

Application Note # ET-19

Metabolic profiling of tea extracts by high-resolution LC in combination with maXis UHR-TOF MS analysis

In this study, we performed untargeted metabolic profiling of different black and green tea extracts by maXis™ ultrahigh-resolution Q-TOF mass spectrometry coupled to a fast UHPLC separation. All relevant information was extracted from the raw data by applying the “Find Molecular Features” (FMF) algorithm. Processed data was submitted to principal component analysis (PCA) in order to differentiate sample groups and identify characteristic compounds of various tea types. Sum formulae were calculated with SmartFormula™ for discriminating metabolites taking accurate mass and isotopic pattern information into account. The number of possible sum formulae in a certain mass accuracy window increases exponentially with higher molecular masses. Therefore, autoMS/MS runs focusing on those analytes accounting for the largest differences between teas as precursor ions were performed. Accurate MS/MS data enabled a reduction of sum formula suggestions by intelligently combining sum formulae and neutral loss information using SmartFormula 3D™. A query of sum formulae in public databases enabled a tentative identification of several compounds characteristic of a green tea extract, which was clearly distinguishable from all other tea samples. Several catechins, identified by comparison to reference standards, differentiated black and green tea samples.

Introduction

Black and green tea account for more than 95% of tea consumed worldwide. Health benefits have been hypothesized for both types of tea. A deeper understanding of potential health promoting effects – as well as an improvement in quality and taste – is a hot topic in academia and food industry. In this study, we used statistical analysis of high resolution ESI mass spectrometry data to study the composition of different teas. High throughput, typically of very complex samples, is a common challenge encountered in metabolomics experiments. Modern UHPLC separations – which enable analysis time to be reduced without sacrificing resolving power required to separate the complex mixtures of compounds present (e.g. in tea extracts) – present an answer to this challenge. Since mass accuracy and resolution of high-resolution TOF instruments are independent of the acquisition rate, such instruments are perfectly suited for a coupling to UHPLC separations. The accurate mass data and isotopic pattern information in MS and MS/MS spectra generated by maXis UHR-TOF analysis permits a reliable sum formula generation. Combined with database queries, this facilitates the identification of unknown compounds — often described as the major bottleneck in metabolomics.

PCA Scores and loadings plot

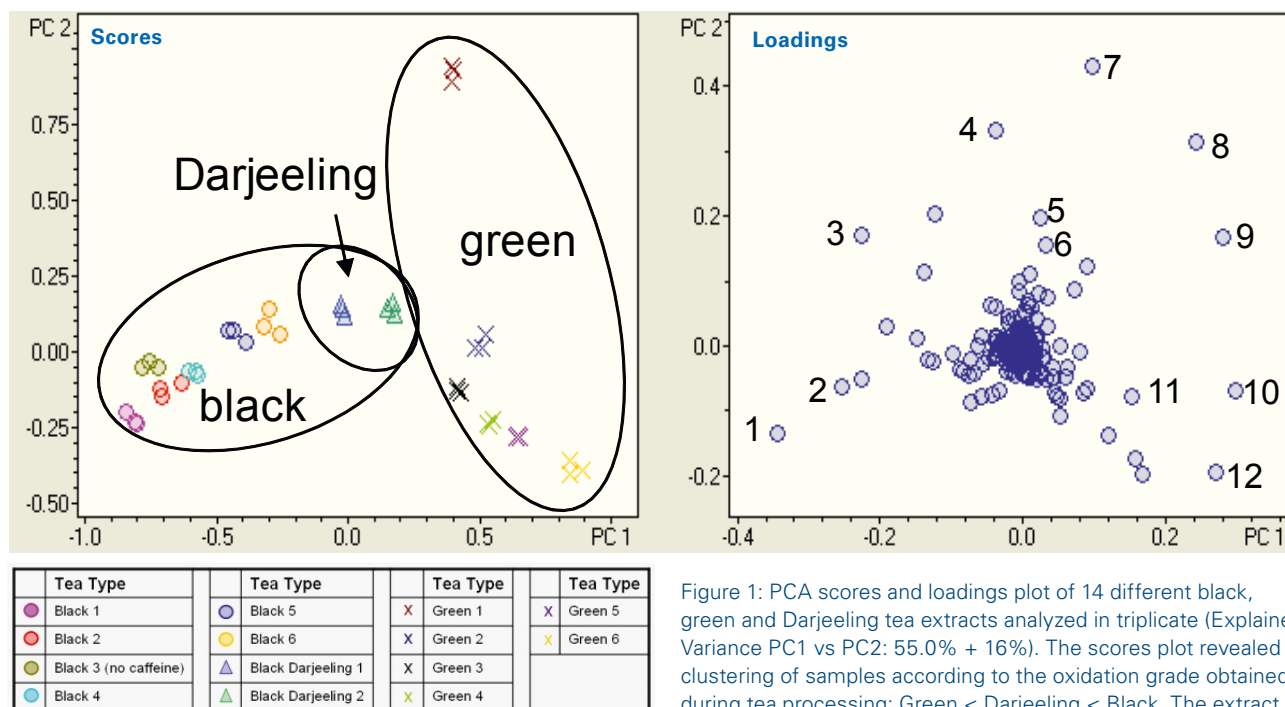


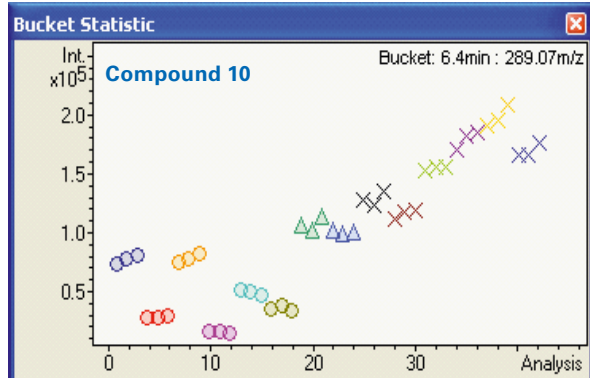
Figure 1: PCA scores and loadings plot of 14 different black, green and Darjeeling tea extracts analyzed in triplicate (Explained Variance PC1 vs PC2: 55.0% + 16%). The scores plot revealed a clustering of samples according to the oxidation grade obtained during tea processing: Green < Darjeeling < Black. The extract of "Green Tea 1" is clearly separated in the scores plot from all other green tea samples. Selected loadings contributing to the difference between the tea extracts are highlighted.

Table 1: Selected loadings differentiating the tea samples (highlighted in Fig.1.). Identities of compounds 8-12 were confirmed by comparison to purified standards.

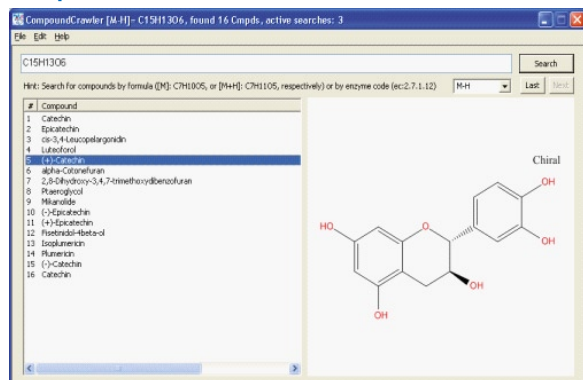
Loadings (see Fig. 1)	Retention time [min]	m/z measured	Error (ppm)	mSigma	Sum formula [M-H] ⁻	Possible compound
1	0.8	191.0560	0.51	6.7	C ₇ H ₁₁ O ₆	Quinic acid
2	6.7	337.0926	0.70	10.4	C ₁₆ H ₁₇ O ₈	Coumaroyl-quinic acid
3	2.4	169.014	-0.59	6.6	C ₇ H ₅ O ₅	Gallic acid
4	2.8	343.067	0.08	7.9	C ₁₄ H ₁₅ O ₁₀	Theogallin
5	6.9	635.0889	0.13	5.4	C ₂₇ H ₂₃ O ₁₈	Trigalloyl-glucose
6	5.3	483.0777	-0.04	6.8	C ₂₀ H ₁₉ O ₁₄	Digalloyl-glucose
7	4.9	633.0727	0.35	13.7	C ₂₇ H ₂₁ O ₁₈	Corilagin
8	7.0	441.0828	-0.22	8.4	C ₂₂ H ₁₇ O ₁₀	Epicatechin-3-gallate
9	6.0	457.0776	0.07	4.0	C ₂₂ H ₁₇ O ₁₁	Epigallocatechin gallate
10	6.4	289.0718	-0.24	6.1	C ₁₅ H ₁₃ O ₆	Epicatechin
11	5.0	305.0665	0.58	2.4	C ₁₅ H ₁₃ O ₇	Epigallocatechin
12	3.5	305.0666	0.16	5.5	C ₁₅ H ₁₃ O ₇	Gallocatechin

Identification of epicatechin by SmartFormula and database query

Bucket Statistics



CompoundCrawler



Smart Formula

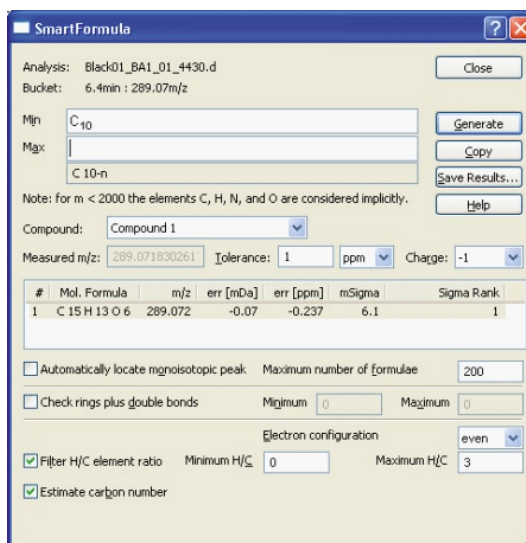


Figure 2: Identification of compound 10 (see PCA loadings plot in Fig.1 & Table 1). Bucket Statistics plot for selected loading (6.4min; 289.07m/z) displays the intensity for all analyses. Sum formula generation by SmartFormula followed by query of public databases highlights catechin and epicatechin as likely structures for this compound.

Experimental

Extracts of 14 different teas (6 black, 6 green and 2 black Darjeeling) were prepared with 100 mL hot water for 5 minutes. 2 μ L undiluted extracts were analyzed in triplicate using a Dionex UltiMate 3000TM Rapid Separation LC-System with an Acclaim C18 column (2.1x100 mm, 2.2 μ m particles) interfaced to a maXis ESI-UHR-Q-TOF-MS. Full scan MS and autoMS/MS data were acquired in ESI negative mode with an acquisition rate of 4 Hz. The following gradient was applied for separation with a flow rate of 400 μ L/min using (A) Water + 0.1% HCOOH (B) Acetonitrile + 0.1% HCOOH as mobile phase. Gradient: 0 min 1% B; 0.5 min 1% B; 6 min 30% B; 6.5 min 98% B; 7.5 min 98% B; 7.6 min 1% B; 10 min 1% B. The datasets were automatically recalibrated and all relevant features were extracted in the Compass DataAnalysisTM 4.0 processing software. A Principal Component Analysis (PCA) of the pre-processed data as well as sum formula generation by SmartFormula were performed in ProfileAnalysisTM 2.0.

Results and Discussion

PCA differentiation of tea extracts according to oxidation grades obtained during production

All relevant information was extracted from the LC-MS data of the different tea samples by applying the "Find Molecular Features" peak detection algorithm. This algorithm efficiently differentiates real signal from background noise. The PCA scores plot (Fig. 1), based on the extracted information, revealed a separation according to the oxidation grade obtained during tea processing: Green < Darjeeling < Black. Although Darjeeling tea is commonly sold as "black tea" it belongs to the group of "oolong teas" for which an incomplete oxidation is obtained during production [1]. Interestingly, the extract of "Green Tea 1" was clearly separated in the scores plot from all other green tea samples.

A decaffeinated black tea (Black 3) clustered close to all other black tea samples indicating that the process of decaffeination did not alter the metabolite composition extensively. Caffeine did not differentiate the samples, since the data was acquired in ESI negative mode where caffeine did not ionize.

Several loadings representing FMF compounds which mainly contributed to the difference between the tea types are highlighted in Fig.1. The sum formulae for these compounds were calculated using SmartFormula. Formulae were ranked according to the fit between measured and theoretical isotopic pattern expressed in the mSigma-Value. The smaller the mSigma-Value the better is the isotopic fit. Elemental compositions given in Table 1 all represent the first Sigma rank within a 1 ppm mass accuracy window. A query of the sum formulae in public databases using the CompoundCrawler™ enabled a tentative identification of these compounds. The identity of epicatechin, gallocatechin, epigallocatechin, epicatechin-3-gallate and epigallocatechin-gallate could be confirmed by comparison to reference compounds (see compounds 8-12 in Table 1).

Identification of epicatechin from green tea

The bucket statistics plot in Fig. 2 displays the peak intensities for loading 10 (see Fig.1) for all samples. This compound with a retention time of 6.4 min and 289.07 m/z differentiated green and black tea samples, with higher intensities being observed in green tea. The two Darjeeling tea extracts exhibited medium intensities.

Sum formula generation based on exact mass (289.0718 m/z) and isotopic pattern information using SmartFormula provided $C_{15}H_{13}O_6$ as a result for $[M-H]^-$. A structural assignment based on a query in public databases using the CompoundCrawler™ returned epicatechin and catechin as likely candidates. Based on a comparison to the retention time of a reference standard, the target compound was identified as epicatechin (Fig. 3). Catechin derivatives are known to be more abundant in green tea than black tea [1].

A Base Peak Chromatogram (BPC) of “Green Tea 1” extract revealed the complexity of the analysed samples (Fig. 3). High-resolution EIC traces (hrEIC), i.e. extracted ion chromatograms with very narrow mass windows of a few mDa, provide a high selectivity even in complex samples. An example is shown for the mass trace of epicatechin (289.0718 +/- 1 mDa).

Tentative identification of trigalloyl-glucose as a characteristic compound for “Green Tea 1”

Loading 5 was among the compounds being responsible for the separation of “Green Tea 1” from all other samples (Fig. 1). Sum formula generation based on the MS data provided 3 possible candidates within 1 ppm mass accuracy (Fig. 4). SmartFormula 3D combining exact mass and isotopic pattern in both MS and MS/MS spectra as well as neutral loss information [2] yielded only one hit: $C_{27}H_{23}O_{18}$. This tentative identification is in accordance with the assigned neutral losses within the MS/MS spectrum. Neutral loss assignments based on maXis data are highly trustworthy as demonstrated by the sub-mDa mass deviations shown in Fig. 4. Further assurance is readily provided by the automatically calculated SmartFormula 3D spectrum including annotated fragment ions (Fig. 4 bottom). The database query for the confirmed sum formula suggested trigalloyl-glucose as a likely structure. The structural candidate is in good agreement to the observed cleavages of ester and neighboring bonds of the gallic acid glucose conjugate. Trigalloyl-glucose has been previously identified in black tea [3].

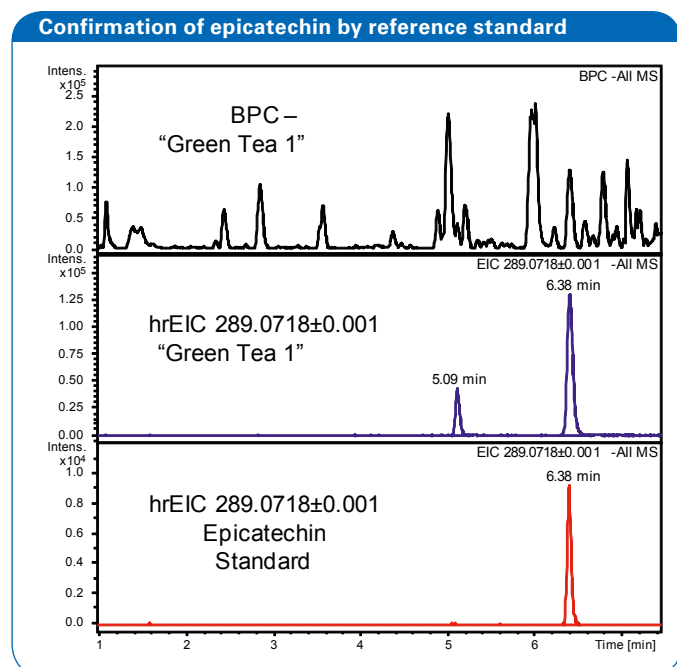


Figure 3: Base Peak Chromatogram from tea extract “Green Tea 1” (top). Mass trace of $m/z=289.0718$ (corresponding to $[M-H]^-$ of epicatechin: $C_{15}H_{13}O_6$) with a trace width of 1 mDa are shown for green tea extract (middle) and epicatechin standard (bottom). The peak at RT 5.09 in the hrEIC of the green tea extract corresponds to catechin.

Towards the Identification of Trigalloyl-Glucose using SmartFormula 3D

SmartFormula

Measured m/z: 635.08890 Tolerance: 1 ppm Charge: -1

#	Mol. Formula	m/z	err [mDa]	err [ppm]	mSigma	Sigma Rank
1	C ₂₇ H ₂₃ O ₁₈	635.089	0.08	0.125	5.4	1
2	C ₂₅ H ₁₁ N ₁₄ O ₈	635.089	0.07	0.109	12.2	2
3	C ₄₀ H ₁₅ N ₂ O ₇	635.088	-0.43	-0.683	66.8	3

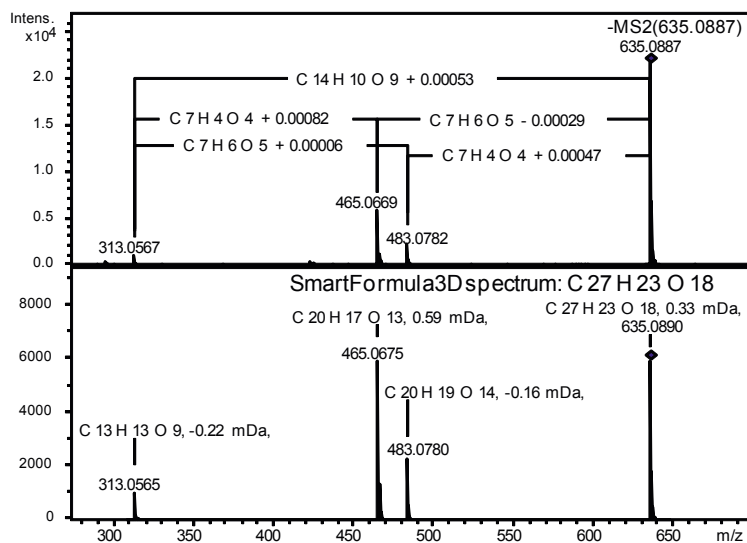
SmartFormula 3D

SumFormula	m/z calc	err[m...]	err[ppm]	mSigma	eConf	Comb.	SumFormula	SumFormula L...	m/z Loss	err[mDa...]	m/z calc	err[mDa]
<input checked="" type="checkbox"/> C ₂₇ H ₂₃ O ₁₈	635.0890	0.3	0.5	13.2	even	3	<input type="checkbox"/> C ₂₀ H ₁₉ O ₁₄	C ₇ H ₄ O ₄	152.0105	0.5	483.0780	-0.2
							<input type="checkbox"/> C ₂₀ H ₁₇ O ₁₃	C ₇ H ₆ O ₅	170.0218	-0.3	465.0675	0.6
							<input type="checkbox"/> C ₁₃ H ₁₃ O ₉	C ₁₄ H ₁₀ O ₉	322.0319	0.6	313.0565	-0.2

One single sum formula of precursor remains

Sum formulae of corresponding fragments

Measured MS/MS spectrum with neutral losses and their mass deviation in Da (top) Simulated SmartFormula 3D spectrum with annotated sum formulae (bottom)



Structural Hypothesis for C₂₇H₂₃O₁₈ Trigalloyl-glucose

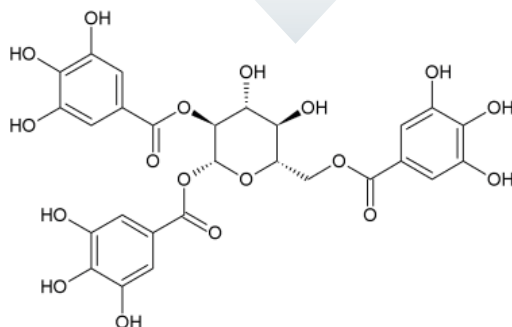


Fig.4: SmartFormula and Smart-Formula 3D result for Compound 5 (6.9 min; 635.089 m/z). Tentative identification as Trigalloyl-glucose based on neutral loss and fragment sum formula information annotated in MS/MS spectrum.

Conclusion

A fast UHPLC separation coupled to maXis UHR-TOF-MS analysis enabled the comprehensive analysis of complex mixtures of small molecules, namely black and green tea samples. PCA analysis allowed for the differentiation of tea samples according to the oxidation grade obtained during tea processing.

The sub-ppm mass accuracy and the True Isotopic Pattern™ information delivered by the maXis instrument facilitated the straightforward sum formula generation of compounds characteristic for different types of tea. By combining accurate mass and isotopic pattern information in MS and MS/MS spectra, SmartFormula 3D extends the capability for unique sum formula generation to compounds with higher molecular weight.

In combination with database queries using the CompoundCrawler, the feasibility of this approach for the identification of unknown compounds in metabolomics studies could be demonstrated.

References

- [1] Graham, H.N.: Green tea composition, consumption, and polyphenol chemistry, *Prev. Med.* 1992, 21(3):334-50.
- [2] Bruker Daltonics; Technical Note #TN-23: Certainty in Small Molecule Identification by Applying SmartFormula3D on a UHR-TOF Mass Spectrometer.
- [3] James Warren Drynan, Michael N. Clifford, Jacek Obuchowicz, Nikolai Kuhnert: The chemistry of low molecular weight black tea polyphenols, *Nat. Prod. Rep.*, 2010, 27, 417-462.

Authors

Aiko Barsch, Gabriela Zurek, Wiebke Lohmann,
Bruker Daltonik GmbH, Bremen, Germany

Keywords

Metabolomics
Metabolic profiling
Principle Component Analysis
Small molecule identification
Catechin derivatives
Green Tea
Black Tea

Instrumentation & Software

maXis UHR-TOF
Dionex UltiMate 3000 RSLC
ProfileAnalysis 2.0
SmartFormula3D

For research use only. Not for use in diagnostic procedures.

www.bdal.com

● **Bruker Daltonik GmbH**

Bremen · Germany
Phone +49 (0)421-2205-0
Fax +49 (0)421-2205-103
sales@bdal.de

Bruker Daltonics Inc.

Billerica, MA · USA
Phone +1 (978) 663-3660
Fax +1 (978) 667-5993
ms-sales@bdal.com