for either TRF or CDDO at 100 nM (data not
shown).

CDDO has been reported to induce the Nrf2/ARE pathway including NAD(P)H
dehydrogenase, quinone-1 (NQO1) and heme oxygenase-1 (HO-1) [33]. These enzymes
reduce reactive oxygen species and play a key role in cytoprotection; NQO1 acts a
superoxide scavenger [47] and HO-1 reduces the production of free radicals via the
catabolism of heme [48]. After 24 h exposure, CDDO significantly increased expression of both NQO-1 and HO-1 in HT29 and CCD 841 CoN cells (Fig. 6), but Nrf2 expression was significantly reduced. The 24 h time-frame was chosen to match with the genoprotective studies but a time-course experiment (see Fig S5) showed that CDDO enhanced Nrf2 expression in both cell lines over shorter exposure times.

TRF increased the expression of the Nrf2 gene in HT29 cells but reduced expression in normal CCD 841 CoN cells (Fig. 6) following 24hr exposure. TRF showed a small, but significant, increase in NQO1 expression in CCD 841 cells but HO-1 expression was significantly reduced, in both HT29 and CCD841 cells. This differential effect on activation of the Nrf2-regulated pathway may be caused by the concentration difference between the pure CDDO and TRF (which is a mixture of triterpenoids), or it may reflect structural differences between the compounds. Indeed, the position of methyl groups in the triterpenic dialcohols, uvaol and erythrodiol, from olive oil altered bioactivity from genoprotective to genotoxic in both normal and breast cancer cell lines (MCF10A & MDA-MB-231) [49]. Alternatively, it could represent a time-dependent effect on up-regulation of Nrf2 expression in both cells as suggested by the results with CDDO (Figure S5).

4 Discussion

The application of non-targeted LC-MSn analysis to ileal fluids after raspberry intake selected previously unknown components that arose from the berry intake. Further extraction and purification work confirmed their presence in raspberries and allowed their identification as p-coumaroyl glucarates and ursolic acid-based triterpenoid glucosides. Their relative stability to GIT conditions explains their revelation in ileal fluid and confirms their bioavailability throughout the GIT and into the colon. However, the non-targeted nature of the LC-MS based approach and experimental robustness of the data from the ileostomy cohort was essential in bringing these previously unknown components to light.
The presence of \( p \)-coumaroyl glucarate derivatives in raspberries, and their survival in ileal fluids, could be of significance as D-glucaric acid has long been known to have anti-cancer effects in carcinogen-induced animal models [50]. Although originally thought to act through inhibition of \( \beta \)-glucuronidase and increased detoxification of carcinogens, it may have more complex effects via anti-inflammatory systems [51] and apoptotic events [52]. As these coumaroyl glucarate esters survive into the ileal fluid, they could deliver potentially-bioactive glucaric acid throughout the GI tract, ultimately to the colon, where it could have a role in chemoprevention of cancer. Future work will purify and assess the potential bioactivity of these colon-available raspberry components.

The fraction from raspberries enriched in the novel triterpenoids was found to be effective in protecting against H\(_2\)O\(_2\)-induced DNA damage at sub-micromolar levels, which could be readily achieved and therefore physiologically relevant in vivo. Indeed, the parent triterpenoid ursolic acid decreased H\(_2\)O\(_2\)-induced DNA damage in the colonic cell line Caco-2 by a similar extent (>40\%) to that of TRF and CDDO, albeit at 50-fold higher concentration of 5 \( \mu \)M [53]. Triterpenoids from *Rubus rosifolius* fruits have previously been shown to have beneficial effects against human colon cancer cells [42] and triterpenoids from *Scoparia dulcis* roots exerted anti-mutagenic activity using the *in vivo* SMART assay [54]. To our knowledge, this study is the first to demonstrate DNA damage reduction and the modulation of Nrf2/ARE pathway by triterpenoids (both CDDO and TRF) in the normal cell line CCD841-CoN. Although recent work [55] reported that sulforaphane decreased H\(_2\)O\(_2\)-mediated oxidative damage and activated the Nrf2/ARE pathway in CCD841 cells, concentrations of 2.5-40\( \mu \)M were required to elicit this effect. Triterpenoids have been reported to have other relevant bioactivities; oleanolic acid glycosides influence gastrointestinal transit in mice [56]; triterpenoids modulate intestinal transport [57]; purified triterpenoids from *Rubus parvifolius*, including suavissimoside R1 and coreanoside F1, have anti-fatigue effects in mice [58] and triterpenoid-rich fractions from Korean raspberries have potent anti-inflammatory effects [59].
In many cases, these triterpenoids are isolated from non-edible plant parts such as roots or leaves, which would not form part of the normal diet.

It is possible that these raspberry fruit components may have contributed to the positive effects noted for raspberry extracts in our previous studies on colon cancer models [17, 22] and may have contributed to bioactivities assigned to berries in other studies [14]. Indeed, after re-examination of previous data we can confirm that the putative triterpenoids survived in vitro digestion procedures used to simulate gastrointestinal conditions (e.g. [22, 37]).

In conclusion, these novel components were discovered by their survival in ileal fluids, they are available in the gut and we can assume that they would enter the colon in vivo and could exert these bioactivities in situ throughout the GI tract. Further work on the contribution of these novel components to potential health effects of raspberries is merited.

The authors responsibilities were as follows: CG, GMcD, RL, IR and AC were involved in study design and CG, and GO’C in study conduct. GMcD, SV, CL, MI, GP, CG, EMB and JWA were involved in experimental and data analysis. The manuscript was prepared by GMcD, CG, IR, AC, DS, GMcM and JWA.

We would like to thank the volunteers for participating in the study. CG, RL and AC acknowledge funding from the National Processed Raspberry Council. GMcD, SV, DS and JWA acknowledge funding from the Scottish Government’s Rural and Environment Science and Analytical Services (RESAS) Division. DS and GMcD acknowledge funding from BachBerry (Project No. FP7-613793). GPC was supported by a postdoctoral fellowship from IFAPA (Programa Operativo del Fondo Social Europeo 2007–2013 de Andalucía).

The authors declare no conflicts of interest.
5 References


Diagram 1

1. *p*-coumaroyl glucarate

2. Trachelosperoside A1 ([1-O-((2α,3β,5ξ,9ξ)-2,3,19,24-tetrahydroxy-24,28-dioxours-12-en-28-yl)-β-D-glucopyranose]); R = glucosyl.

(http://www.chemspider.com/Chemical-Structure.10273095.html)

Legends

Figure 1. Comparison of abundance of selected MS signals in ileal fluids before and after raspberry intake. Panel A = m/z 934 = sanguiin H-6 at RT 17.9 min; B = m/z 679 at RT 23.5; C = m/z 355 at RT 12.7. Peak areas are in arbitrary MS units.

Figure 2. Mass spectral characteristics of m/z 679 signal in raspberry seed extracts. Panel A = UV trace at 280 nm; B = MS base peak; C = MS spectra of peak T1; D = MS spectra of peak T2. MS spectra were obtained on ion-trap MS which does not detect doubly charged ions. Figures in the top right corners represent the full scale deflection of the MS detector.

Figure 3. Comparative abundance of peaks T1 and T2 in sequential extracts from whole puree (WP), unseeded puree (USP) and seeds. Peak areas are average of three determinations ± SE. Areas are in arbitrary MS units.

Figure 4. LC-MS profiles of fractions from purification scheme. Panel A = 80 % ethanol seed extract; B = LH20 Sephadex -unbound fraction, C = LH20 Sephadex 10 % ethanol wash fraction, D = 80% ACN SPE fraction. Figures in the top right corners represent the full-scale deflection of the MS detector.

Figure 5. MS spectra of LC peaks T1 and T2. Panel A = peak T1, B = peak T2. m/z values in bold are discussed in the text and in Table 2. Figures in the top right corners represent the full-scale deflection of the MS detector.

Figure 6. Anti-genotoxic effect of triterpenoids on H2O2-induced DNA damage. Anti-genotoxic effects of 100 nM CDDO and TRF after 24 hr pre-incubation on DNA damage in HT29 and CCD 841 cells challenged with 75 µM H2O2 and 25 µM H2O2 respectively. Data is presented as mean of 3 independent experiments ± SD compared to the untreated cells as control. One-way ANOVA and Post Hoc test Dunnett’s T * p < 0.05.

Figure 7. Effect of triterpenoids on antioxidant signalling pathway gene expression. Graphs show absolute fold change values (i.e. change in gene expression when compared to normalised untreated cell as control). Data is presented as the mean of 3 individual experiments (n=3) ± SD. Student t-test. * p<0.05,** p<0.01, *** p<0.001, **** p<0.001.
Table 1. Properties of selected unknown components in ileal samples

<table>
<thead>
<tr>
<th>Peak</th>
<th>[M-H(^-)] (m/z)</th>
<th>MS(^2)</th>
<th>Predicted formula</th>
<th>Putative identity</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>m/z 355</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>355.0648</td>
<td>337, <strong>209</strong>, 191</td>
<td><strong>C(<em>{15})H(</em>{15})O(_{10})</strong></td>
<td>Caffeoyl glucuronide (<a href="http://www.hmdb.ca/metabolites/HMDB41705">http://www.hmdb.ca/metabolites/HMDB41705</a>) but MS(^2) data does not fit</td>
</tr>
<tr>
<td>Positive</td>
<td>357.0809*</td>
<td>339, <strong>147</strong></td>
<td><strong>(C(<em>{15})H(</em>{15})O(<em>{9}), C(</em>{9})H(<em>{7})O(</em>{2}))</strong></td>
<td>(\rho)-coumaroyl glucarate [39, 40]</td>
</tr>
<tr>
<td><strong>m/z 679</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Negative</td>
<td>679.3648*</td>
<td>661, 559, 541, <strong>517</strong>, 499, 455, No MS(^2)</td>
<td><strong>C(<em>{36})H(</em>{55})O(_{12})</strong></td>
<td>Components matching with C(<em>{36})H(</em>{56})O(_{12}) formula were all triterpenoid glycoside derivatives based on oleanolic acid structures</td>
</tr>
<tr>
<td></td>
<td>1359.7403</td>
<td></td>
<td><strong>C(<em>{72})H(</em>{111})O(_{24})</strong></td>
<td>Tenuifolin (<a href="http://www.chemspider.com/Chemical-Structure.10205970.html?rid=52e5e4b2-c256-4cea-b530-813b9921295d">http://www.chemspider.com/Chemical-Structure.10205970.html?rid=52e5e4b2-c256-4cea-b530-813b9921295d</a>)</td>
</tr>
<tr>
<td>Positive</td>
<td>681.3833*</td>
<td>No MS(^2)</td>
<td><strong>C(<em>{36})H(</em>{57})O(_{12})</strong></td>
<td>Trachelosperoside A1 (<a href="http://www.chemspider.com/Chemical-Structure.10273095.html">http://www.chemspider.com/Chemical-Structure.10273095.html</a>)</td>
</tr>
<tr>
<td></td>
<td>1361.7603*</td>
<td>No MS(^2)</td>
<td><strong>C(<em>{72})H(</em>{113})O(_{24})</strong></td>
<td>Suavissimoside F1 (<a href="http://www.chemspider.com/Chemical-Structure.10251895.html">http://www.chemspider.com/Chemical-Structure.10251895.html</a>)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Ilexoside XLV [44]</td>
</tr>
</tbody>
</table>

\(\Delta\) ppm < 2 for all formulae. MS\(^2\) fragments in bold italics = most abundant signals. *-[M-H\(^-\)]\(^2\) ion, has 0.5 amu variants, + = positive mode data, \(b\)formula in brackets are for the MS\(^2\) fragments.
<table>
<thead>
<tr>
<th>Peak</th>
<th>[M-H] (m/z)</th>
<th>MS²</th>
<th>Predicted formula</th>
<th>Putative identity</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>679.3451*</td>
<td>661.3, 559.2, 541.4, <strong>517.3</strong>, 499.3, 455.2, 437.3</td>
<td>C_{36}H_{55}O_{12} (+formate)</td>
<td>Triterpenoid glycoside</td>
</tr>
<tr>
<td></td>
<td>1020.0237**</td>
<td>1358.9, 517.3, 499.3, 455.2</td>
<td>None</td>
<td>Unknown</td>
</tr>
<tr>
<td></td>
<td>1359.7463**</td>
<td>679.3</td>
<td>C_{72}H_{111}O_{24}</td>
<td>Triterpenoid glycoside dimer</td>
</tr>
<tr>
<td>T2</td>
<td>1357.6812**</td>
<td>1313.2, 1151.2, 1018.7, <strong>679.3</strong>, 517.1, 455.2</td>
<td>None</td>
<td>Triterpenoid glycoside dimer (+formate)</td>
</tr>
<tr>
<td></td>
<td>679.3441*</td>
<td>541.3, <strong>517.4</strong>, 499.3, 455.3</td>
<td>C_{36}H_{55}O_{12}</td>
<td>Triterpenoid glycoside</td>
</tr>
<tr>
<td>1</td>
<td>191.0130</td>
<td><strong>172.9</strong>, 111.01</td>
<td>C_{6}H_{7}O_{7}</td>
<td>Citric acid</td>
</tr>
<tr>
<td>2</td>
<td>575.0983</td>
<td>557.1, 449.2, <strong>423.1</strong>, <strong>407.0</strong>, 289.2</td>
<td>C_{30}H_{25}O_{12}</td>
<td>A-type EC dimer</td>
</tr>
<tr>
<td>3</td>
<td>577.1143</td>
<td>559.3, 451.2, <strong>425.2</strong>, 407.2, 289.2</td>
<td>C_{30}H_{25}O_{12}</td>
<td>B-type EC dimer</td>
</tr>
<tr>
<td>4</td>
<td>121.0252</td>
<td>93.1</td>
<td>C_{7}H_{6}O_{2}</td>
<td>Benzoic acid</td>
</tr>
<tr>
<td>5</td>
<td>289.0612</td>
<td><strong>245.1</strong>, 205.1, 179.1</td>
<td>C_{15}H_{13}O_{6}</td>
<td>EC (+formate)</td>
</tr>
<tr>
<td></td>
<td>(335.652)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>561.1202</td>
<td>543.1, 435.1, <strong>407.0</strong>, 289.1</td>
<td>C_{30}H_{25}O_{11}</td>
<td>EfEC dimer (+formate )</td>
</tr>
<tr>
<td></td>
<td>(607.1239)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>849.1729</td>
<td>723.3, <strong>561.1</strong>, 407.2, 289.1</td>
<td>C_{45}H_{57}O_{17}</td>
<td>Ef EC EC trimer</td>
</tr>
<tr>
<td>8</td>
<td>934.0376*</td>
<td>1566.9, 1234.9, 897.0, 633.1, 301.3</td>
<td>None</td>
<td>Sanguin H-6</td>
</tr>
<tr>
<td>9</td>
<td>833.1788</td>
<td>815.1, 707.2, <strong>561.0</strong>, 289.2</td>
<td>C_{45}H_{57}O_{16}</td>
<td>EfEfEC trimer</td>
</tr>
<tr>
<td>10</td>
<td>833.1788</td>
<td>815.0, 707.2, <strong>561.0</strong>, 289.1</td>
<td>C_{45}H_{57}O_{16}</td>
<td>EfEfEC trimer</td>
</tr>
<tr>
<td>11</td>
<td>1105.2369</td>
<td>979.1, 951.1, <strong>833.1</strong>, 815.1, 707.3, 561.1, 543.1</td>
<td>C_{50}H_{69}O_{21}</td>
<td>EfEfEfEC tetramer</td>
</tr>
<tr>
<td>12</td>
<td>147.0400</td>
<td><strong>118.9</strong>, 84.9</td>
<td>C_{10}H_{16}O_{2}</td>
<td>Cinnamic acid</td>
</tr>
<tr>
<td></td>
<td>Mass M/z</td>
<td>Formula</td>
<td>Structure</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---------</td>
<td>---------</td>
<td>----------------------------</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>709.3553</td>
<td>663.2, <strong>501.3</strong></td>
<td>C_{37}H_{57}O_{13} Formate adduct of triterpenoid glycoside [60]</td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>709.3553</td>
<td>663.2, <strong>501.3</strong></td>
<td>C_{37}H_{57}O_{13} Formate adduct of triterpenoid glycoside [60]</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>725.3490</td>
<td><strong>679.3</strong>, 517.3</td>
<td>C_{37}H_{57}O_{14} Formate adduct of triterpenoid glycoside</td>
<td></td>
</tr>
<tr>
<td>B</td>
<td>741.3428</td>
<td><strong>695.3</strong>, 533.2</td>
<td>C_{37}H_{57}O_{15} Formate adduct of unknown triterpenoid glycoside</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>493.2106</td>
<td>447.3, <strong>315.1</strong></td>
<td>C_{22}H_{37}O_{12} Unknown</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>327.2056</td>
<td>309.3, 291.3, 229.2, <strong>171.1</strong></td>
<td>C_{18}H_{31}O_{5} Trihydroxyoctadecadiene derivative (CS)</td>
<td></td>
</tr>
</tbody>
</table>

All predicted formula derived with $< 2 \Delta$ ppm mass accuracy data; *[M-H]^2- ion, has 0.50 amu variants, **[M-H]^3- and [M-H]^3-ions, has both 0.50 and 0.33 amu variants. MS^2 fragments in bold italics = most abundant signals

Shaded entries are the main triterpenoid peaks or are peaks enriched during purification, Underlined = major MS^2 fragments. EF = epiafzelechin; EC = epicatechin.
Fig. 1

Sanguin H6; m/z 934, RT = 17.9

m/z 679, RT = 23.5
$m/z$ 355 main peak, RT = 12.7
Fig. 2

A

B

C

D

FSD: 1.32e6

FSD: 1.87e9

FSD: 5.45e8

FSD: 6.89e8
Fig. 3
Fig. 4
Fig. 5

A

B

FSD = 4.53e6

FSD = 8.31e5
Fig. 6

**HT29 Cells**

<table>
<thead>
<tr>
<th></th>
<th>Untreated</th>
<th>100nM CDDO</th>
<th>100nM TRF</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Tail DNA</td>
<td>50</td>
<td>*20</td>
<td>*20</td>
</tr>
<tr>
<td>Treatment + 75μM H₂O₂</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**CCD841 Cells**

<table>
<thead>
<tr>
<th></th>
<th>Untreated</th>
<th>100nM CDDO</th>
<th>100nM TRF</th>
</tr>
</thead>
<tbody>
<tr>
<td>% Tail DNA</td>
<td>50</td>
<td>*20</td>
<td>*20</td>
</tr>
<tr>
<td>Treatment + 25μM H₂O₂</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Fig. 7

Expression change after treatment with 100nM CDDO

- **Nrf2**
  - HT29
  - CCD 841

- **NQO1**
  - HT29
  - CCD 841

- **HO-1**
  - HT29
  - CCD 841

Expression change after treatment with 100nM TRF

- **Nrf2**
  - HT29
  - CCD 841

- **NQO1**
  - HT29
  - CCD 841

- **HO-1**
  - HT29
  - CCD 841

**ORIGINAL FIGURE 7 (below)**

Expression change after treatment with 100nM CDDO

- **Nrf2**
  - HT29
  - CCD 841

- **NQO1**
  - HT29
  - CCD 841

- **HO-1**
  - HT29
  - CCD 841

Expression change after treatment with 100nM TRF

- **Nrf2**
  - HT29
  - CCD 841

- **NQO1**
  - HT29
  - CCD 841

- **HO-1**
  - HT29
  - CCD 841